OECD Education Working Papers No. 117

The Efficiency of Secondary Schools in an International Perspective: Preliminary Results from PISA 2012

Tommaso Agasisti,

Pablo Zoido

https://dx.doi.org/10.1787/5js1t53mrdf0-en

OECD

Unclassified **EDU/WKP(2015)6**

Organisation de Coopération et de Développement Économiques
Organisation for Economic Co-operation and Development **05-May-2015**

English - Or. English

DIRECTORATE FOR EDUCATION AND SKILLS

Cancels & replaces the same document of 23 April 2015

THE EFFICIENCY OF SECONDARY SCHOOLS IN AN INTERNATIONAL PERSPECTIVE: PRELIMINARY RESULTS FROM PISA 2012

Education Working Paper No. 117

by Tommaso Agasisti and Pablo Zoido

Tommaso Agasisti, Politecnico di Milano School of Management, Italy
(e.tommaso.agasisti@polimi.it)
Pablo Zoido, Analyst, EDU/ECS (Pablo.Zoido@oecd.org)

JT03375588

Complete document available on OLIS in its original format

OECD EDUCATION WORKING PAPERS SERIES

Comment on the series is welcome, and should be sent to edu.contact@oecd.org.

This working paper has been authorised by Andreas Schleicher, Director of the Directorate for Education and Skills, OECD.

----

www.oecd.org/edu/workingpapers

TABLE OF CONTENTS

Tables

Figures

THE EFFICIENCY OF SECONDARY SCHOOLS IN AN INTERNATIONAL PERSPECTIVE: PRELIMINARY RESULTS FROM PISA 2012

Tommaso Agasisti*
Politecnico di Milano School of Management
* Corresponding author
Department of Management, Economics and Industrial Engineering
OECD Thomas J. Alexander Fellow
e. tommaso.agasisti@polimi.it

Pablo Zoido
OECD, Organisation for Economic Co-operation and Development

Keywords: Efficiency, equity, international comparison, bootstrap Data Envelopment Analysis
JEL Codes: I21, I24, I28, C14

ABSTRACT

As governments around the world struggle with doing more with less, efficiency analysis climbs to the top of the policy agenda. This paper derives efficiency measures for more than 8,600 schools in 30 countries, using PISA 2012 data and a bootstrap version of Data Envelopment Analysis as a method. We estimate that given current levels of inputs it would be possible to increase achievement by as much as 27% if schools improved the way they use these resources and realised efficiency gains. We find that efficiency scores vary considerably both between and within countries. Subsequently, through a second-stage regression, a number of school-level factors are found to be correlated with efficiency scores, and indicate potential directions for improving educational results. We find that many efficiency-enhancing factors vary across countries, but our analysis suggests that targeting the proportion of students below low proficiency levels and putting attention to students' good attitudes (for instance, lower truancy), as well as having better quality of resources (i.e. teachers and educational facilities), foster better results in most contexts.

RÉSUMÉ

Alors que les gouvernements du monde entier tentent de faire toujours plus avec moins, l'analyse de l'efficience occupe le haut de l'agenda politique. Ce document s'appuie sur des mesures d'efficience effectuées dans plus de 8600 écoles dans 30 pays, en utilisant les données PISA de 2012 et une version bootstrap d'une méthode d'analyse par enveloppement de données. Nous estimons qu'au regard des niveaux actuels des contributions, il serait possible d'augmenter les performances de 27% si les écoles amélioraient la façon dont elles utilisent les ressources en réalisant des gains d'efficience. Nous constatons que les scores d'efficience varient de manière considérable entre les pays et au sein des pays. En conséquence, par le biais d'une régression de deuxième étape, il se trouve qu'un certain nombre de facteurs scolaires sont corrélés aux scores d'efficience et indiquent de possibles orientations visant à améliorer les résultats en matière éducative. Nous constatons que de nombreux facteurs favorisant l'efficience varient d'un pays à l'autre, mais notre analyse indique que l'on obtient de meilleurs résultats dans la plupart des domaines en se concentrant sur les étudiants dont les compétences sont faibles et en mettant l'accent sur les bonnes attitudes (réduire l'absentéisme par exemple) tout en ayant des ressources de meilleure qualité (professeurs et établissements scolaires).

ACKNOWLEDGMENTS

A previous version of this paper has been presented at the Workshop on "Efficiency in Education" in London on 19-20 September, 2014 and at the OECD EDU Forum in Paris on 8 October 2014. We are grateful for the useful comments received in that occasion. Tommaso Agasisti acknowledges funding from the OECD Thomas J. Alexander fellowship program for carrying out this work. The work should not be reported as representing the official views of the OECD or of its member countries. The opinions expressed and arguments employed herein are those of the author(s).

Introduction

Analysing the efficiency of educational systems and organisations is today at the forefront of the policy and academic debate. Various factors make efficiency more important than ever: declining public budgets, raising competition across public services for limited public expenditures, increasing demand for transparency in information about the costs and results of schools' activities. In this setting, providing clear quantitative information about the efficiency of educational institutions has become more important than ever. This information can contribute to a better understanding of the school factors that are associated with better results and fostering improvements in school outcomes even in the face of declining resources.

Most research and analysis of schools performance and efficiency builds on within-country comparisons of schools (see Johnes, 2004 and Worthington, 2001 for early reviews). A single-country focus has many advantages. For one, the institutional setting can be considered fairly constant and therefore schools share many of the features that characterise them as organisations that transform inputs into outputs. Comparable data of schools operating in different countries is also quite limited. However, the development of programs measuring the achievement of students and schools through standardised tests in several countries offers the opportunity to take an international perspective when conducting efficiency research.

Focusing on student performance, a well-developed, relatively recent stream of literature uses a microeconomic setting for estimating international educational production functions (EPFs). This strand of literature conceptualises several characteristics of students, schools and countries' educational systems as inputs that can explain differences in standardised tests, the outputs (see early work by Wößmann, 2003; 2007, and a wide discussion in Hanushek & Wößmann, 2010). While this research investigates how inputs are statistically related with outputs, it does not analyse efficiency in a technical sense – i.e. the capacity of educational institutions of transforming inputs (resources) into outputs (achievement).

An approach that keeps schools as the objects of study is more appropriate when analysing the efficiency of educational institutions In fact, the specification of an education production function assumes efficiency in production, while the theoretical arguments (Levin et al., 1974) as well as the empirical evidence (Johnes, 2004) suggest that schools tend to incorporate some inefficiency in operation. In addition, studies employing an education production function approach do not consider the school as a unit of analysis (i.e. they do not measure the organisational performance); rather they focus on students instead. A focus on schools is most appropriate when the focus moves towards understanding which organisational factors can help schools to be more effective or efficient.

The literature about educational efficiency in an international setting has generated two parallel streams. The first strand focuses on countries as units of analysis. Clements (2002) used a simple setting where two inputs are indicated as proxies for human and financial resources as inputs (student-teacher ratio and expenditure per student, respectively), and TIMSS (Trends in International Mathematics and Science Study) scores are outputs. This research finds that some countries can achieve their present levels of achievement scores with 25% fewer resources. Afonso and St. Aubyn (2006) consider a country's average PISA score as output, and various measures of resources (i.e. time spent at school and the student-teacher ratio) as inputs. The empirical results of a two-stages Data Envelopment Analysis (hereafter, DEA) where a Tobit regression is used in the second stage, conducted on 25 countries, reveals substantial efficiency gains in public spending on education are common across countries (the authors find that countries can improve results by 11.6% using the same resources). These improvements are related to a country's stock of human capital (measured by attainment of adult population) and wealth (as measures by GDP per capita). Gimenez et al. (2007) use one similar non-parametric model (DEA), but add measures of students' socioeconomic background as inputs. Again, the results help in identifying those countries, where average test score should be higher when taking into consideration available inputs. Agasisti (2014) innovates the state-of-the-art by extending the empirical analysis to two subsequent periods (instead than a single year), considering PISA 2006 and 2009 scores as outputs, and student-teacher ratio and expenditure per student as inputs. Not only DEA efficiency (of spending) scores are derived (they estimate 10% of inefficiency, on average), but also a Malmquist index is computed to analyse if efficiency gains occurred or not, and if they were driven by pure efficiency (movements towards the frontier) or technology improvements (frontier shifts). Overall, this group of contributions agrees in considering that, with the same resources, the countries analysed can obtain higher average achievement scores.

The second strand of literature, still in its infancy, considers schools as the unit of analysis. In this case, the research analyses how schools can be organised or resourced for the purpose of improving educational outcomes. To the best of our knowledge, only Sutherland et al. (2009) use international data for measuring efficiency at school level[1]. Their research comprises the 30 OECD countries, and is based on PISA 2003 as outputs (test scores); the student-teacher ratio and the school-average socio-economic status (an index for the Economic, Social and Cultural Status of students) are used as inputs. They use a Stochastic Frontier (SF) model, instead of the non-parametric techniques widely used in the other studies. The results demonstrate that inefficiency is indeed present in several countries, although it is more relevant in others. The authors argue that differences in school efficiency also exist within countries, albeit these are quite limited in general (around 10% of output expansion), and input savings are substantial (more than 15%) in a number of countries. An econometric model about the determinants of inefficiency reveal that "(…) indicators for higher quality teaching resources appear to be correlated with better performances at the school level" (p. 24), and this is especially true in making schools able to use less resources in obtaining a given level of achievement scores.

This paper extends contributes to this second line of research, and compares the efficiency scores of more than 8 600 schools in 30 countries using PISA 2012 data. It tackles the following key research questions:

1. How relevant are the differences in the efficiency of schools across the selected 30 countries? (a_1) Are these differences driven by between-schools or between-countries variance?

2. Which factors are associated with schools' efficiency scores? (b_1) Are these factors common across all countries?

3. Is there a trade-off between efficiency and equity at school level?

This paper makes at least three important contributions to efficiency research. First, it illustrates the efficiency of schools from an international perspective using the most up-to-date data available; while the only other attempt in this direction dates back to PISA 2003 (Sutherland, et al., 2009) the results included here provide an updated picture about the frontier of efficient schools in 30 countries. Second, we explicitly compare efficiency scores and measures of equity at the school level, and discuss potential trade-offs (or complementarities) for schools pursuing these two objectives. Third, we investigate if the school factors associated to efficiency scores are similar or structurally different across countries, and if differences in efficiency are more relevant between or within countries.

We find that efficiency scores vary widely both between and within countries. When considering all schools together – so allowing for the existence of an international common benchmark – we find that on average schools can raise their scores by 27%, ranging across countries from 15% for schools in Singapore to more than 33% for those in Slovenia). The picture is quite different when we compare each school with

[1] Sutherland *et al.* (2009) also includes a study at country-level, similar the ones of Afonso and St. Aubyn (2006) and Gimenez *et al.* (2007).

those operating in the same country. In this case, the average improvement in output is at 15% (and ranges from 6% on average among schools in Ireland to 22% for those in Slovenia). We also find that efficiency scores of schools within countries encompass the entire range of the international distribution of efficiency, underlying the fact that country average efficiency scores mask substantial internal variation.

Among the factors that are associated with efficiency, we find that the characteristics of the student intake in each school (i.e. the proportion of females and immigrants, the diversity of socioeconomic background, etc.) explain most of the variation in efficiency across schools; however school-related factors (i.e. practices such as extracurricular activities, principal's leadership style, etc.) also play a role in describing differences in efficiency across schools.

Lastly, we find no evidence of trade-offs between efficiency and equity: in other words, more efficient schools tend to be more inclusive. Efficiency scores are related with higher inclusion, as measured by the percentage of students in the school who score above proficiency Level 2, the baseline level of performance in PISA.

Data

The first choice to be made is about the set of countries to be included in the analysis. We opted for the group of countries whose cumulative expenditure per student is above $ 50 000 PPPs (purchasing power parity), as characterised by OECD (2013a). The choice of focusing on these countries is justified by the opportunity of comparing realities with some similarities in the "intensity of investment" in education. There are 30 countries; specifically, all OECD countries including Singapore and excluding Chile, Greece, Hungary, Mexico and Turkey.

The dataset used for this paper is PISA 2012, where data from student and school questionnaires (with students and school level information, respectively) were merged. The selection of inputs and outputs followed the well-established literature about school efficiency, and was constrained by data availability. More specifically, the indicators included in the efficiency analyses are (I and O denote inputs and outputs, respectively):

- *ESCS* (I_1), the average socio-economic status of students in the school, ESCS stands for the Economic, Social and Cultural Status, and provides a measure of family background that includes several dimensions, namely parents' occupation and education, as well as home possessions. It is built to be internationally comparable, and it is been constrained to have a mean of 0 and a standard deviation of 1.

- *StRatio* (I_2), the (inverse of) the student-teacher ratio. This indicator is a proxy measure of the quantity of human resources employed by each school.

- *Computer_n* (I_3), the number of computers per student at school level; it is a proxy for the quantity of material resources (and facilities) available for the school.

- *pv1math* (O_1), is the average score in mathematics obtained by the students in the school. PISA reports five plausible values for each student as a measure for the test score, with the aim of approximating the true distribution of the latent variable being measured (cognitive skills) – for details about the methodology, see OECD (2012). For the baseline model, the first plausible value is used as output, and the robustness of the results are checked with other values. By construction, plausible values have an international mean of 500 and standard deviation of 100.

- *pv1read* (O_2) as the measure for test scores in reading.

We are aware that the data are collected not to be representative at school level in PISA. For this reason, it can be the case that the “true” average PISA score at school level is slightly different from that calculated here. Thus, we consider our measure a proxy for efficiency more than a precise estimate of the true efficiency of each school.

Schools for which at least one of these indicators was missing are excluded from the sample. Table 1 represents the descriptive statistics for these five variables[2] across the 8,640 schools included in the final sample. On average, the schools in the sample have a student population with socio-economic status close to the OECD average (that is, close to zero), around 0.8 computers per student, and 9.5 students per teacher. The average score in mathematics and reading is close to the OECD average of 496 and 495, respectively. However, the differences between countries' averages are large. Annex 1 reports the mean and standard deviation for each variable in each country. Interestingly, standard deviations of these measures within countries are similar or even higher than that in the entire distribution, which highlights the importance of taking into account the variability within each country when interpreting the results of this analysis.

Table 1. Inputs and outputs, descriptive statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
ESCS	8,640	0.063	0.528	-2.636	1.578
StRatio	8,640	0.105	0.141	0.007	5.988
Computer_n	8,640	0.836	1.250	0	55
pv1math	8,640	496.583	63.736	98.232	782.373
pv1read	8,640	495.393	66.350	163.594	734.684

Table 2 provides the list of variables selected for a second-stage analysis that focuses on the factors that are statistically associated with efficiency scores, including a short definition and description (Panel A) and descriptive statistics (Panel B). The second-stage variables have been classified in three main groups: students' characteristics (other than socio-economic status), general characteristics of the schools, and school's practices and processes.

[2] When building the variable at school level (i.e. school averages) we use student weights W_FSTUWT.

Table 2. Second-stage variables

Panel A. Labels and definitions

Variable's label	Description
School's general characteristics	
Orgen	School's orientation is: GENERAL (not vocational)
isced2	*The school's modal level (grade) is ISCED2*
pv1mathsd	SD of math score (pv1math), within school
pv1_belowprof2	Proportion of students below proficiency level 2, within school
Private	The school is private
clsize_small	The classes in the school have, on average, less than 15 students
Size	School's size (number of students)
Students' characteristics	
immig_1	Proportion of immigrant students (1st generation)
Female	Proportion of female students
hwork_h	Number of hours devoted to homework (set by teachers), per week
Repeater	Proportion of students who were repeater
st_truancy	Proportion of students who reported to have skipped school days
ESCSsd	SD of ESCS, within school
School's practice and processes	
Poor relations	Principal reports that "Learning hindered by poor students/teachers relationships: a lot"
sc_matbui	Index for the principal's perception about the quality of educational infrastructures
prop_cert	Proportion of certified teachers
budget_2	The principal has a major/relevant responsibility in budget allocation
tc_part	Index of teachers' participation to governance
leadership_5	The principal leads or attend in-service activities concerned with instruction: once a month or more
accountability_1	Achievement data are posted publicly (e.g. in the media): YES
qa_ext	There is a quality assurance system that involves external evaluations
eval_teach	Achievement scores used to make judgments about teachers' effectiveness
Volunt	The school organises volunteering activities
select_1	The school uses always at least one factor to select students
Grouping	Ability grouping for some or all classes
Competition	*The school competes with two schools or more for resources*

Panel B. Descriptive statistics

Variable	#	Mean
isced2	8,640	0.504
Orgen	8,640	0.756
Private	8,640	0.192
clsize_small	8,640	0.068
immig_1	8,640	0.063
Female	8,640	0.489
Repeater	8,640	0.147
st_truancy	8,640	0.192
poorrelations	8,640	0.005
prop_cert	8,124	0.922
budget_2	8,640	0.778
leadership_5	8,640	0.241
accountability_1	8,640	0.409
qa_ext	8,640	0.611
eval_teach	8,640	0.457
Volunt	8,640	0.712
select_1	8,640	0.393
Grouping	8,640	0.513
competition	8,640	0.588

Variable	#	Mean	Std. Dev.	Min	Max
pv1mathsd	8,579	72.826	16.532	0.000	199.707
pv1_belowprof2	8,640	0.218	0.226	0.000	100.000
ESCSsd	8,573	0.754	0.174	0.000	2.033
Size	8,462	676.476	475.508	2	4300
hwork_h	8,522	5.297	2.868	0.000	30.000
sc_matbui	8,374	0.030	0.996	-2.755	1.305
tc_part	8,624	0.131	0.926	-1.847	4.027

A set of students' characteristics investigates if efficiency is related to the proportion of female students (*female*), first-generation immigrants (*immig_1*), repeaters, and students who are reported to have skipped school days (*st_truancy*). In this set of variables, we also considered the number of hours devoted to homework set by teachers (*hwork_h*). With the aim of controlling for the diversity of the student population – and somehow peer effects – we computed the standard deviation of socio-economic status within a school (*ESCSsd*). All these indicators can be thought as proxies for additional inputs of the schools; however, for keeping the model of school "(technical) productive process" as simplest as possible, we introduce these in the second stage for seeing if and how they influence the efficiency scores as derived through the baseline specification.

Among the schools' general characteristics, we inserted a dummy (*orgen*) to indicate whether the orientation of the school is academic (compared with technical and vocational), and another dummy for schools where the modal grade across students is lower secondary, ISCED 2 (*isced2*). A complete set of country fixed effects are used, together with a set of dummies for program types as classified in the PISA dataset with the variable *progname*. One dummy (*private*) distinguishes between private schools (both State-dependent and independent) and public ones. Indicators for size and class size are included in the model, the former as the number of students per schools (*size*), the latter through a dummy that identifies if the average dimension of classes within a school is lower than 15 students per class (*clsize_small*). Two indicators are intended to capture dimensions of equity: *pv1mathsd* measures the standard deviation of test scores within school, and can be regarded as an indicator of equality, while *pv1_belowprof2* calculates the proportion of students who obtain a score below the proficiency Level 2, which is indicated by OECD as the baseline level to participate to the modern economic and social life – in this sense, it is a proxy for the concept of inclusion as defined by Schleicher (2014).

Among the set of variables for school's practices and processes, there is a number of indicators taken from the school questionnaire: if learning is hindered a lot by poor relationships between teachers and students (*poorelations*), and index that measures the quality of educational infrastructures at school (*sc_matbui*), a dummy that indicates if the principal has a major responsibility in budget allocation (*budget_2*), an index of teachers' participation to governance (*tc_part*), a variable indicating if the principal exerts instructional leadership by leading or attending meeting together with teachers for discussing about instructional methods and contents (*leadership_5*). Some dummies signal if the school's achievement data are publicly available (*accountability_1*), if is there a quality assurance system that involves external examiners (*qa_ext*), if the achievements scores are used in some way for teachers' evaluation (*eval_teach*), if the school organises volunteering as an extracurricular activity (*volunt*), if the school always uses at least one factor to select students at the entrance (*select_1*), and if the school practices ability grouping between classes systematically (*ability*). Lastly, an index of competition is measured as a dummy if the school reports that it competes with two or more schools for the same students (*competition*).

Coherently with the spirit of this work, all these variables have the objective of measuring school-level (i.e. within-country) variations, not country-level ones. In other words, we are not interested to check if an educational system's level of accountability structurally affects the efficiency of schools that operate there in a general equilibrium perspective, but if different grades of accountability in the same country are associated with different levels of efficiency (technically, the variation across countries is instead captured by the country fixed effects in the second-stage regression). There is substantial within country variation for these indicators. Basic statistics (available upon request from the authors) demonstrate that it is indeed the case –the mean, the standard deviation and the coefficient of variation (sd/mean) of these variables reveal that the standard deviation within countries is as high if not higher than the standard deviation for the entire sample.

Methodology

The baseline methodology used here is the computation of efficiency scores based on Data Envelopment Analysis (for an analytical description of the methodology, see Thanassoulis et al., 2008). DEA allows the inclusion of multi inputs and outputs in the model, and its objective is to calculate the efficiency score of each j-th unit (in this case, school). In the generic case where each school is characterised by a combination of $i(i = 1, \dots, m)$ inputs x_i and $r(r = 1, \dots, s)$ outputs y_r such as $(x_{0i}; y_{0r})$, then the concept of the j-th school's efficiency can be defined as the ratio of (weighted) outputs over the (weighted) inputs:

$$max \left\{ h_0 = \frac{\sum_{r=1}^{s} u_r y_{r0}}{\sum_{i=1}^{m} v_i x_{i0}} \middle| \begin{array}{c} \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \leq 1 \\ j = 1, \dots, n \quad u_r, v_i \geq 0 \end{array} \right\} \qquad (1)$$

Where u_r and v_i are the weights associated to the outputs and inputs, respectively; they are not set by the analyst, but are determined by the computation of the efficiency score h_0 (through linear programming) so that it turns out as "optimal" – in the sense of the highest as possible (for mathematical details about the use of the linear programming technique, see Thanassoulis *et al.* (2008; 260-265). By construction, h_0 denotes an efficient school, while $h_0 < 1$ means that the school is inefficient (and the difference between each school's h_0 and 1 measures its degree of inefficiency)[3]. It is crucial to indicate that the efficiency

[3] Originally, the efficiency score is estimated in the interval $[1, \infty]$ and is successively standardised in the interval $[0,1]$.

score that is used for each school is not the raw one calculated through (1), but the bias-corrected one, after having performed a bootstrap determination of the schools' efficiency scores and their reference confidence intervals, following the procedure suggested by Simar & Wilson (2000).

With the aim of decomposing the dimensions that affect a school efficiency score, we rely on the intuition of Portela & Thanassoulis (2001), who split pupils' efficiency scores between parts related to their own activity and others due to being enrolled in a specific school. Mancebon *et al.* (2012) adopt the same idea for comparing the efficiency of schools belonging to two different subgroups (private and public), by attributing parts of inefficiency to the management of schools and others to the subsector specificities. Here, we separate efficiency attributable to the single school from that attributable to being operating in a given country; for each school j-th we calculate an "overall" efficiency score eff_j (calculated through the calculus of h_0 as discussed in the equation (1) above):

$$eff_{j(overall)} = eff_{j(C)} \times eff_{j(INT)} \qquad (2)$$

Where $eff_{j(INT)}$ is the baseline score obtained using the international frontier as a benchmark, and $eff_{j(C)}$ is the score when each school is compared only with other schools in its country.

Using the (bootstrapped, bias-corrected) efficiency scores as dependent variables, a second-stage Tobit regression has been performed[4], to observe if some variables (see previous section) are correlated with efficiency. We also considered the concerns raised by Simar & Wilson (2007), and we show the results of using the double-bootstrap procedure suggested by them instead of the baseline second-stage regression (specifically, we use the algorithm #2).

For the purpose of exploring the heterogeneity of the relationships between school features and efficiency scores along the latter distribution, we estimate the second-stage regression also in a quantile fashion (Koenker & Hallock, 2001):

$$min_\beta \sum \theta \,|y_i - \beta \bar{X_\iota}| + \sum (1-\theta)\,|y_i - \beta \bar{X}_i| \qquad (3)$$

Where y_i is the bootstrapped, bias-corrected efficiency score as obtained in the previous step, $\bar{X_\iota}$ is the vector of school-level factors that are potentially associated with efficiency, as listed in the section about Data, and θ is the quantile of the efficiency distribution at which estimating the relationship between covariates and efficiency scores (here: 25th, median and 75th).

[4] In the literature, there is still a open debate about the preferable type of regression to be employed in the second-stage (for instance, deciding between OLS and Tobit) – see McDonald (1999).

Results

An international perspective on the efficiency of educational institutions

Table 3 reports the efficiency scores (summarised by country) calculated using an international benchmark, the entire sample of schools across all countries. These scores are the bias-corrected ones. Given that the mean efficiency score (in the international distribution) is 0.73, schools can on average increase their PISA scores by around 27%, holding inputs constant. This result varies substantially across countries, ranging from 15% for Singapore schools (on average), to 33% for Slovenian ones.

Table 3. Efficiency scores: descriptive statistics and decomposition of efficiency scores

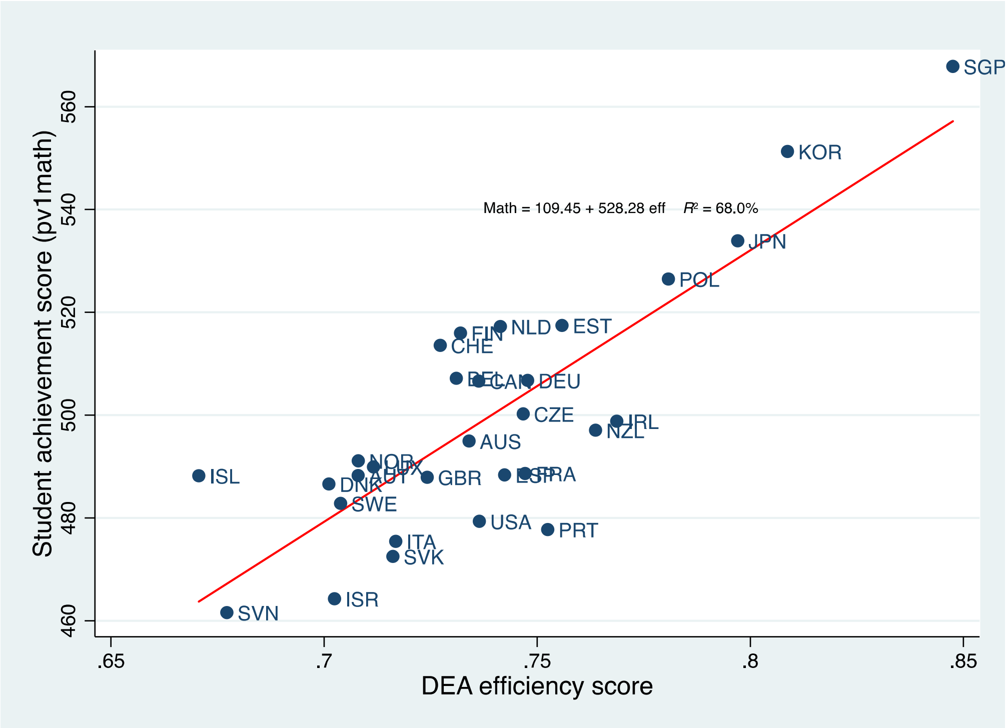
	International frontier					Decomposition of overall efficiency	
Country	**Mean**	**Median**	**s.d.**	**Min**	**Max**	**Eff. scores (country-specific)**	**Eff. scores (overall)**
	(a)	(b)	(c)	(d)	(e)	(f)	(g)
AUS	0.734	0.736	0.070	0.316	0.971	0.791	0.584
AUT	0.708	0.717	0.078	0.472	0.860	0.860	0.613
BEL	0.731	0.744	0.089	0.451	0.909	0.849	0.627
CAN	0.736	0.739	0.064	0.475	0.965	0.842	0.623
CHE	0.727	0.722	0.069	0.526	0.923	0.836	0.612
CZE	0.747	0.752	0.081	0.490	0.916	0.865	0.651
DEU	0.748	0.765	0.086	0.488	0.891	0.876	0.662
DNK	0.701	0.704	0.052	0.520	0.890	0.869	0.611
ESP	0.742	0.746	0.055	0.508	0.918	0.877	0.654
EST	0.756	0.750	0.054	0.614	0.959	0.871	0.661
FIN	0.732	0.734	0.060	0.412	0.924	0.833	0.613
FRA	0.747	0.760	0.091	0.391	0.977	0.860	0.649
GBR	0.724	0.724	0.060	0.506	0.924	0.854	0.621
IRL	0.769	0.775	0.048	0.557	0.876	0.943	0.726
ISL	0.671	0.679	0.064	0.462	0.811	0.873	0.589
ISR	0.702	0.715	0.091	0.392	0.883	0.852	0.606
ITA	0.717	0.730	0.086	0.369	0.903	0.813	0.590
JPN	0.797	0.801	0.085	0.554	0.970	0.867	0.695
KOR	0.809	0.810	0.067	0.574	0.945	0.885	0.720
LUX	0.712	0.707	0.054	0.586	0.815	0.943	0.673
NLD	0.741	0.757	0.096	0.457	0.910	0.876	0.656
NOR	0.708	0.708	0.060	0.438	0.910	0.886	0.630
NZL	0.764	0.767	0.068	0.572	0.952	0.865	0.665
POL	0.781	0.781	0.061	0.470	0.963	0.803	0.630
PRT	0.752	0.753	0.067	0.467	0.931	0.860	0.651
SGP	0.848	0.853	0.059	0.691	0.978	0.882	0.750
SVK	0.716	0.718	0.076	0.516	0.926	0.839	0.605
SVN	0.677	0.677	0.084	0.440	0.887	0.788	0.541
SWE	0.704	0.701	0.067	0.488	0.894	0.857	0.606
USA	0.736	0.737	0.055	0.532	0.878	0.885	0.654
Total	0.734	0.738	0.077	0.316	0.978	0.847	0.626

Notes. Output-oriented, VRS, bootstrap (bias-corrected) efficiency scores. Baseline model (3 inputs; 2 outputs). International frontier as benchmark. Overall efficiency is decomposed as $eff_{j(INT)}*eff_{j(C)}$ – see the section about Methodology.

Figure 1 illustrates the relationship between efficiency and mathematics performance across countries. Figure 2 displays the distribution of efficiency scores by country. These figures make clear one of the main findings in this paper: efficiency scores are much more dispersed within countries than between them. In this sense, comparing the efficiency of educational systems as a whole is less useful exercise than considering efficiency of each school from international perspective. This approach compares each school with others that share similar characteristics but may be located in different countries, and uses them as a reference to compare practices, features and activities – and eventually discusses their efficiency.

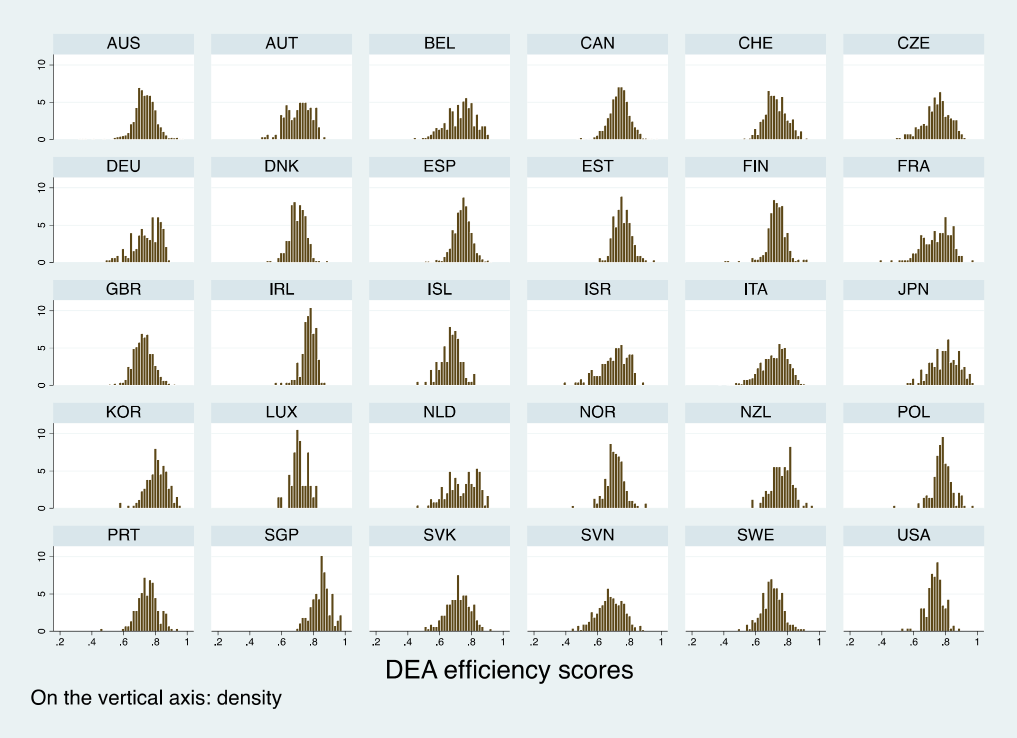
Figure 1 also shows that there is a positive relationship between performance and efficiency. In general, countries where the average PISA score is higher also tend to show better efficiency scores. And yet, some countries with similar average efficiency have very different levels of performance (see the case of Portugal and Estonia, or even strikingly those of the United States and Canada). Other countries have similar levels of average performance but a different average efficiency (see for example the cases of New Zealand and Australia).

Figure 1. Efficiency scores versus performance in mathematics, summarised by country



Notes. The achievement score is based on the first plausible value in mathematics (pv1math), and is the mean of the schools' average scores.

Figure 2. Distribution of efficiency scores, by country



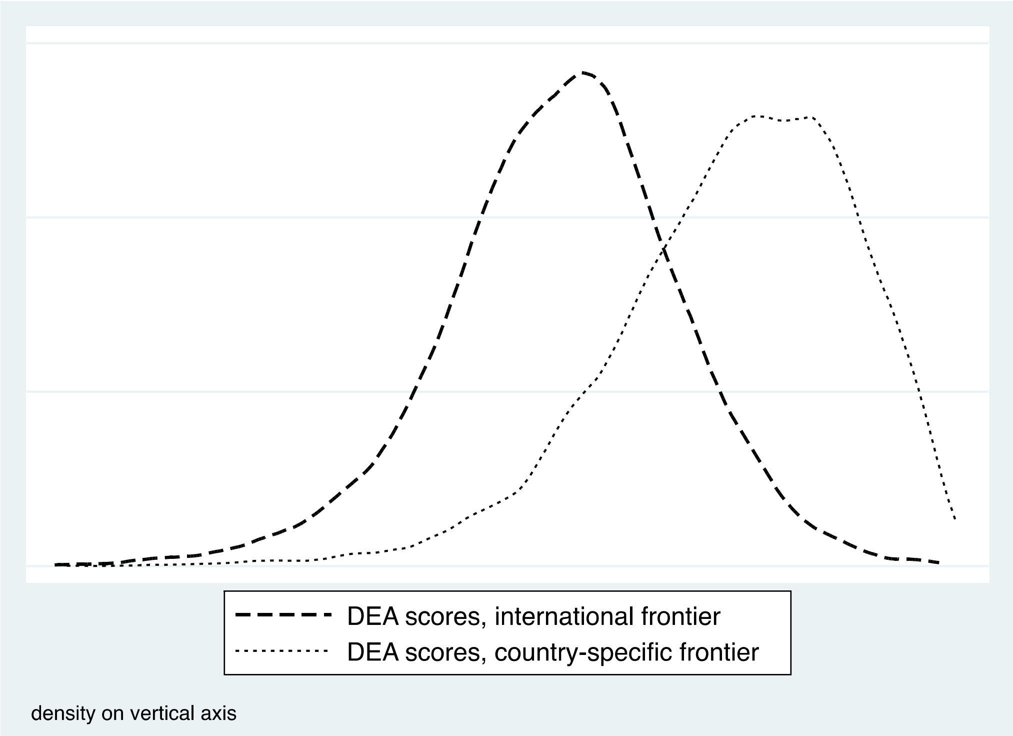
The policy question is whether countries are willing to improve overall performance even if that implies an inefficient amount of resources or instead they are more interested in maximising efficiency at a lower overall level of achievement. Clearly, different strategies at system level can have direct and indirect effects for the incentives and activities of single schools, but this is outside the scope of this paper.

The same positive relationship between performance and efficiency is apparent between schools within countries. However, in this case the correlation is weaker – at international level, pairwise correlation is 0.815, while between schools within countries it is 0.659. Within each country, a large proportion of schools that are efficient, despite their low scores, while others are inefficient even if they reach high achievement levels. A school is defined as "high performance" when the average performance of the students in the school is higher than the median performance of all schools in the sample (497.6). We define a "high efficiency" school as one in which the efficiency score is higher than the mean efficiency score in the entire sample of schools. (0.737). When classified this way, we find that 11% of schools are high performance, low efficiency schools. Meanwhile, 17% of schools are classified as high efficiency, low performance. As with countries, school can benefit from comparing themselves with other educational institutions, in their country or abroad, and try to understand their own position in the efficiency/performance set of possibilities.

We then re-estimated the efficiency scores of schools considering each country-specific frontier separately as a benchmark (Portela & Thanassoulis, 2001). The results are presented in the last column of Table 3**.** They suggest that overall efficiency is determined by a higher frontier when the international benchmark is taken into consideration. When considering only the country-specific frontier, the average efficiency scores are higher because the set of units to be compared with each country is both smaller and more homogeneous.

Figure 3 shows the distribution of efficiency scores is substantially different when considering the country-specific frontier. The gap between the two distributions provides strong evidence for the argument in favour of considering the international frontier as a benchmark, as it widens the options for schools to consider efficient combinations of inputs and outputs. In this sense, the amount of inefficiency that is attributable to being operating in a specific country is more relevant than the inefficiency associated with each school's activity. At the same time, the distribution of efficiency scores within countries is wider than that the differences between countries.

Figure 3. The distribution of efficiency scores, international frontier vs country-specific frontier



In what follows, we only consider the efficiency scores estimated when using the international frontier. This choice implies that the estimated efficiency scores incorporate the role of country-specific features on schools' activities and results. Indeed, in our framework $eff_{j(overall)}$ represents an efficiency measure that takes into account the effect of operating in a concrete educational system. When the objective is to compare real efficiency scores obtained by the schools in an international perspective, the use of $eff_{j(INT)}$ is preferable. The scores obtained under the two different hypotheses are strongly correlated, and to a certain extent the analysis presented in this paper shares many similarities with an analysis using country specific comparisons[5].

The large differences in efficiency scores across schools within countries highlight the importance of the distribution of efficiency, beyond mean efficiency. Therefore, the description of the main results of this paper focuses on the proportion of schools than can be considered efficient and inefficient. Columns (a-b) in Table 4 report for each country the number and proportion of efficient schools. A school is defined as efficient if its estimated efficiency score is statistically significantly above the international mean efficiency. In efficient schools, the estimated confidence interval around their efficiency score does not

[5] Pearson's correlation is >0.81, and Spearman's rank correlation is >0.75. These correlations range from 0.68 (SWE) to 0.97 (ITA) and from 0.60 (LUX) to 0.96 (ITA), respectively.

cross the confidence interval estimated for the international mean. Columns (c-d) contain the number and proportion of schools that are less efficient than average, in each country.

Table 4. The number of efficient schools, by country

Country	N	More efficient than avg n	%	Less efficient than avg n	%	Schools in the 5th percentile of efficiency distribution n	% (intl')
		(a)	(b)	(c)	(d)	(e)	(f)
AUS	718	369	51%	290	40%	18	7%
AUT	178	67	38%	88	49%	0	-
BEL	264	135	51%	104	39%	14	5%
CAN	753	391	52%	264	35%	9	3%
CHE	369	156	42%	173	47%	9	3%
CZE	249	144	58%	79	32%	13	5%
D	194	110	57%	64	33%	2	1%
DNK	283	74	26%	175	62%	1	-
ESP	841	477	57%	246	29%	8	3%
EST	199	128	64%	51	26%	5	2%
FIN	294	145	49%	110	37%	6	2%
FRA	193	111	58%	64	33%	9	3%
GBR	447	187	42%	205	46%	5	2%
IRL	152	125	82%	18	12%	1	0%
ISL	112	14	13%	86	77%	0	-
ISR	141	60	43%	69	49%	0	-
ITA	1,044	488	47%	459	44%	12	5%
JPN	190	143	75%	33	17%	44	17%
KOR	154	131	85%	10	6%	20	8%
LUX	39	12	31%	25	64%	0	-
NLD	143	76	53%	58	41%	7	3%
NOR	177	55	31%	107	60%	2	1%
NZL	149	98	66%	34	23%	5	2%
POL	166	136	82%	18	11%	10	4%
PRT	171	99	58%	40	23%	5	2%
SGP	163	155	95%	3	2%	54	20%
SVK	195	74	38%	94	48%	3	1%
SVN	317	84	26%	205	65%	2	1%
SWE	193	60	31%	112	58%	2	1%
USA	152	73	48%	41	27%	0	-
Total	8,640	4,377	51%	3,325	38%	266	100%

Notes. Average efficiency, as calculated in the international frontier is 1.378 [or 0.725 in the scale (0;1)].

How does the distribution of school efficiency vary across countries? How many schools are efficient? In some countries, the proportion of efficient schools is much higher than inefficient ones. For example, this is the case in Estonia (64% vs 26%), Japan (75% vs 17%), Poland (80% vs 11%), Korea (85% vs 6%) and Singapore (95% vs 2%). In many countries the distribution is more balanced. Australia (51% vs 40%) and the Netherlands (53% vs 41%) are good examples. Lastly, in a few countries, there are many more inefficient schools than efficient ones: Denmark (26% vs 62%), Norway (31% vs 60%), Slovenia (26% vs 65%).

What does school efficiency looks like? Where are the most efficient schools? How frequently are schools from different countries determining the international efficiency frontier? We look into the 5% of schools with the highest efficiency score. To do this we consider the schools that have an efficiency score that is statistically better than the mean efficiency score at 5th percentile (eff_j = 1.170583). In column e of Table 5, the number of these schools by country is reported (total=266); the column (f) calculates how this number by country as a percentage of all of the most efficient schools internationally (that is it computes (e)/266). Schools in Singapore account for 20% of this group and Japan another 17%. Schools located in

European countries account for 35%, Australia for 7%, Canada for 3%. No schools in the United States are present in the 5% of most efficient schools.

Table 5. The number of efficient schools, by country

Country	N	More efficient than avg n	More efficient than avg %	Less efficient than avg n	Less efficient than avg %	Schools in the 5th percentile of efficiency distribution n	Schools in the 5th percentile of efficiency distribution % (intl')
		(a)	(b)	(c)	(d)	(e)	(f)
AUS	718	369	51%	290	40%	18	7%
AUT	178	67	38%	88	49%	0	-
BEL	264	135	51%	104	39%	14	5%
CAN	753	391	52%	264	35%	9	3%
CHE	369	156	42%	173	47%	9	3%
CZE	249	144	58%	79	32%	13	5%
DEU	194	110	57%	64	33%	2	1%
DNK	283	74	26%	175	62%	1	-
ESP	841	477	57%	246	29%	8	3%
EST	199	128	64%	51	26%	5	2%
FIN	294	145	49%	110	37%	6	2%
FRA	193	111	58%	64	33%	9	3%
GBR	447	187	42%	205	46%	5	2%
IRL	152	125	82%	18	12%	1	0%
ISL	112	14	13%	86	77%	0	-
ISR	141	60	43%	69	49%	0	-
ITA	1,044	488	47%	459	44%	12	5%
JPN	190	143	75%	33	17%	44	17%
KOR	154	131	85%	10	6%	20	8%
LUX	39	12	31%	25	64%	0	-
NLD	143	76	53%	58	41%	7	3%
NOR	177	55	31%	107	60%	2	1%
NZL	149	98	66%	34	23%	5	2%
POL	166	136	82%	18	11%	10	4%
PRT	171	99	58%	40	23%	5	2%
SGP	163	155	95%	3	2%	54	20%
SVK	195	74	38%	94	48%	3	1%
SVN	317	84	26%	205	65%	2	1%
SWE	193	60	31%	112	58%	2	1%
USA	152	73	48%	41	27%	0	-
Total	8,640	4,377	51%	3,325	38%	266	100%

Notes. Average efficiency, as calculated in the international frontier is 1.378 [or 0.725 in the scale (0;1)].

Table 5 also highlights that school efficiency comes in many shapes and forms. Schools in the top efficiency group present very different characteristics. The table summarises their inputs and outputs by country. Given the small number of schools in each country, these estimates have to be interpreted cautiously but the differences are very strong. For example, the 18 very efficient Australian schools have relatively socioeconomically advantaged students (they have a high average index of socioeconomic status), but the average test score is very high (higher than 635 in both subjects, approximately the equivalent of two schooling years ahead of the average OECD student). Conversely, the relatively few Spanish and Portuguese top efficiency schools have only modest average scores, but their students come from relatively disadvantaged socioeconomic backgrounds. The scores of the 14 top efficiency Belgian schools are not at the very top of the performance distribution, in particular when taking into account their relatively socioeconomic advantage, but the number of computers per student is very low, so these are likely to be poorly equipped institutions. An efficiency analysis of this kind can be a useful tool for helping schools evaluate their relative position and eventually reflecting strategically about which parts of the efficiency frontier they are aiming at. A school efficiency profile in this sense has the advantage of comparing schools across inputs and outputs and providing valuable information well beyond the average score of students in a standardised test on a particular subject. Table 5b shows similar evidence when considering schools in the 90th percentile, instead of the 95th.

Table 5b. Widening the understanding of the efficient schools' characteristics: focus on very efficient schools when score is in the 90th percentile

Country	#schools 10% most efficient	#schools – total	% schools 10% most efficient	ESCS	StRatio	Computer_n	pv1math	pv1read
	(a)	(b)	(c) =(a)/(b)	(d)	(e)	(f)	(g)	(h)
AUS	36	718	5.0%	0.673	0.078	1.634	602.69	618.90
AUT	2	178	1.1%	0.652	0.090	0.491	589.63	598.54
BEL	28	264	10.6%	0.711	0.082	0.668	626.48	615.37
CAN	36	753	4.8%	0.679	0.061	1.047	581.30	598.86
CHE	21	369	5.7%	0.557	0.091	0.730	628.56	606.16
CZE	32	249	12.9%	0.431	0.087	0.999	602.98	600.59
DEU	23	194	11.9%	0.524	0.068	0.555	599.64	592.31
DNK	2	283	0.7%	0.227	0.070	1.789	553.93	561.38
ESP	29	841	3.4%	-0.202	0.072	0.834	524.51	548.65
EST	9	199	4.5%	0.437	0.113	0.776	602.75	608.77
FIN	7	294	2.4%	0.672	0.124	0.361	648.48	640.10
FRA	31	193	16.1%	0.264	0.081	0.527	571.40	601.43
GBR	12	447	2.7%	0.829	0.063	0.767	592.06	608.70
IRL	6	152	3.9%	0.364	0.075	0.619	532.70	585.07
ISR	1	141	0.7%	-0.570	0.097	0.001	554.92	518.61
ITA	50	1044	4.8%	0.315	0.086	0.480	580.15	595.07
JPN	57	190	30.0%	0.113	0.086	0.342	605.69	609.13
KOR	42	154	27.3%	0.152	0.065	0.350	611.39	587.87
NLD	20	143	14.0%	0.568	0.060	0.737	609.70	606.22
NOR	4	177	2.3%	0.173	0.164	1.747	581.58	626.03
NZL	16	149	10.7%	0.174	0.151	1.175	560.12	601.26
POL	20	166	12.0%	-0.027	0.122	0.407	599.79	603.52
PRT	16	171	9.4%	-0.552	0.089	0.392	529.51	537.76
SGP	93	163	57.1%	-0.209	0.074	0.636	599.37	568.27
SVK	6	195	3.1%	0.533	0.073	0.511	605.59	605.54
SVN	6	317	1.9%	0.063	0.076	1.045	563.77	577.35
SWE	4	193	2.1%	0.864	0.073	0.564	597.48	622.24
USA	2	152	1.3%	-0.262	0.050	1.587	501.62	530.65
Total	611	8640	7.1%	0.248	0.082	0.704	592.14	593.12

Notes: a school is defined in the group of "very efficient" ones when its score is in the 90th percentile.

The factors associated with efficiency

Table 6 presents the results of an analysis of efficiency enhancing factors across all schools. These results need to be interpreted cautiously. In particular, an important caveat to keep in mind is that factors that may be associated with higher efficiency across schools may not be associated with higher efficiency across school systems. Another caveat is that we find considerable heterogeneity across countries in the strength of the relationship between efficiency and some of these school level characteristics or policies, a topic discussed in the next subsection.

Table 6. The inputs and outputs of very efficient schools, by country

Country	ESCS	StRatio	Computer_n	pv1math	pv1read	n
AUS	0.676	0.073	1.340	640.264	636.830	18
BEL	0.718	0.082	0.579	638.665	625.506	14
CAN	0.485	0.064	1.351	581.148	610.201	9
CHE	0.647	0.093	0.632	637.518	627.761	9
CZE	0.632	0.078	0.848	635.675	621.509	13
DEU	0.101	0.073	1.004	580.602	595.574	2
DNK	-0.686	0.069	2.800	508.408	537.994	1
ESP	-0.436	0.073	1.105	524.552	539.409	8
EST	0.635	0.108	0.810	622.729	637.337	5
FIN	0.669	0.128	0.254	664.666	645.177	6
FRA	0.392	0.084	0.505	590.715	631.255	9
GBR	0.838	0.064	1.018	602.717	621.805	5
IRL	0.914	0.073	0.545	574.026	627.753	1
ITA	0.245	0.082	0.378	597.066	610.742	12
JPN	0.130	0.086	0.352	613.103	617.450	44
KOR	0.218	0.072	0.411	632.269	604.796	20
NLD	0.646	0.070	0.811	624.941	625.671	7
NOR	-0.105	0.202	2.750	602.806	642.300	2
NZL	0.264	0.321	1.453	577.351	647.235	5
POL	-0.058	0.112	0.407	602.568	611.190	10
PRT	-0.667	0.079	0.362	527.473	540.924	5
SGP	-0.086	0.075	0.637	625.673	592.792	54
SVK	0.591	0.073	0.507	619.998	618.872	3
SVN	0.065	0.086	2.209	560.010	585.683	2
SWE	0.759	0.069	0.831	600.166	634.517	2
Total	0.248	0.086	0.696	614.485	611.199	266

Notes: a school is defined in the group of “very efficient” ones when its score is in the 95th percentile.

In Table 6 we report the results of the second-stage regression; we use both automatic backward and forward procedure (columns a and b, respectively) for maximising the explanatory powers of the model and selecting the statistically significant variables. As a robustness check, we also performed a comparison with the procedure suggested by Simar & Wilson (2007) – i.e. the double-bootstrap technique (available on request from the authors); the coefficients and main findings are substantially confirmed both qualitatively and quantitatively. The results are presented in a stepwise fashion, by adding the three groups of controls once a time.

Table 6b. Robustness check – excluding pv1_belowprof2 and pv1mathsd

	Model 1 (baseline)	Model 2 (without pv1_belowprof2 and pv1mathsd)
Students' characteristics		
immig_1	0.11724***	0.07358***
	0.006	0.007
female	0.02053***	0.01686***
	0.003	0.004
hwork_h	0.00295***	0.00726***
	0.000	0.000
repeater	-0.02175***	-0.08347***
	0.004	0.005
st_truancy	-0.01587**	-0.06804***
	0.005	0.006
ESCSsd	-0.06964***	-0.07489***
	0.004	0.004
School's general characteristics		
isced2		-0.01817***
		0.004
Orgen	0.01967***	0.02812***
	0.003	0.004
pv1mathsd	-0.00054***	
	0.000	
pv1_belowprof2	-0.17576***	
	0.004	
private	-0.01533***	-0.00607**
	0.002	0.002
clsize_small	-0.0043	
	0.002	
Size	0.00001***	0.00002***
	0.000	0.000
School's practices and processes		
Poor relations	-0.04716***	-0.05086***
	0.006	0.007
sc_matbui	0.00177**	0.00132
	0.001	0.001
prop_cert	0.02359***	0.03212***
	0.003	0.004
budget_2	0.00558***	0.00569***
	0.001	0.002
tc_part	-0.00201**	-0.00308***

Table 6b. Robustness check – excluding pv1_belowprof2 and pv1mathsd

	0.001	0.001
leadership_5	0.00817***	
	0.001	
accountability_1	-0.00616***	
	0.002	
qa_ext	0.00340*	
	0.001	
eval_teach	0.00454***	
	0.001	
Volunt	0.01166***	0.00835***
	0.002	0.002
select_1		0.00282
		0.002
competition		-0.00311*
		0.002
Constant	0.74996***	0.64887***
	0.007	0.009
Progname dummies	Y	Y
Country FE	Y	Y
Sigma	0.05052***	0.05933***
	0.000	0.000
N	7590	7590
Ll	7675.661	6456.05

Notes. Sigma is the estimated standard error of the regression (it is the equivalent to the root mean squared error in OLS). AUT, DNK and POL are excluded because of missing data or multicollinearity. In italics, standard errors. *, ** and *** mean statistically significant at .1, 1 and 5% respectively

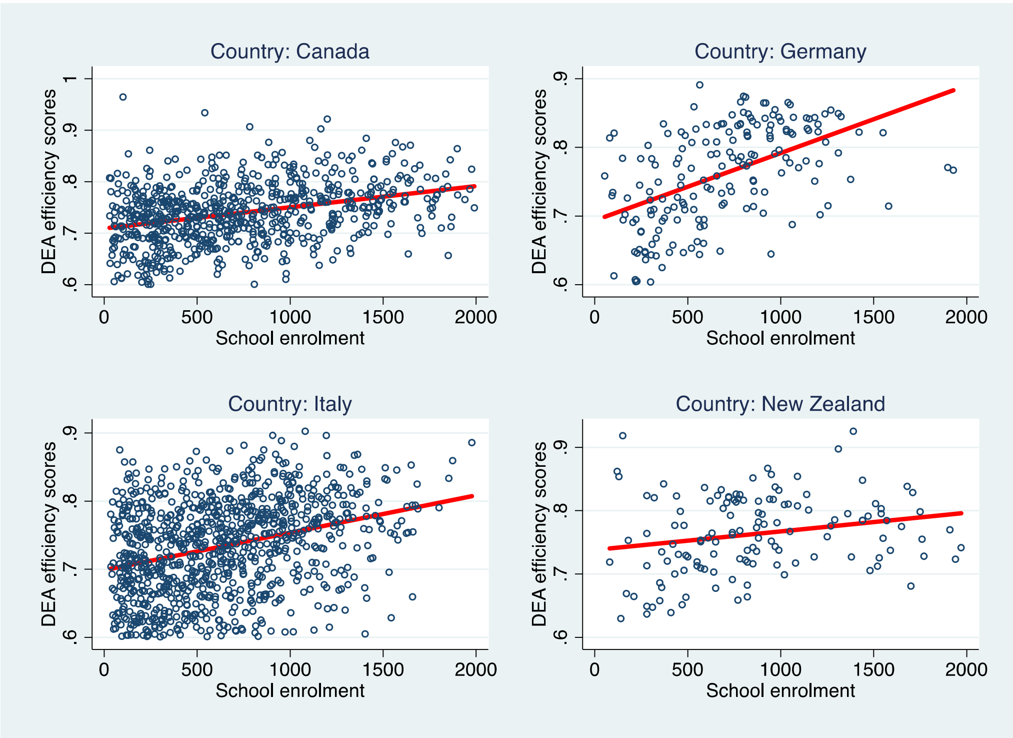
Model 1 (column a) presents the results for the characteristics of the student population in each school. The results highlight the complex relationship between efficiency and homogeneity of backgrounds (across socioeconomic status or country of origin, for example). Schools where the population of students have a more diverse socioeconomic background (as measured by *ESCSsd*) have lower efficiency scores; however, this negative association is counterbalanced by the positive effect on efficiency exerted by higher proportions of immigrant students (*immig_1*). Probably, the mechanism through which these variables impact efficiency is mediated by peer effects (Epple & Romano, 2011). These results suggest that having more homogenous classes implies using less resources for obtaining equal academic achievement – after having controlled for the students' socioeconomic background. A higher proportion of female students is associated with higher efficiency. Conversely, a higher proportion of students who reported to have skipped school days (*st_truancy*) in negatively related to efficiency (evidence about this point is also presented graphically in the figure 6). The number of hours that students devote for homework (*hwork_h*) shows a positive, albeit small in magnitude, relationship with efficiency of the school; however, it is impossible to disentangle how much of it is pure efficiency effects, and how is an indirect effect of students engagement, for example.

Model 2 (column (b)) adds schools' general characteristics to the general predictive model of efficiency. Schools with general/academic orientation are more efficient than their counterparts with vocational or technical focus. Conversely, there is no statistical difference in efficiency between schools where modal grade is ISCED 2 (lower secondary education) or ISCED 3 (upper secondary education).

Model 2 also explores equality issues. Schools with a wider dispersion of test scores (*pv1mashsd*) within the school have lower efficiency scores, but the magnitude of the estimated effect is negligible. In the same direction, but with a substantial magnitude, the model highlights the negative correlation between efficiency and the proportion of students with a test score below the baseline level of performance (Level 2; see *pv1_belowprof2*), which can be considered as a measure of inclusion or school failure (OECD, 2012b). Table 6b reports results of a model estimated without these measures of inclusion. The table shows that results (sign, and to a lesser extent, the coefficient) are not substantially affected. In this sense, taking them into consideration in the overall model does not distort the overall results.

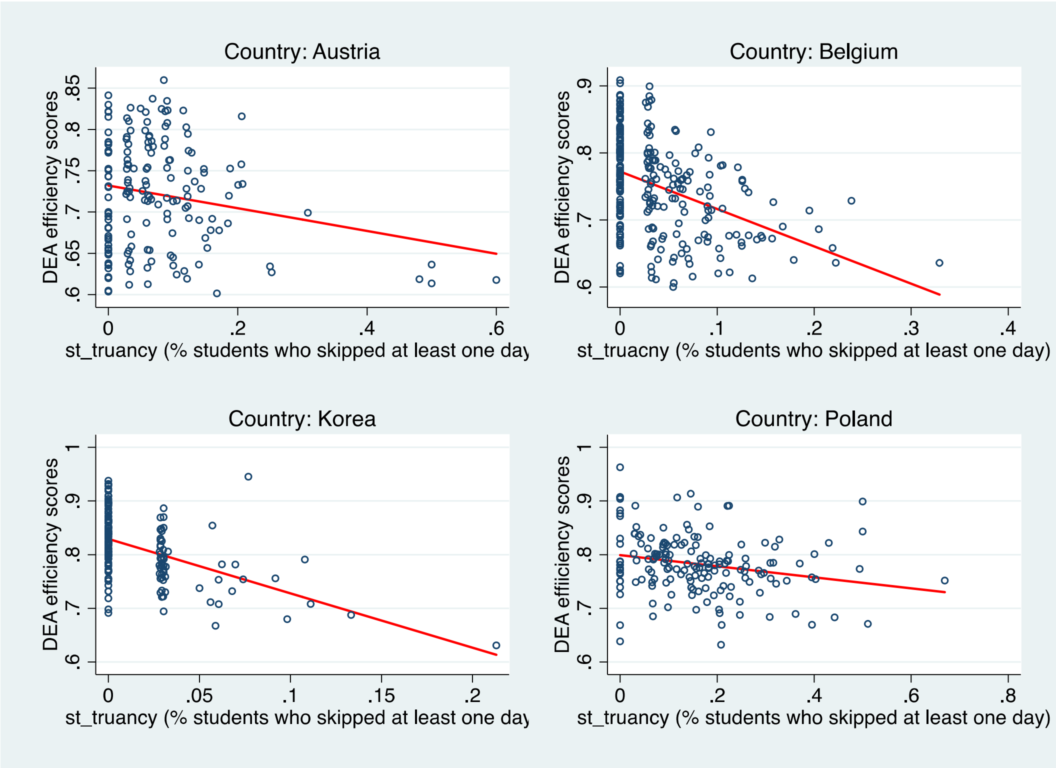
Model 2 also shows that having classes of a small average dimension (*clsize_small*) is negatively associated with efficiency, probably because it implies higher level of resources (i.e. teachers) and the benefits in achievement not always emerge (see for example, Krueger, 2003). We find no relationship between school size and efficiency (after taking into account all other student characteristics and school factors). This suggests that only very lager differences in size may be related to differences in efficiency. Sutherland et al. (2009), using PISA 2003 data, found that small schools tend to be less efficient. In Figure 4 we show that in some countries (see, for example, the case of Germany) there seem to be a clearer (positive) correlation between size and efficiency. In general, however, the correlation is very low.

Figure 4. The relationship between efficiency and size, selected countries



Model 3, column (c) explores the associations between efficiency scores and schools' practices and processes are explored. Two indicators of school climate are related with efficiency. The perception of non-positive relations between students and teachers (*poorelations*) and the index for participation of teachers to governance and decisions (*tc_part*) are negatively associated with efficiency scores. This finding suggests that when teachers have a negative attitude towards students or when they are not included in school matters obtaining higher scores for each unit of resources invested is more difficult. On the other side, figure 5 seems suggesting that, at least in some countries, the proportion of certified teachers is positively associated with efficiency.

Figure 5. The relationship between efficiency and student truancy



Model 3 also reveals that a number of activities undertaken by school principals can have a very positive effect on efficiency. More autonomy in allocating budgets across schools' activities (*budget_2*), exerting instructional leadership through meetings with teachers about the educational contents and strategies (*leadership_5*), involving external evaluators for quality assurance procedures (*qa_ext*) and using achievement scores for evaluating teachers (*eval_teach*) are all positively related with efficiency.

Model 3 also includes the quality of educational resources, measured by an index for the quality of educational infrastructures (*sc_matbui*) and the proportion of certified teachers (*prop_cert*), and shows that they are both positively associated with efficiency. These findings suggest that – all else equal – transforming inputs into achievement is easier if the resources (especially teachers) are of good quality[6].

Lastly, Model 3 includes some indicators for school practices, which have been explored in previous empirical work, and found related to schools' performances (Hanushek & Raymond, 2005; Guyon et al., 2012) – it is interesting to see if statistical significance holds also with efficiency. Selecting students at admissions (*select_1*) and making achievement results public (*accountability_1*) are associated with lower efficiency, while schools that organise volunteering as extracurricular activity (*volunt*) and grouping on ability between classes (*grouping*) tend to be more efficient (but, in this case, the size of effect is almost zero). There is no evidence of a relationship between efficiency scores and the measure of competition, as

[6] Our measure is a conventional proxy for the (certified) skills of teachers; although they usually have scarce relationship with students' achievement, the even more recent literature shows that when properly measured, teachers' activities and certifications do make a difference for students' results (for instance, see the discussion in Harris & Sass, 2011).

instead identified pointed out in similar settings by Agasisti (2013) for Italy, Bradley *et al.* (2001) for England, and Millimet & Collier (2008) for the United States.

Considering the whole picture together, this can suggest that giving more autonomy to school can allow them to obtain better or worse results, given the available resources, depending on how their managers use them. The findings support the idea that judgments about the various determinants of efficiency must be taken cautiously. Indeed, the magnitude of the coefficients is often very low, even lower than 0.01 (when the average efficiency is measured as high as 0.7). So, even taking all these factors into account (in addition to structural differences between country and programs), there is substantial unexplained differences across schools in terms of efficiency that are due to unobserved factors, and more research is needed in this respect.

Overall, the main factors in terms of magnitude are those related to the school composition in terms of students; even after having accounted for their background (ESCS is included among inputs, indeed), peer effects of various kinds are likely to improve or reduce the ability of schools to maximize students' achievement. As a matter of fact, the coefficients of the variables that measure student-related characteristics are higher than those that consider schools' features and practices. A graphical illustration that gives a sense of the magnitude of association between different variables and efficiency is reported in figure 7.

Figure 6. The relationship between efficiency and the proportion of certified teachers

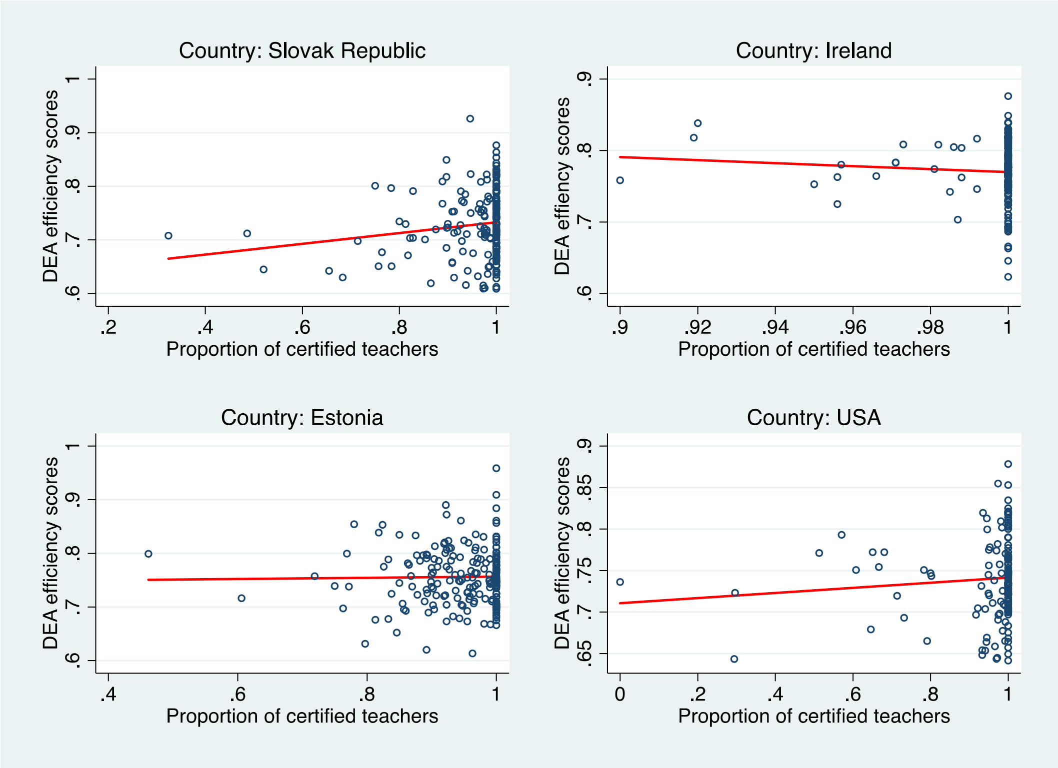
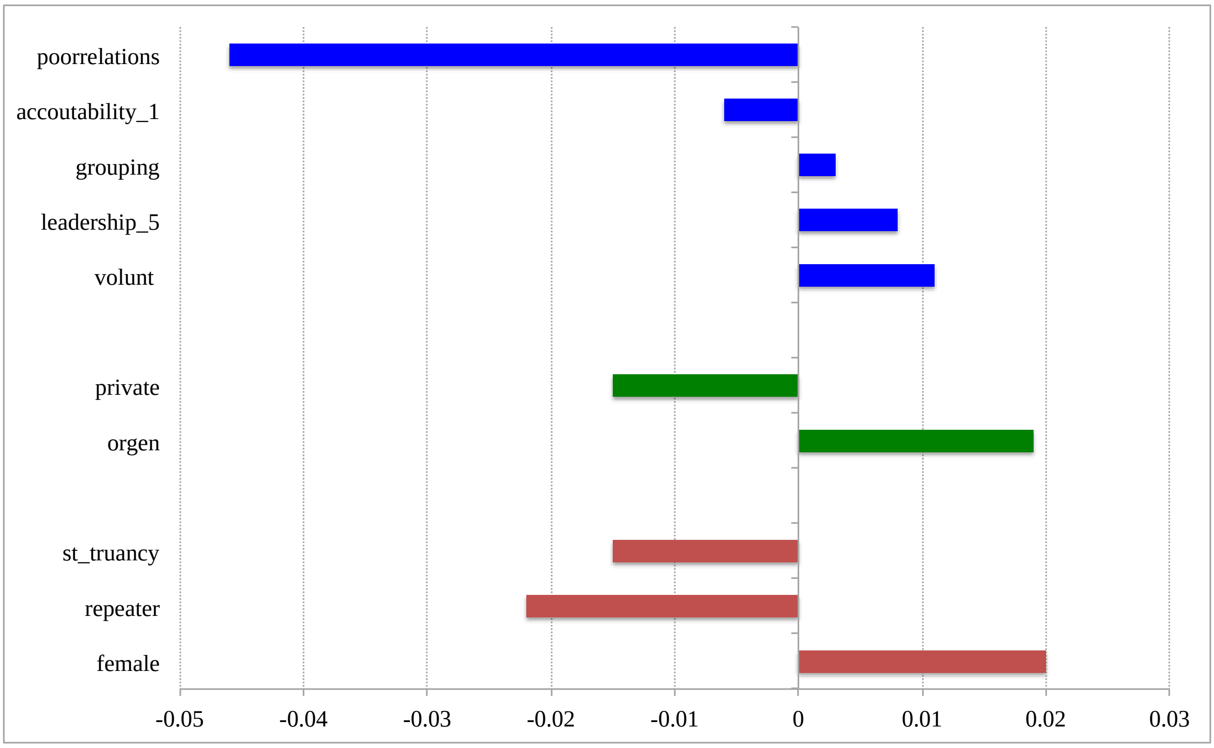


Figure 7. The magnitude of variables' association with efficiency



Notes: histograms built on the basis of marginal impact of variables, from Tobit regression.

Do these relationships hold across efficient and inefficient schools? And across countries?

We now explore differences across the efficiency frontier, at different levels of efficiency, in the relationship between efficiency and school characteristics, policies and practices. We also explore differences across countries. Table 7 reports the results of a quantile regression (columns a-c) and of regressions realised separately by country (columns d-f).

The results from the quantile regressions show that for some variables, whether a school is among the most or least efficient matters for the relationship between efficiency and these factors. For examples, schools with a general/academic orientation (*orgen*) are much more efficient than others when comparing among the most inefficient schools (those at the 25th percentile of the efficiency distribution). On the contrary, private schools (*private*) are relatively less efficient when comparing among the most efficient schools (those at the 75th percentile).

The positive relationship between the proportion of immigrant students (*immig_1*) and efficiency is detectable in the subgroup of relatively inefficient schools, while, the positive effect of the proportion of certified teachers as a proxy for school teachers' quality (*prop_cert*) is almost double for the subgroup of efficient than inefficient schools.

Making achievement scores public (*accountability_1*) is statistically negatively associated with efficiency only for least efficient schools, while involving externals in quality assurance procedures (*qa_ext*) is positively relevant only for most efficient schools. Volunteering as extracurricular activities (*volunt*) are associated with better efficiency for the group of schools at the lower end of distribution; perhaps indicating that it could be a viable way to improve their results in subsequent years. Interestingly, the quantile regression analysis reveals that the index that reports if the school is competing with two schools or more for the same students (variable *competition*) is negatively associated with efficiency at the top of the efficiency distribution. This does not necessarily mean that competition harms achievement; it can be the case that the eventually positive effects on achievement are not compensated by the higher amount of resources that are needed for competition. It may also signal that at the higher end of the distribution of performance, competition leads to sorting and cream-skimming that is based on socioeconomic status, not on performance (a discussion on a similar topic is in Hsieh & Urquiola, 2006 who studied the effects of a school choice program in Chile; see also Rutkowski et al., 2012).

Table 7 also reports on an analysis of the second-stage regressions within each country. In this case, we allow the relationship of efficiency with each factor to differ across countries[7]. In column (d), we report the number of countries (out of 30) for which each variable shows a statistically significant correlation with efficiency; columns (e) and (f) indicate the number of countries for which the correlation is positive or negative respectively.

[7] The dependent variable, in this case, is the efficiency score calculated when the frontier is the country-specific one. Complete results are available on request from the authors.

Table 7. The factors associated with efficiency scores, second-stage regression

	Model 1	Model 2	Model 3
	(a)	(b)	(c)
Students' characteristics			
immig_1	0.07960***	0.12650***	0.11719***
	0.006	*0.006*	*0.006*
Female	0.02280***	0.01734***	0.02064***
	0.003	*0.003*	*0.003*
hwork_h	0.00825***	0.00300***	0.00295***
	0.003	*0.000*	*0.000*
Repeater	-0.1076***	-0.02958***	-0.02208***
	0.004	*0.003*	*0.004*
st_truancy	-.06878***	-0.01525**	-0.01553**
	0.006	*0.005*	*0.005*
ESCSsd	-0.06787***	-0.07134***	-0.06998***
	0.004	*0.003*	*0.004*
School's general characteristics			
Orgen		0.02233***	0.01994***
		0.003	*0.003*
pv1mathsd		-0.00047***	-0.00053***
		0.000	*0.000*
pv1_belowprof2		-0.17334***	-0.17599***
		0.003	*0.004*
Private		-0.01528***	-0.01536***
		0.002	*0.002*
clsize_small		-0.00591**	-0.00434
		0.002	*0.002*
Size		0.00002***	0.00001***
		0.000	*0.000*
School's practices and processes			
Poor relations			-0.04674***
			0.006
sc_matbui			0.00177**
			0.001
prop_cert			0.02389***
			0.003
budget_2			0.00559***
			0.001
tc_part			-0.00200**
			0.001
leadership_5			0.00819***
			0.001
accountability1			-0.00626***
			0.002
qa_ext			0.00352*
			0.001
eval_teach			0.00464***
			0.001
Volunt			0.01166***
			0.002
select_1			0.00011
			0.001
Grouping			0.00383**
			0.001
Constant	0.6703***	0.75496***	0.72321***
	0.021	*0.018*	*0.020*
Progname dummies	*Y*	*Y*	*Y*
Country FE	*Y*	*Y*	*Y*
Sigma	0.06108***	0.05129***	0.05050***
log likelihood	7049.875	8302.278	7678.601
LR Chi2	4251.99	7032.98	6686.86

Notes. Sigma is the estimated standard error of the regression (it is the equivalent to the root mean squared error in OLS). AUT, DNK and POL are excluded because of missing data or multicollinearity. In italics, standard errors. *, ** and *** mean statistically significant at .1, 1 and 5% respectively.

For some variables, the relationship with efficiency is consistent across countries. For example, the proportion of students below proficiency level 2 (*pv1_belowprof2*) is negatively related to efficiency in all countries. Truancy, as measured by the proportion of students who reported to skip school days is negatively correlated with efficiency – all else equal in 12 out of 17 countries with a significant relationship. The proportion of female students is significantly positively related to efficiency in 10 out of 16 countries with a significant relationship.

For other variables, the number of countries for which the statistical correlation with efficiency holds is quite limited and sometimes inconsistent. In addition to the extreme case of *poor relations*, see for instance the index for the participation of teachers to governance (*tc_part*) and that of extracurricular activities (*volunt*). In these and other cases, the number of countries in which the variable plays a statistical role is lower than 10, and equally distributed between cases where the relationship is positive and others where is negative.

The main message from this country-by-country analysis is that while there are some factors, namely among student characteristics, that are consistently related to efficiency across countries, other relationships need to be analysed and interpreted from the specific context of the country under study. While attempts to draw general policy lessons for improving educational systems were made in previous studies using educational production functions (Wößmann, 2007), the analyses here call for caution when interpreting the relationship between schools' characteristics and school efficiency.

On efficiency and equity at school level

This paper analyses two different indicators of equality and equity and their relationship with efficiency. The first one is a proxy of equality in outputs (i.e. if the scores are very different within school or narrowly distributed across the mean). We measure within school equality with the standard deviation of efficiency scores (pv1mathsd). The other measure captures "inclusion", i.e. the ability of keeping the proportion of students below the baseline level of proficiency (Level 2) as low as possible (pv1_belowprof2).

While the statistical correlation between the efficiency scores and the former appears very low in magnitude, our measure of inclusion is one of the factors with a clearer, and stronger, association with efficiency at school level. On one side, this relationship is intuitive – even mechanical – as efficiency scores include achievement among its outputs, they must be higher if the proportion of low performers is lower (because all else equal, it means that the school average score is higher). Nonetheless, it is not always the case that a school with a lower proportion of low performers automatically has an average score lower higher than one with more low performers, because it depends by the entire distribution of scores. For instance, it can be that a school with a high proportion of very low and very high performers turns out to have a similar average score than a school when the distribution of test scores is narrow around the mean; these differences in distributions (that can be attributed to different teaching styles or strategies, for instance) then would be not reflected in differences on the mean. Conditional to equal inputs, this would also imply similar efficiency scores. In this perspective, a positive relationship between efficiency scores (effj) and pv1_belowprof2 can be interpreted as a key managerial finding: that schools are more efficient when are able to keep the proportion of low performers as low as possible, independently by the mean performance achieved.

Figure 8 illustrates the relationship between efficiency and inclusion across countries. The (negative) correlation at the country level is strong (-0.733). However, when looking at the same relationship at country level (see selected countries in Figure 9), the relationship between these two variables is not always as strong (ranging from -0.525 in Spain to -0.778 in Japan). Within a particular country, there are some schools with high levels of inefficiency despite the low proportion of students below proficiency level, and vice versa. This evidence suggests that the relationship between efficiency and inclusion is not straightforward. It shows that in general there is no trade-off between efficiency and equity.

Figure 8. The relationship between DEA efficiency scores and inclusion, country average

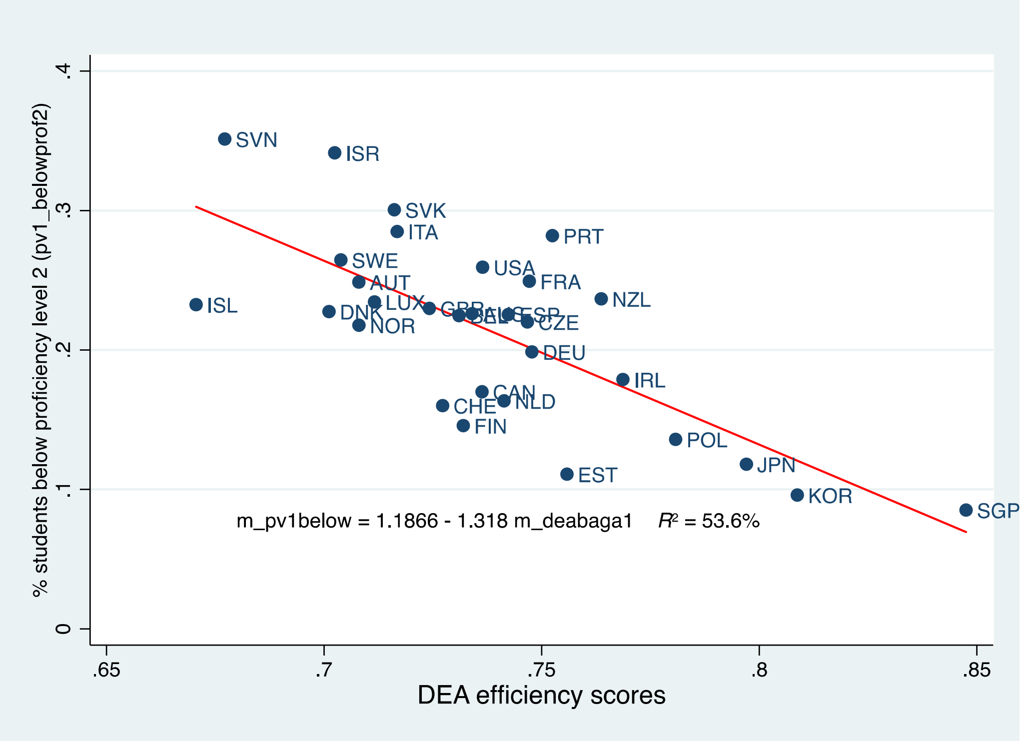


Figure 9. The relationship between DEA efficiency scores and inclusion, selected counties

Robustness checks

We performed a number of robustness checks on the results reported here.

First, we derived efficiency scores with an alternative frontier technique, namely Stochastic Frontier Analysis (SFA) – see Greene (2008) for technical details about the methodology, and Johnes (2004) for a discussion of relative advantages and shortcomings with respect to DEA. As the method allows the specification of a single output at a time, we estimate the production (frontier) function for the two subjects separately, adopting a translog functional form; mathematically:

$$\ln\left(y_{(M;R)j}\right) = \beta_0 + \sum_j \beta_j lnX_j + \sum_j \beta_j lnX_j^2 + \frac{1}{2}\sum_j \sum_i lnX_j \, lnX_i + \varepsilon_j \qquad (4)$$
$$\varepsilon_j = v_j - u_j$$

Where lnX_j and lnX_i are (log of) vectors of inputs (as specified in the DEA analysis), and $\varepsilon_j = v_j - u_j$ is the typical error term in a SFA contexts, decomposed in the idiosyncratic error v_j and one-sided inefficiency term u_j. Assuming an half-normal distribution for u_j, the j-th school's efficiency score is estimated via $E\{\exp(-u_j)|\epsilon_j\}$ where u_j are the estimates of minus the natural log of the technical efficiency via $E(u_j|\epsilon_j)$. Two sets of efficiency scores are then generated, one for mathematics and one for reading as outputs (*SFA_eff$_j$_Math* and *SFA_eff$_j$_Read*, respectively).

Then, we estimated three new sets of efficiency scores by using alternative DEA specifications, where the set of inputs is changed. Recalling that the three inputs included in the baseline are school-average ESCS, the inverse of students/teachers ratio (*StRatio*) and the number of computers per student (*Computer_n*), the set of alternative DEA specifications used here are: (i) only StRatio and ESCS (*eff$_j$_rc1*), (ii) only StRatio and Computer_n (*eff$_j$_rc2*), and (iii) only ESCS and Computer_n (*eff$_j$_rc3*).

Lastly, we perform four additional DEA analyses, using the plausible values 2, 3, 4 and 5 as outputs and the three baseline inputs (the resulting efficiency scores are labelled *eff$_j$_pv2*, *eff$_j$_pv3*, *eff$_j$_pv4* and *eff$_j$_pv5*). In all the alternative DEA specifications, we use a bootstrap computation (number of replicates equal to 2,000) and keep the bias-corrected efficiency score.

Table 8 contains the descriptive statistics of the efficiency scores computed as robustness checks (panel A), and the correlation indexes between them and the baseline scores discussed in previous sections (panel B). Two types of correlation indexes are reported: Pearson's product-moment correlation coefficient, and Spearman's rank correlation coefficient. All the correlations are strong and indicate that the results are robust to alternative specifications. Among DEA scores, the lower coefficients are those for the model *eff$_j$_rc2* (around 0.88), where ESCS in not included among inputs, and this corroborates the importance of taking students' background into account when measuring efficiency of schools. In all other cases, correlations are >0.95. The correlation indexes with efficiency scores derived through SFA are negative, because the latter method build technical efficiency scores in the interval [0;1] where 1 is maximum efficiency. Albeit high (>0.7) and statistically significant, the coefficients are lower than those for alternative DEA specification, so suggesting that results are more robust within methods than across them.

Table 8. The heterogeneity of factors associated with efficiency – quantile regression

	Dependent variable: bias-corrected efficiency score, international frontier			Dependent variable: bias-corrected efficiency score, country-specific frontier		
	25th percentile	Median	75th percentile	statistically significant	(+) w/eff	(-) w/eff
	(a)	(b)	(c)	(d)	(e)	(f)
isced2	-0.00279	-0.01904***	0.00828	11	3	8
	0.004	*0.003*	*0.005*			
Orgen	0.01700***	0.01319***	0.02858***	11	9	2
	0.004	*0.003*	*0.005*			
pv1mathsd	-0.00023***	-0.00017***	-0.00070***	12	8	4
	0.000	*0.000*	*0.000*			
pv1_belowprof2	-0.19111***	-0.20931***	-0.17332***	30	0	30
	0.004	*0.003*	*0.005*			
Private	-0.01835***	-0.01541***	-0.00584*	15	3	12
	0.002	*0.002*	*0.003*			
clsize_small	0.00838**	0.00264	-0.01709***	11	3	8
	0.003	*0.002*	*0.003*			
Size	0.00002***	0.00002***	0.00001***	12	6	6
	0.000	*0.000*	*0.000*			
ESCSsd	-0.05671***	-0.04019***	-0.08406***	15	6	9
	0.004	*0.003*	*0.005*			
immig_1	0.10924***	0.05445***	0.24182***	14	9	5
	0.007	*0.006*	*0.009*			
female	0.02815***	0.02716***	0.00747	16	16	0
	0.004	*0.003*	*0.005*			
hwork_h	0.00339***	0.00271***	0.00388***	11	9	2
	0.000	*0.000*	*0.000*			
repeater	-0.01385**	0.01533***	-0.04470***	13	5	8
	0.005	*0.004*	*0.006*			
st_truancy	-0.02031***	-0.01249**	-0.02563***	16	4	12
	0.006	*0.005*	*0.007*			
poorrelations	-0.02930***	0.00263	-0.06866***	2	1	1
	0.007	*0.005*	*0.009*			
sc_matbui	0.00119	0.00325***	-0.00086	10	3	7
	0.001	*0.001*	*0.001*			
prop_cert	0.02059***	0.02947***	0.02353***	9	7	2
	0.004	*0.003*	*0.005*			
budget_2	-0.00035	-0.00124	0.00395	8	6	2
	0.002	*0.001*	*0.002*			
tc_part	-0.00244**	-0.00086	-0.00164	4	4	0
	0.001	*0.001*	*0.001*			
leadership_5	-0.00146	-0.00099	0.00168	10	6	4
	0.002	*0.001*	*0.002*			
accountability_1	-0.00624***	-0.00122	-0.00999***	7	3	4
	0.002	*0.001*	*0.002*			
qa_ext	0.00055	0.01174***	-0.00215	7	3	4
	0.002	*0.001*	*0.002*			
eval_teach	0.00991***	0.00338**	0.01313***	10	6	4
	0.002	*0.001*	*0.002*			
volunt	0.01191***	0.00377*	0.01747***	9	6	3
	0.002	*0.001*	*0.002*			
select_1	-0.00441**	0.00176	-0.00204	8	3	5
	0.002	*0.001*	*0.002*			
grouping	0.00391*	-0.00146	0.00603**	9	3	6
	0.002	*0.001*	*0.002*			
competition	0.00169	-0.00751***	0.00395*	16	7	9
	0.002	*0.001*	*0.002*			
Constant	0.70130***	0.67661***	0.77155***			
	0.023	*0.018*	*0.029*			
Progname dummies	Y	Y	Y			
Country FE	Y	Y	Y			
pseudoR2	0.4494	0.3858	0.3462			

Notes. In italics, standard errors. *, ** and *** mean statistically significant at .1, 1 and 5% respectively.

Table 9. Robustness checks

Panel A. Descriptive statistics of alternative efficiency scores

Variable	Mean	Std. Dev.	Min	Max
eff$_j$	0.734	0.077	0.316	0.978
SFA_eff$_j$_Math	0.913	0.047	0.404	0.989
SFA_eff$_j$_Read	0.882	0.070	0.424	0.989
eff$_j$_rc1	0.730	0.078	0.315	0.994
eff$_j$_rc2	0.694	0.093	0.235	0.985
eff$_j$_rc3	0.718	0.073	0.283	0.977
eff$_j$_pv2	0.745	0.076	0.326	0.985
eff$_j$_pv3	0.736	0.078	0.274	0.983
eff$_j$_pv4	0.728	0.073	0.305	0.982
eff$_j$_pv5	0.744	0.075	0.332	0.986

Panel B. Correlation indexes between alternative efficiency scores

Pearson's corr. index	eff$_j$	SFA_eff$_j$_Math	SFA_eff$_j$_Read	eff$_j$_rc1	eff$_j$_rc2	eff$_j$_rc3	eff$_j$_pv2	eff$_j$_pv3	eff$_j$_pv4	eff$_j$_pv5
eff$_j$	1									
SFA_eff$_j$_Math	0.7262*	1								
SFA_eff$_j$_Read	0.8732*	0.7362*	1							
eff$_j$_rc1	0.9841*	0.7330*	0.8500*	1						
eff$_j$_rc2	0.8814*	0.5964*	0.7525*	0.9043*	1					
eff$_j$_rc3	0.9677*	0.7541*	0.8755*	0.9575*	0.8722*	1				
eff$_j$_pv2	0.9726*	0.7257*	0.8857*	0.9618*	0.8524*	0.9571*	1			
eff$_j$_pv3	0.9701*	0.6975*	0.8822*	0.9502*	0.8592*	0.9515*	0.9701*	1		
eff$_j$_pv4	0.9531*	0.7735*	0.8797*	0.9427*	0.8014*	0.9535*	0.9596*	0.9411*	1	
eff$_j$_pv5	0.9671*	0.7399*	0.8810*	0.9493*	0.8244*	0.9535*	0.9745*	0.9649*	0.9614*	1

Spearman's corr. index	eff$_j$	SFA_eff$_j$_Math	SFA_eff$_j$_Read	eff$_j$_rc1	eff$_j$_rc2	eff$_j$_rc3	eff$_j$_pv2	eff$_j$_pv3	eff$_j$_pv4	eff$_j$_pv5
eff$_j$	1									
SFA_eff$_j$_Math	-0.6863*	1								
SFA_eff$_j$_Read	-0.8913*	0.6996*	1							
eff$_j$_rc1	0.9817*	-0.6903*	-0.8641*	1						
eff$_j$_rc2	0.8588*	-0.5226*	-0.7051*	0.8892*	1					
eff$_j$_rc3	0.9617*	-0.7206*	-0.8954*	0.9500*	0.8467*	1				

Panel B. Correlation indexes between alternative efficiency scores

eff$_j$_pv2	0.9700*	-0.6884*	-0.9099*	0.9569*	0.8283*	0.9523*	1			
eff$_j$_pv3	0.9674*	-0.6538*	-0.8968*	0.9448*	0.8330*	0.9450*	0.9682*	1		
eff$_j$_pv4	0.9443*	-0.7552*	-0.9104*	0.9312*	0.7641*	0.9473*	0.9525*	0.9312*	1	
eff$_j$_pv5	0.9636*	-0.7095*	-0.9122*	0.9415*	0.7930*	0.9488*	0.9728*	0.9622*	0.9571*	1

Notes: * means that the correlation is statistically significant at 0.1% level.

Concluding remarks

An international efficiency benchmark suggests substantial gains, measured as improvements in PISA score points, are possible given the current allocation of resources to schools. On average, we estimate that schools could raise their score by 27% when considering the distribution of performance and resources internationally, and as much as 15% when comparing themselves with schools in their own country. These estimates are consistent with those in the literature (Sutherland et al. 2009).[8]

The use of international, comparable data provides a unique opportunity to analyse efficiency in the provision of education services. This paper represents one of the first attempts of using PISA for a comparative efficiency analysis of schools with a large number of schools (more than 8,600) and countries (30). The analysis included here provides evidence on the degree of variation in efficiency between and within countries. It also includes an exploration of the potential determinants of efficiency within countries.

The efficiency scores estimated here are only proxies for true efficiency, given the potential imprecisions in measurement of some variables. Most importantly, they cannot be used to as precise measures of efficiency at the school level. For example, any attempt to use these measures to rank schools according to their efficiency score would be ill conceived. And yet, the main strength of the analysis presented in the paper lies in that it provides a clear picture of the distribution of schools' efficiency scores across and within countries.

A key message from these findings is that in terms of efficiency the "average" school in a country does not exist. In fact, we find that the heterogeneity is higher within countries than between them. A second key finding is that an international benchmark is most promising and perhaps the most appropriate because a wider set of alternative combinations of inputs and outputs can be considered to define efficiency, and ways to pursue it.

In terms of the efficiency-enhancing factors, we find that there is not a one-size-fits-all set of factors that improve schools' efficiency across countries. Although second-stage regressions reveal that some

[8] The estimates of Sutherland et al. (2009) are lower, between 5% and 10%, but are obtained through Stochastic Frontier Analyses. When considering SFA estimates obtained in our paper, the results are comparable, as the average efficiency score is 0.91 (i.e. the correspondent degree of inefficiency is around 9%).

school-level indicators are associated with higher (or lower) efficiency, there is wide variation across countries in the nature and strength of those relationships. Factors such as the proportion of female students and the proportion of students above the baseline level, are positively associated with school efficiency. Other factors, such as ability grouping between classes, principal's autonomy in allocating budget, leadership style (i.e. dealing with instructional tasks), quality assessment and teachers' evaluation, etc, the relationship with efficiency varies considerably across countries, both the direction and the significance. Moreover, an analysis of these relationships across the entire distribution of performance highlights that some of these variables are more important in explaining efficiency at lower tail of scores' distribution, while others at the top of it. These findings suggest that, the way the schools implement policies, such as school autonomy, is as important, if not more, than the actual policy itself.

We also find no evidence in favour of a trade-off between efficiency and inclusion. Schools with higher efficiency scores are also those with a smaller proportion of students who perform below proficiency level 2 (*pv1_belowrprof2*). While the data is not sufficiently powerful to distinguish if having a higher proportion of low-performing students reduces the productivity of educational activities and resources or more efficient schools raise average achievement standards, it is clear that it is possible and indeed quite common to have efficient schools where all students achieve at levels above the baseline.

An analysis of structural differences in efficiency across countries is outside the scope of this paper. Several areas for future research include; for instance, adding country-level varying factors to the second-stage regression (together with country fixed-effects) which can help in understanding if are there institutional factors such as accountability, competition, central examinations or tracking policies that affect the structural differences of schools' efficiency scores across countries. Also, it would be interesting to compare if these factors share analogies with those that have been proven to influence achievement *per se,* as discussed by Hanushek & Wößmann (2010).

Data availability affect the quality of efficiency analyses from an international perspective. They provide insights into how to enhancing PISA and other international large-scale assessments. For example, the absence of indicators about students' prior achievement imposes the assumption about the index of socio-economic status is a good proxy for early learning opportunities; the development of measures about prior achievement level would help disentangling school effects and students' ability. The lack of data about school expenditures prevents the estimation of cost efficiency, which would be a natural and important extension of efficiency analyses from a comparative international perspective.

PISA 2012 includes a number of measures of non-cognitive skills; indeed, OECD (2013b) analyses a number of indicators on students' perseverance, locus of control, motivation, etc. Forthcoming research including these variables among outputs will explore how schools efficiency estimates change, when adding these dimensions into the analysis.

REFERENCES

Afonso, A. and M. St Aubyn (2006), “Cross-country efficiency of secondary education provision: A semi-parametric analysis with non-discretionary inputs”, *Economic Modelling*, Vol 23 (3), pp. 476-491.

Agasisti, T. (2014), “The efficiency of public spending on education: An empirical comparison of EU countries”, *European Journal of Education*, Vol 49 (4), pp.543-557.

Agasisti, T. (2013), “The efficiency of Italian secondary schools and the potential role of competition: A data envelopment analysis using OECD-PISA2006 data”, *Education Economics*, Vol 21 (5), pp. 520-544.

Bradley, S., G. Johnes and J. Millington (2001), “The effect of competition on the efficiency of secondary schools in England”, *European Journal of Operational Research*, Vol. 135 (3), pp. 545-568.

Clements, B. (2002), “How efficient is education spending in Europe?” *European Review of Economics and Finance*, Vol. 1 (1), pp. 3-26.

Epple, D. and R. Romano (2011), “Peer effects in education: A survey of the theory and evidence”, *Handbook of Social Economics*, Vol. 1(11), pp. 1053-1163.

Giménez, V., D. Prior and C. Thieme (2007), “Technical efficiency, managerial efficiency and objective-setting in the educational system: an international comparison”, *Journal of the Operational Research Society*, Vol 58 (8), pp. 996-1007.

Greene, W. H. (2008), “The econometric approach to efficiency analysis”, *The measurement of productive efficiency and productivity growth*, pp. 92-250.

Guyon, N., E. Maurin and S. McNally (2012), “The effect of tracking students by ability into different schools: A natural experiment. *Journal of Human Resources*, Vol 47 (3), pp.684-721.

Hanushek, E. A. and L. Woëßmann (2010), “The economics of international differences in educational achievement”, *National Bureau of Economic Research*, (No. w15949)

Hanushek, E. A. and M.E. Raymond (2005), “Does school accountability lead to improved student performance?” *Journal of Policy Analysis and Management*, Vol 24 (2), pp. 297-327.

Harris, D. N. and T. R. Sass (2011), “Teacher training, teacher quality and student achievement”, *Journal of Public Economics*, Vol 95 (7), pp. 798-812.

Hsieh, C. T. and M. Urquiola (2006), “The effects of generalised school choice on achievement and stratification: Evidence from Chile's voucher program”, *Journal of Public Economics*, Vol 90 (8), pp. 1477-1503.

Johnes, J. (2004), “16 Efficiency measurement” *International Handbook on the Economics of Education*, pp. 613.

Koenker, R. and K. Hallock (2001), Quantile regression: An introduction. *Journal of Economic Perspectives*, Vol 15 (4), pp. 43-56.

Krueger, A. B. (2003), Economic considerations and class size. The Economic Journal, 113(485), F34-F63.

Levin, H. M., D. T. Jamison and R. Radner (1976), "Concepts of economic efficiency and educational production", *In Education as an Industry,* NBER, pp. 149-198.

Mancebón, M. J., et al. (2012), "The efficiency of public and publicly subsidized high schools in Spain: Evidence from PISA-2006", *Journal of the Operational Research Society*, Vol. 63 (11), pp. 1516-1533.

McDonald, J. (2009), "Using least squares and tobit in second stage DEA efficiency analyses", *European Journal of Operational Research*, Vol. 197 (2), pp. 792-798.

Millimet, D. L. and T. Collier (2008), "Efficiency in public schools: Does competition matter?", *Journal of Econometrics*, Vol. 145 (1), pp. 134-157.

OECD (2013a), *PISA 2012 Results*: *What makes schools successful (Volume IV): Resources, Policies and Practices,* PISA, OECD Publishing Paris, http://dx.doi.org/10.1787/9789264201156-en

OECD (2013b), *PISA 2012 Results: Ready to Learn (Volume III): Students' Engagement, Drive and Self-Beliefs*, PISA, OECD Publishing Paris, http://dx.doi.org/10.1787/9789264201170-en

OECD (2012a), PISA 2009 Technical Report, PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264167872-en

OECD (2012b), *Equity and Quality in Education: Supporting Disadvantaged Students and Schools*, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264130852-en

Portela, M. C. A. and E. Thanassoulis (2001), "Decomposing school and school-type efficiency", *European Journal of Operational Research*, Vol. 132 (2), pp. 357-373.

Rutkowski, L., D. Rutkowski, D. and J. Plucker (2012), "International determinants of private school attendance", *Educational Research and Evaluation*, Vol 18 (4), pp. 375-391.

Schleicher , A. (2014), *Equity, Excellence and Inclusiveness in Education: Policy Lessons from Around the World*, International Summit on the Teaching Profession, OECD Publishing, Paris. DOI: http://dx.doi.org/10.1787/9789264214033-en

Simar, L. and P.W. Wilson (2007), "Estimation and inference in two-stage, semi-parametric models of production processes", *Journal of Econometrics*, Vol. 136 (1), pp. 31-64.

Simar, L., and P.W. Wilson (2000), "A general methodology for bootstrapping in non-parametric frontier models", *Journal of Applied Statistics*, Vol 27 (6), pp. 779-802.

Sutherland, D., R. Price and F. Gonand (2009), "Improving public spending efficiency in primary and secondary education", *OECD Journal: Economic Studies*, 2009(1), pp. 1-30.

Thanassoulis E., M.S.C. Portela and O. Despic (2008), "The measurement of productive efficiency and productivity growth", *Data envelopment analysis: the mathematical programming approach to efficiency analysis*, pp. 251-420.

Wößmann, L. (2008). Efficiency and equity of European education and training policies. International Tax and Public Finance, 15(2), 199-230.

Wößmann, L. (2007). International evidence on school competition, autonomy, and accountability: A review. Peabody Journal of Education, 82(2-3), 473-497.

Wößmann, L. (2003), “Schooling resources, educational institutions and student performance: the international evidence”, *Oxford Bulletin of Economics and Statistics*, Vol. 65 (2), pp. 117-170.Worthington, A. C. (2001). An empirical survey of frontier efficiency measurement techniques in education. Education Economics, 9(3), 245-268.

ANNEX 1. INPUTS AND OUTPUTS – DESCRIPTIVE STATISTICS, BY COUNTRY

Country	n	ESCS (mean/sd)	StRatio (mean/sd)	Computer_n (mean/sd)	pv1math (mean/sd)	pv1read (mean/sd)
AUS	718	0.191	0.079	1.540	495.909	503.471
		0.46	0.02	1.40	59.28	60.59
AUT	178	0.020	0.138	1.754	490.204	475.238
		0.55	0.13	5.37	73.27	71.78
BEL	264	0.104	0.133	0.683	508.464	500.772
		0.53	0.08	0.65	81.12	81.43
CAN	753	0.369	0.082	1.132	506.668	508.946
		0.43	0.10	1.76	46.89	48.64
CHE	369	0.086	0.095	0.630	514.480	491.514
		0.41	0.04	0.67	56.29	55.90
CZE	249	-0.057	0.102	0.960	500.460	496.301
		0.44	0.14	0.96	77.19	71.08
DEU	194	0.125	0.080	0.698	506.127	499.447
		0.54	0.07	1.50	78.30	76.23
DNK	283	0.274	0.128	0.956	487.781	484.696
		0.44	0.20	1.14	42.96	46.61
ESP	841	-0.166	0.103	0.723	488.871	488.364
		0.54	0.07	0.63	46.89	47.61
EST	199	0.027	0.110	0.835	517.874	514.117
		0.46	0.07	1.08	38.65	41.94
FIN	294	0.392	0.102	0.509	515.609	519.333
		0.36	0.03	0.44	46.32	48.59
FRA	193	-0.071	0.090	0.565	491.619	502.184
		0.45	0.03	0.55	75.19	84.79
GBR	447	0.233	0.070	0.978	487.539	496.491
		0.38	0.02	0.64	49.97	50.79
IRL	152	0.102	0.074	0.669	499.761	521.593
		0.42	0.02	0.54	40.22	46.43
ISL	112	0.598	0.121	0.986	488.954	479.208
		0.43	0.04	0.94	45.24	50.53
ISR	141	0.133	0.100	0.392	463.691	482.825
		0.49	0.04	0.34	72.42	80.74
ITA	1,044	-0.125	0.128	0.540	477.247	477.940
		0.55	0.09	0.48	72.72	80.18
JPN	190	-0.091	0.123	0.619	534.274	535.426
		0.37	0.11	0.88	71.12	70.62
KOR	154	0.007	0.068	0.388	551.658	534.127
		0.37	0.03	0.47	63.50	54.12
LUX	39	0.128	0.119	0.842	491.747	487.234
		0.62	0.02	1.00	56.64	60.23
NLD	143	0.191	0.066	0.671	512.702	501.646
		0.37	0.03	0.53	78.22	80.74
NOR	177	0.453	0.102	0.822	491.185	505.514
		0.28	0.03	0.45	40.31	46.35
NZL	149	0.013	0.080	1.175	498.820	513.478
		0.44	0.10	0.67	54.37	59.74
POL	166	-0.093	0.131	0.377	525.942	525.380
		0.60	0.08	0.29	60.64	59.55
PRT	171	-0.505	0.132	0.506	479.325	480.033
		0.66	0.09	0.48	58.22	57.43
SGP	163	-0.288	0.075	0.659	568.773	537.067
		0.46	0.02	0.39	66.20	62.30
SVK	195	-0.289	0.082	0.812	472.479	451.335
		0.61	0.02	0.51	71.95	84.06
SVN	317	-0.142	0.253	0.857	462.550	439.830
		0.50	0.59	0.89	74.57	81.55
SWE	193	0.303	0.088	0.737	484.947	489.264
		0.37	0.03	1.39	46.56	55.41
USA	152	0.194	0.076	0.927	481.553	497.942
		0.55	0.09	0.74	48.53	52.19
Total	8,640	0.063	0.105	0.836	496.583	495.393
		0.53	0.14	1.25	63.74	66.35