OECD Economics Department Working Papers No. 546

Performance Indicators for Public Spending Efficiency in Primary and Secondary Education

Douglas Sutherland, Robert Price,
Isabelle Joumard, Chantal Nicq

OECDPUBLISHING

Organisation de Coopération et de Développement Economiques
Organisation for Economic Co-operation and Development

27-Feb-2007

_____
English - Or. English

ECONOMICS DEPARTMENT

PERFORMANCE INDICATORS FOR PUBLIC SPENDING EFFICIENCY IN PRIMARY AND SECONDARY EDUCATION

Economics Department Working Paper No. 546

by Douglas Sutherland, Robert Price, Isabelle Joumard and Chantal Nicq

All Economics Department Working Papers are available through OECD's internet web site at www.oecd.org/eco/Working_Papers

JT03222661

Document complet disponible sur OLIS dans son format d'origine
Complete document available on OLIS in its original format

TABLE OF CONTENTS

Tables

Figures

ABSTRACT/RÉSUMÉ

Performance indicators for spending efficiency in primary and secondary education

This paper assesses the potential to raise public spending efficiency in the primary and secondary education sector. Resource availability per pupil has increased significantly over the past decade in a number of countries; often in attempting to exploit the link between educational attainment and growth. However, available evidence reveals only a weak correlation between increased resource availability and pupil performance. In order to draw cross-country comparisons of the efficiency in the provision of education, the paper develops a set of comparable indicators which reflect international differences in the levels of efficiency in the primary and secondary education sector both within and among countries. The paper identifies significant scope to improve efficiency by moving towards best practice. These results are robust to a variety of approaches and are most certain when the potential gain from eliminating inefficiency is larger.

JEL classification: C14; C21; D24; H41; H52; I21; I22; I28

Keywords: Education; public spending; efficiency; nonparametric methods; data envelopment analysis; stochastic frontier analysis.

*****

Indicateurs de performance de l'efficacité des dépenses publiques d'éducation dans l'enseignement primaire et secondaire

Ce document évalue les possibilités d'accroître l'efficacité des dépenses publiques d'éducation dans l'enseignement primaire et secondaire. Les ressources disponibles par élève se sont accrues sensiblement au cours de la dernière décennie dans un certain nombre de pays, souvent dans l'espoir d'exploiter le lien entre niveau d'instruction et croissance. Pourtant les études disponibles ne révèlent qu'une faible corrélation entre la disponibilité croissante des ressources et la performance des élèves. Afin d'établir des comparaisons internationales de l'efficacité des prestations en matière d'éducation, cette étude met au point un ensemble d'indicateurs comparables qui reflètent les différences internationales entre les niveaux d'efficacité dans le secteur de l'enseignement primaire et secondaire à la fois dans chaque pays et entre pays. Elle met en évidence des marges significatives pour améliorer l'efficacité en se rapprochant des meilleures pratiques. Ces résultats sont robustes au regard de plusieurs approches et sont d'autant plus fiables que les gains potentiels résultant de l'élimination de l'inefficacité sont importants.

Classification JEL : C14 ; C21 ; D24 ; H41 ; H52 ; I21 ; I22 ; I28

Mots clés : Éducation ; dépenses publiques ; efficacité ; méthodes non paramétriques ; analyse par enveloppement des données ; analyse de frontière stochastique.

PERFORMANCE INDICATORS FOR PUBLIC SPENDING EFFICIENCY IN PRIMARY AND SECONDARY EDUCATION

By Douglas Sutherland, Robert Price, Isabelle Joumard and Chantal Nicq[1]

1. Introduction: synthesis and main results

1. This paper is presented as part of a larger ongoing project aimed at drawing cross-country comparisons of efficiency in the provision of public services. It focuses on primary and lower secondary education, on which governments in the OECD area spend on average around 3% of GDP. Influenced by the alleged link between higher levels of educational attainment and growth, the education sector has seen significant reform efforts in recent years in a number of countries and spending per student has often risen markedly since the mid-1990s (Figure 1, upper panel). In some countries, the ratio of teachers to pupils increased steadily over this period, often reflecting efforts to reduce class size; in others, teachers' relative pay was raised substantially. However, cross-sectional evidence reveals only a weak correlation between either spending or teachers per student and mean pupil performance in standardised tests (Figure 1, lower panels). While the apparently weak link between educational resources and outcomes may reflect the influence of a wide range of socio-economic factors, it raises questions about possible failures to use inputs effectively or to contain their relative cost.

2. To investigate these issues, this paper develops a set of indicators which reflect international differences in technical efficiency in the primary and lower-secondary sector within and among countries, by using, respectively, data for individual schools and internationally-comparable country-level data. The two sets of results are complementary -- the larger data set for schools allows for a more disaggregated specification of the basic input-output model for example. At the country level, indicators of cost efficiency are also developed.

3. Much of the paper is about definition and method. Inputs are described both in physical terms, the relationship between input volumes (teachers) and outputs being a measure of technical efficiency, and in terms of the amount spent per pupil, which determines cost efficiency. The main difficulty on the input side is to define inputs not under the control of education providers, such as those related to pupils' socio-economic background. Outputs are conceptually more complex. They could be measured narrowly, in terms of hours of schooling or diplomas or, very broadly, in terms of the increments to economic welfare or GDP. This paper looks at intermediate outcomes, as measured by the mean of PISA scores in four academic disciplines, which may be seen as some of the basic elements in human capital accumulation. Since educational systems may also be charged with distributional goals such as mitigating

---

inequality, a measure of the homogeneity of the PISA scores is also treated as an outcome at the national level. The weight given to these objectives may vary across countries, reflecting social preferences, but, subject to the fact that pupils' socio-economic background may differ, there do not seem to be trade offs between the level of educational attainment and the homogeneity of student achievement (Figure 2).

4.   The main findings of the paper with respect to the performance of OECD economies in primary and lower-secondary education provision are as follows:

- At the school level, there appear to be significant opportunities to improve technical efficiency by moving towards best practice, controlling for the effects of socio-economic and linguistic background on educational attainment. The scope identified for resource savings rises to as much as one third for the median school if the school raised its level of efficiency to that of best practice. Raising efficiency levels to at least the 95th percentile or the median level would still give significant savings. The scope identified for boosting output while holding inputs constant is estimated to be somewhat smaller at around one fifth. Also, within some countries -- but not generally -- the variation in school-level technical efficiency is quite large.

- At the national level, controlling for the effects of socio-economic background on educational attainment, the results are broadly similar to the school-level results. Holding resources constant while moving towards best practice, PISA scores could be boosted by an average of 5% for OECD countries and by around 10% for the least efficient. The conclusions are similar for potential improvements in cost efficiency.

- Proportionally greater efficiency gains would accrue from reducing inputs while holding outputs constant, the potential saving in resources being around 19% on average, rising to over two-fifths for the least efficient -- equivalent to over 1% of GDP. One implication is that the potential returns from higher outputs need to be substantial to justify not saving resources for use elsewhere.

- There is a weak tendency for better PISA performance to be related to better cost performance but in a number of cases, for example Mexico, efficiency is relatively high but PISA scores are low. Decisions about the amount of resources devoted to education would in this case need to be made not only in terms of efficiency performance, but taking account of the possibly significant returns to investment in education, which this paper does not discuss.

5.   The paper comes to the following conclusions as to the statistical robustness of the findings:

- Statistical approaches that construct an efficiency frontier by defining this directly from the best-performing schools or countries (so-called non-parametric techniques, of which the most prominent is data envelopment analysis) permit quite robust inferences to be made about *relative* inefficiencies. They are subject to shortcomings with respect to possible measurement errors, but techniques for detecting outliers or sample biases can be used to estimate confidence intervals for individual units.

- Importantly, the estimates of potential efficiency gains are more certain when the estimated potential gain from eliminating inefficiency is greater.

- The efficiency rankings for schools have been compared with those derived from "stochastic frontier analysis", which directly accounts for "statistical noise" but requires the assumption of a specific functional relationship. Non-parametric techniques have been preferred to stochastic frontier analysis mainly due to their simplicity, their added attraction being that the uncertainty

surrounding the estimates of technical efficiency can be assessed for individual school or country observations.

6.   The structure of the paper is as follows: the first two sections discuss issues related to identifying the appropriate outputs and inputs in the education sector. Then follows a consideration of different approaches to measuring technical and cost efficiency and their relative strengths and shortcomings, as well as complementary approaches to ensure that efficiency measures are robust. The final section presents the results of estimates of technical efficiency at the national level and for school level data and then estimates of cost efficiency at the national level.

2.   Measuring public spending efficiency in the education sector

2.1.   Outputs vs. outcomes

7.   At its most basic level, output can be measured by "quantity" indicators such as course enrolment and completion rates, study duration, the level of education reached, or even equated with the quantity of inputs (Box 1). However, the narrow nature of academic exams has been criticised in that it may not link directly to welfare objectives or growth.[2] An approach which takes the quality of teaching (and learning) into account would focus more on outcomes, such as literacy rates at particular ages, learning achievement, or longer-term earnings and occupation and would give a better understanding of educational contributions to human capital.

8.   The use of PISA results for 15 year old students as a measure of the cumulative output of primary and lower-secondary education goes some way to treating outputs as educational outcomes. The PISA questions are designed to evaluate real-life aptitudes and not just academic attainment. Assessment at the age of 15 is designed to capture student abilities at the end of their compulsory schooling, thus avoiding the difficulties of comparing students across countries when participation is voluntary. Specifically, the measures adopted to capture achievement and equality outcomes -- the two main objectives of educational systems -- are as follows:

- The average level of education is derived as a synthetic indicator of the components of the 2003 PISA evaluations; results for the four PISA fields -- mathematics, reading, science, and problem solving -- are highly correlated. Principal component analysis suggests that a synthetic indicator that accords almost equal weighting to the results of the four tests would account for almost all the variance (see Annex 1).

- Homogeneity in educational attainment is measured as the ratio of the scores of the 25th to the 75th percentiles. Thus, the smaller the difference between the percentiles the higher the score.[3]

---

2.   See Hanushek (1986) for a discussion on the definition of outputs. Hanushek (2002) reviews the importance of teaching quality on educational outcomes. On the relation between measures of quality and growth see Barro (2002).

3.   This measure was adopted principally because it approximates most closely to a normal distribution (for the importance of this criterion, see Annex 1 for discussion of non-parametric techniques). Alternative measures were tested, but country rankings according to estimates of efficiency appear to be only marginally sensitive to alternative measures of homogeneity or dispersion. A variable assessing the proportion of students attaining basic competency in mathematics was also examined, though this variable did not appear to have a significant additional effect on results.

Box 1. **Measuring outputs from public spending on education in national accounts**

The convention applied in national accounts has been to equate the output of public education services with the cost of inputs ("output equals input") -- the approach adopted for all non-market outputs. Changes in the volume of education provided were derived from the value increase deflated by the relevant cost index and not measured directly. To overcome shortcomings of the traditional approach, the System of National Accounts (SNA93) recommended the development of new indicators to obtain volume measures of outputs, particularly for education and health. These volume measures, it recommended, should be based on an observable flow of service provision -- such as the teaching hours provided to students -- and not on outcomes that depend on additional factors.

More recently, Eurostat (2001) has defined output as the "quantity of teaching received by the students, adjusted to allow for the quality of the services provided for each type of education".[1] With regard to implementation, the preferred measure is the number of hours of teaching instruction received by pupils, though, in practice, if this is unavailable, the number of pupils taught could be used as the measure when teaching hours are relatively stable over time.

When the volume of educational output is defined in this way, a reduction in student numbers implies deterioration in productivity. However, lower class size may serve as an instrument for improving teaching quality and accounting for these changes is clearly needed in order to assess accurately changes in productivity of delivering education services.[2] There are two main approaches for doing this: one based on inputs and the other one based on outputs.

Using an input approach -- such as that chosen by Italy -- the measure of educational output is adjusted to reflect the implied difference in the teaching quality when class sizes change (Collesi, 2000).

The United Kingdom is one of the front-runners in implementing the "output-based" approach, a direct measure of education output introduced in 1998. The current measure reflects pupil attendance and a quality adjustment of +0.25% per year, based on past trends in exam results (*Atkinson Review*, 2005).

Although these approaches are conceptually more attractive than the pure input-based approach, they both contain arbitrary elements and in the absence of a common definition for the volume and quality of outputs, it becomes virtually impossible to compare productivity in the education sector across countries.[3]

_____

1. EU countries are required to introduce direct measures of output for certain government services including health care and education with the dissemination of 2006 national accounts.

2. Fraumeni *et al.* (2004) report, from a review of the literature, that class size is generally negatively correlated with academic outputs, though this is typically stronger for pupils in earlier grades.

3. The OECD has begun work aimed at harmonising national-accounts measures of output volumes in the education and health care sectors. This should eventually permit comparisons to be made over time and across countries.

---

9.   The use of PISA results does have some drawbacks. In a small number of OECD countries enrolment rates are low at that age (under 60% in Mexico and Turkey compared with an OECD average of 95%). In those cases, PISA mean scores may tend to overestimate the average level of human capital for the overall cohort. Cross-country comparisons could also be distorted by high truancy rates in some countries.[4] Related to these concerns is the fact that educational outcomes should also be judged by how many of the young continue their education to the end of the compulsory stage, which has obvious implications for human capital accumulation. However, the PISA questionnaire has been designed to collect background and institutional detail on a comparable basis, permitting the analysis of efficiency to be situated in a policy context.

10.   An additional complication is that national authorities may be targeting other outputs that are more difficult to take into account in the analysis. In particular, non-cognitive skills can be difficult to measure and particularly so in an international context. As such, differences in efficiency based on a

---

4.   A related problem is the exclusion of poorer performing students during exams which are important for school appraisal, though this should be less of a problem for the PISA study (see Figlio, 2006).

narrower set of outputs may be biased to the extent that the unmeasured outputs differ across countries and are uncorrelated with other outputs.

2.2. *Defining inputs*

11. Two main types of input determine educational outcomes. The first type covers discretionary factors under the control of the education system, such as teacher numbers, teacher-student ratios, class sizes, instruction time, teacher quality and, to a lesser extent, other resources in schools. The second covers so-called non-discretionary or environmental inputs.[5] On the discretionary input side, the main indicator used here is the ratio of teaching staff per 100 students. Estimates of the average class size are quite similar across most countries at the primary and lower secondary level, while variation in the more easily measured student/teacher ratio is somewhat greater (Figure 3, panel B). The data used in the empirical analysis are constructed from the PISA 2003 or the OECD's *Education at a Glance* databases, which are both designed so that data series are comparable across countries.

12. Estimates of cost efficiency are derived from estimates of cumulative spending on primary and lower education, which differ somewhat more among OECD countries than do physical inputs. As an illustration of the magnitude of the extent of cross-country differences, Luxembourg spent approximately eight times more than Mexico on the average 15-year old pupil in 2002 (Figure 3, panel A). International differences are due mainly to disparities in unit labour costs (employee compensation accounted for three-quarters of total expenditure on educational institutions on average in OECD countries in 2002 and teacher compensation for almost 60%). These disparities may reflect teaching quality[6] and the availability of other potentially important resources available in schools, including computer resources, as well as labour-market factors unrelated to efficiency.[7]

13. The indicators of student characteristics are taken as indices of the socio-economic background of the pupil, as well as the pupil's immigrant status or the language spoken at home. Typically, student achievement is considered to be dependent on family and peer-group effects and also innate ability. While difficult to measure, these factors are often proxied by measures of socio-economic status and in some cases indicators of immigrant or language status. Indeed, a pupil's socio-economic background, as measured using the index of economic, social and cultural status (ESCS) derived from the PISA 2003 study, appears to be strongly correlated with PISA scores (Figure 4).[8]

---

5. Other non-discretionary factors or characteristics, such as broader institutional settings, are sometimes considered to be inputs. Discussion of these factors, which may be amenable to policy determination, is contained in Gonand *et al.* (2007).

6. Card and Krueger (1992), in a study using labour market outcomes, suggest that important indicators of school quality are pupil-teacher ratios, longer school years, and higher relative teacher pay. Rivkin *et al.* (2005) find that teacher quality is an important determinant of student outcomes, but those characteristics such as possessing a postgraduate qualification or years of experience after the first few years are generally weak measures of teacher quality.

7. International comparisons of spending on education are complicated by the fact that relative prices vary with the level of development (discussed further below).

8. The PISA index of economic, social and cultural status (ESCS) is designed to capture broad aspects of a student's family and home background. It is derived from sub-indices based on: *i)* the highest occupational status of the student's parents; *ii)* the highest level of education of the parents; and *iii)* an index based on home possessions. Analysis of PISA results (OECD 2004, 2005b and 2005c; Hampden-Thompson and Johnston, 2006) demonstrate the importance of variables measuring students' socio-economic and language background on PISA results.

2.3   Approaches to measuring efficiency

14.   The fact that outputs in the public sector are amorphous and intangible in many respects makes it difficult to define a supply function in the conventional sense, while the fact that public sector organisations produce goods that are free at the point of use means that the prices of outputs are not determined by market forces. As economic efficiency cannot be directly measured, a technique is needed to proxy an efficiency frontier which would allow relatively accurate benchmarking. This section looks at two main approaches for achieving this: the first is a non-parametric technique and the second stochastic frontier analysis.

15.   In considering the choices that may apply with respect to reducing or sustaining educational resource inputs once efficiency shortfalls have been identified, the paper provides two perspectives on technical efficiency. The first is an *input oriented* measure which estimates how much inputs could be scaled back without reducing the level of outputs. The second -- the *output orientation* -- considers how much outputs could be boosted given current levels of inputs. Differences between the two occur because of the shape of the efficiency frontier, which may be subject to diminishing returns to scale after resource inputs reach certain levels.

Non-parametric approaches to measuring efficiency have advantages…

16.   Non-parametric approaches construct an efficiency frontier based on the input and output data from all the countries/schools of a sample. In essence, as shown in Figure 5, the frontier is constructed from the schools or countries that *envelop* the remaining observations and thus provides a benchmark by which the others can be judged. By assumption, the frontier determines best practice, and potential efficiency gains for specific countries or schools are measured by their position relative to the envelope. In the "one input-one output" case given in Figure 5, a measure of the efficiency shortfall in terms of unachieved output is given by the ratio of a school's output to the output on the frontier for the same level of inputs (*i.e.* the point on the frontier vertically above the school/country observation). Conversely, the ratio of inputs on the frontier to the school's inputs at the same output (measured horizontally) is a measure of inefficiency in terms of potentially excess inputs. In the case of multiple inputs or outputs, the measures of efficiency are determined in a similar fashion by holding the relative proportions of either inputs or outputs constant in measuring the distance to the frontier. Countries or schools can then be benchmarked on the basis of potential efficiency gains compared to the measures of best practice.

17.   Some arbitrary assumptions are needed to define the shape of the frontier, which can lead to different measures of efficiency losses. In particular, the results are sensitive to whether the frontier is characterised by constant or variable returns to scale (see Box 2).

18.   A considerable attraction of non-parametric approaches to estimating efficiency is their relative transparency. They also permit the identification of the number of schools, or countries, in the sample that produce at least as much output with less or the same amount of input. Generally, the larger the number of other schools with higher levels of efficiency the more relevant and reliable the benchmark is likely to be. Non-parametric approaches to measuring efficiency are thought to have two additional advantages. Most

Box 2. Assumptions about the shape of the efficiency frontier

Data Envelopment Analysis (DEA) and Free Disposal Hull (FDH) are two related methods for non-parametric estimation of efficiency frontiers. The techniques measure efficiency relative to the observed most efficient units. In essence, the non-parametric approaches are constrained optimisation problems.
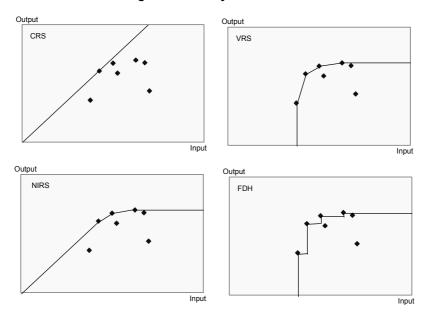
The aim of DEA and FDH is to construct a frontier such that all observations lie either on or within the frontier. The shape of the DEA efficiency frontier depends on the assumptions about returns to scale:

- *Constant returns to scale (CRS)*. This assumption describes the efficiency frontier as a ray from the origin through the observation(s) with the highest output/input ratio (Box figure, upper left panel). One consequence of this assumption would be that the efficiency estimate will be the same whether measured as excess inputs or deficient outputs. Constant returns to scale are, however, an extreme assumption in practice, with the implication that all observations are operating at their technically most efficient scale.

- *Variable returns to scale (VRS)*. This approach identifies the schools or countries that define the frontier by starting from the observations of units that use the least of each input and ending with the observations producing the highest amount of each output (upper-right panel). This approach assumes convexity, so that for the one input/one output case the points that define the frontier are linked by straight lines.

- *Non-increasing returns to scale (NIRS)*. This assumption combines the constant returns to scale assumption between the origin and the observation with the highest output/input ratio, and variable returns to scale thereafter (lower left panel).

The free disposal hull (FDH) approach is intimately related to the variable returns to scale DEA analysis (lower right panel). It makes no assumptions about convexity, but instead constructs the efficiency frontier in a stepwise fashion between the units identified to be on the frontier. As a result, more units are likely to lie on the FDH frontier and estimates of inefficiency are likely to be smaller than DEA estimates. Comparison with the results from DEA estimates can give an insight into the importance of the assumption of convexity.

By comparing the results under the different assumptions about returns to scale it is possible to identify whether the school or country would be at its technically most efficient scale if inefficiency was eliminated. In the one input, one output case shown in the figure, schools or countries on the efficiency frontier where it flattens out under the variable returns to scale would be operating on a portion of the efficiency frontier characterised by diminishing returns to scale. On the vertical portion of the variable returns to scale efficiency frontier, schools or countries could potentially exploit increasing returns to scale.

Figure. Efficiency frontiers

importantly, they do not require assumptions about the *specific* functional form of the educational production function, although they do assume that the production function is common to all units. In addition, the framework can easily accommodate multiple outputs, which is a particularly attractive feature for public services. Exploiting these attractions, non-parametric measures have been employed in a number of studies, looking at relative school performance, both within a country and across countries, as well as at the aggregate national level.[9]

...and drawbacks

19.   Non-parametric approaches have three main drawbacks:

- First, non-parametric estimates of efficiency are particularly sensitive to measurement error, statistical noise and outliers.[10] An observation that determines a segment of the efficiency frontier erroneously will in turn affect the measure of inefficiency for all the schools or countries lying within that segment.

- Second, small-sample bias -- omission from the sample of the relevant school/country with best practice -- may lead to the underestimation of the degree of inefficiency. This problem is aggravated as the number of inputs and outputs rises.

- Third, omitted or irrelevant inputs and outputs will affect estimates of inefficiency. Omitted variables will tend to lead to estimates of inefficiency that are larger than "true" inefficiency, while irrelevant variables will result in estimates of inefficiency that are too small.

Statistical measures can address some concerns

20.   In DEA analysis, two main statistical approaches can be adopted -- and have been adopted in this paper -- to address problems of measurement error and sampling bias. Potential outliers can be identified by using a so-called "jackknife" procedure to conduct empirical influence tests as part of exploratory data analysis and data cleaning. This method assesses the influence on the estimated statistics of sequentially dropping each observation and can thus help identify particular observations that have a large effect on the measures of efficiency. Confidence intervals and bias correction for efficiency scores can be derived from the bootstrap procedure, which is a statistical method for estimating the sampling distribution of an estimator (See Annex 1).[11]

21.   There are two additional methodological issues that complicate the non-parametric approach. First, the inclusion of non-discretionary inputs can lead to schools or countries operating in the most adverse environments being assessed as more efficient. Secondly, and perhaps more importantly, the results can depend heavily on the choice of inputs and outputs considered, but there is no *a priori* means of choosing the correct number to include. For example, in the DEA context, as the number of outputs rises, measured efficiency will increase and in the extreme all schools or countries will be regarded as 100% efficient.[12] There is thus an onus on limiting the number of dimensions.[13] The approach adopted in this

---

9.   See, respectively, Bessent *et al.* (1982); Wilson (2005a) and Afonso and St Aubyn (2005).

10.   See Greene (1997); Coelli *et al*. (2003) and Murillo-Zamorano (2004).

11.   This approach follows Simar and Wilson (2000) and Wilson (2005a).

12.   The intuition behind this is that if outputs are specified very finely there will eventually be one output unique to each organisation, which it will, by definition, be uniquely efficient at producing.

paper has been to estimate a baseline specification and then augment the analysis with variables identified through stochastic frontier analysis.

Stochastic frontier analysis provides a check on the DEA results

22.   Stochastic frontier analysis uses statistical methods to fit a frontier, as shown in Figure 6. By making assumptions about the distribution of inefficiency, the regression error term can be decomposed into statistical noise and a measure of inefficiency. Consequently, stochastic frontier analysis is less sensitive to the influence of measurement error. However, this feature comes at the cost of having to specify the functional form and making assumptions about the distribution of inefficiency (Annex 1), which can be an important drawback given their importance in determining the degree of estimated inefficiency. Stochastic frontier analysis is used below to provide a check on the robustness of the DEA results and assess the significance of variables included in the analysis.

3.   Estimates of technical efficiency

3.1   *Coverage and specification*

23.   Although non-parametric methods have been used in many studies of efficiency in the education and health care sectors, the approach adopted here extends previous research by conducting the analysis at both the national and school levels. At the school level the much larger sample reduces the sensitivity of the results to small-sample bias that may affect the national comparisons.[14] In addition, school-level analysis provides insights into the extent that inefficiency varies within countries, which the analysis of average or median results is unable to capture, and can help identify how efficiency varies across different categories of schools (*e.g.* between small and large schools or public and private schools).

24.   At the school level, the large sample allows DEA results to be tested against a stochastic frontier approach and specification. The large sample permits a broad specification of inputs, which includes the ratio of teaching staff per 100 students, a proxy for capital inputs based on the availability of computer resources and indicators of student characteristics,.[15] The indicators of student characteristics are the index of the socio-economic background of the pupil, and indices of immigrant status or the language spoken at home.[16]

---

13.   This has made for a number of rules of thumb, such as limiting the number of inputs and outputs (combined) to under a third of the number of observations. One approach has been to assess whether the inclusion or exclusion of a potential variable has an impact on a sufficient number of observations that exceeds a predetermined threshold (Pastor *et al.*, 2002). However, this test can be sensitive to the distributions of the variables.

14.   At the same time, school-level analysis may be affected by endogeneity: if higher ability students select more efficient schools (at least in the output direction) the measure of efficiency may be biased.

15.   Input indicators were also constructed based on the minutes of instruction time per week and teacher qualifications. Specification tests did not suggest these were influential for the DEA results.

16.   As the non-parametric techniques do not work with zero or negative numbers, the ESCS index was rebased on the sub-index of highest level of educational attainment of the parents. Immigration and language indices were constructed by assigning a value of 1 if the student was not assessed to be an immigrant or if a native language was spoken at home and 0 otherwise. The school index is the average of the pupil scores. As a result, schools that have only pupils that are immigrants or speak a non-native language at home are excluded from the analysis. This index is only weakly correlated with the school level socio-economic index (a correlation coefficient of 0.11).

25.   To provide a point of comparison between the country and school results, the paper reports a baseline specification of technical efficiency at both levels. Due to the limited number of observations at the national level, the specifications have been kept as simple as possible in order to minimise the problems associated with small samples. The baseline specification includes the synthetic indicator of the PISA scores as a measure of the output and two variables as inputs, one measuring teaching resources and the other socio-economic background of parents (see above). However, the preferred specification at the national level is actually a broader measure, with a second output assessing the homogeneity of PISA scores as a measure of the attainment of equity objectives. The results from the addition of the school enrolment rate as a third output are also reported (Annex 2).

3.2   *Technical efficiency at the school level*

Stochastic frontier results

26.   As noted above, the larger dataset at the school level allows a wider set of inputs to be considered in the estimates of technical efficiency, including capital inputs and language spoken at home which acts as a measure of the difficulties facing children of immigrants and minority groups in education.[17] A translog (transcendental logarithmic) production function was estimated for the stochastic frontier analysis. This production function allows greater flexibility, including non-constant returns to scale and varying input elasticities across schools.[18] The basic results are presented in Table 1, which presents the results for two specifications with different assumptions about the distribution of inefficiency (half-normal and exponential). The results suggest a significant role is played by student socio-economic background in determining performance in the PISA tests. Student performance also increases with teacher numbers, though with diminishing returns. Capital inputs -- as proxied by computer availability -- make a positive contribution to student performance only in conjunction with teaching inputs. The efficiency estimates for the median school in each country and a measure of the variation of efficiency across schools within countries are presented in Table 2.[19]

Data Envelopment Analysis results

27.   Data Envelopment Analysis was conducted using the same inputs and outputs as identified in the stochastic frontier analysis. Overall, the estimates of school-level efficiency are highly correlated with the parametric estimates from stochastic frontier analysis The DEA estimates of inefficiency are larger than the parametric estimates (Figure 7), which is often a feature of comparison of the two approaches though also the result of a richer specification in this case. Nonetheless, the different estimates of inefficiency are highly correlated, with a rank correlation of 0.9.

28.   The main results of the DEA are:

---

17.   Although similar results are derived if alternative proxy variables are used, such as measures of immigrant status, the language indicator emerges as the preferred variable after comparing the variables under stochastic frontier analysis estimation and using tests for variable relevancy in the DEA framework.

18.   The estimation included quadratic and cross-product terms for teaching inputs and computer inputs (serving as a proxy variable for capital inputs), and also controlled for the socio-economic and language background of the pupil. The likelihood ratio test also supported this functional form over the Cobb-Douglas production function. The final specification was partly determined by problems of multicollinearity encountered in an initial more general specification. Multicollinearity is a frequent problem for translog estimation.

19.   The translog production function does not satisfy globally the regularity conditions of monotonicity and concavity (the partial elasticities of each input are positive and the bordered Hessian matrix is negative definite, respectively). These conditions were checked and found not to hold for each observation.

- The scope for reducing inputs while holding outputs constant is around one-third for the median school (Table 3). Potential gains from maximising outputs from the current level of inputs are slightly smaller. The average PISA scores of students in the median school is around one-fifth below the level suggested possible by the efficiency frontier.[20] To put this into perspective, these potential gains if realised could be worth in excess of an additional year's tuition, with positive implications for growth.[21]

- Differences in estimates of efficiency for different types or sizes of school tend to be modest. The median public school in the overall sample is slightly less efficient than both the median government-dependent private school and median independent private school.[22] Schools that rely on public sources for the majority of their funding also tend to be slightly less efficient than other schools. Small schools tend to be less efficient than larger schools, particularly in the input direction.

- There are significant differences across countries (Figure 8). The estimates suggest that the potential to eliminate inefficiency by reducing inputs, while holding outputs constant, or increasing outputs, holding inputs constant, are substantial for Greece, Hungary, Iceland, Luxembourg, Norway and the United States. The potential to eliminate inefficiency in the median schools in Korea and Japan appears comparatively modest. For the median school, the uncertainy around the point estimate is noticeable in the case of Turkey.

- There are significant differences in the estimates of technical inefficiency within some countries, though not generally (Figure 9). For example, there is considerably more heterogeneity in school-level technical efficiency in Japan than in Ireland -- both countries where the median level of inefficiency is relatively low.

- The estimated technical inefficiency is generally much larger with the constant returns to scale assumption than with the other assumptions. This is consistent with most schools being clustered along the portion of the efficiency frontier characterised by diminishing returns to scale.[23] In the majority of cases, eliminating inefficiency by achieving the maximum possible outputs from available inputs would locate schools on this portion of the efficiency frontier. However, scaling back inputs to eliminate technical inefficiency would leave a significant portion of schools operating below their most efficient scale in Greece, Hungary and Turkey. (Figure 10).

29.   The non-parametric approaches can identify how many schools can produce at least as many outputs with fewer or the same amount of inputs. Figure 11 presents information on the average number of "reference" schools, which ranges from just under 20 schools in Turkey and Germany to over 600 schools in Iceland and Norway. This information complements the DEA results, by indicating that although

---

20.   The interpretation of the input orientation results is complicated as schools cannot alter the non-discretionary inputs, at least in the short run. This complication does not affect output oriented results. The convexity assumption is quite important in determining the level of inefficiency, the free disposal hull efficiency estimates suggesting that for the median school the portion of the efficiency frontier that provides the benchmark is not well defined by actual observations.

21.   OECD (2004) estimates that an additional year of tuition is equivalent to 41 points in the mathematics test The long-run effect of this could raise output *per capita* by around 6% (Bassanini and Scarpetta, 2001).

22.   However, within countries the differences between different types of schools can be more marked. For example, the median private school in the United Kingdom is noticeably more efficient in the output direction (*i.e.* higher outputs for a given level of inputs) than the median public school.

23.   See Hanushek (2003) on returns to scale in US schools.

schools in Greece, for example, are among the least efficient there are fewer schools in comparison with Norway and Iceland transforming fewer inputs into more outputs. This may indicate that identifying relevant best practice may be more difficult for Greek schools.

30.   The implications of aggregate resource savings or aggregate gains in PISA scores are presented in Figure 12. The savings or gains assume that schools attain either at least the degree of technical efficiency of the school at the 95th percentile (as special factors may account for the most efficient schools) or the 50th percentile.[24] The implications in terms of staffing levels given the student teacher ratios and increases in PISA scores are then computed for each school. Overall, the potential gains are relatively modest. In the case of raising performance to the level of that achieved by the 95th percentile, the potential for resource savings is still substantial in Belgium, Greece and Italy, while the gains in PISA scores are potentially substantial in the Czech Republic and Turkey. In the case of Belgium (the French community) and Greece, even raising performance to that of the median would yield sizeable gains.

3.3   *Estimates of technical efficiency from country data*

31.   At the national level the baseline and equity-augmented specifications are reported assuming non-increasing returns to scale.[25] The small-sample bias is likely to be more pronounced when there are few countries with similar input and output mixes, making estimates of the portion of the efficiency frontier that is relevant for these observations less certain. The techniques used to estimate technical efficiency can also shed some light on the uncertainty surrounding the estimates. As discussed above (and in Annex 1), statistical procedures can make a correction for the small-sample bias and generate confidence intervals around the point estimates. In this light, bias-corrected estimates and the 95% confidence intervals are also reported.

32.   For the purpose of comparison, a baseline specification was also estimated for school-level data based on a common sample (Figure 13).[26] The most notable difference between the results is that estimates of inefficiency are generally much larger at the school level, a result which is the natural outcome of the DEA method given the greater variation in the school data. The best schools will be more efficient than the average, which will generate larger estimates of inefficiency, whereas, at the national level, aggregation masks the within-country variation. However, the relative rankings are broadly consistent with the national results.

33.   The estimates of technical efficiency lead to the following conclusions:

- The potential for efficiency gains is around 19% on average and largest in Greece, Iceland, Luxembourg, Norway and the United States (Figure 14, panel A). The uncertainty surrounding the efficiency estimates is often considerable and much larger than that surrounding school-level

---

24.   The calculations also exclude schools below the 5th percentile, as the most inefficient schools may be affected by special factors that may make realisation of greater efficiency problematic.

25.   Preliminary testing suggested Korea may be an outlier. As there may be a case to consider it unusual because of the degree of private tuition, results excluding Korea are presented in Annex 2 for comparison.

26.   The overall rank correlation between the estimates for the baseline and wider specification is high, at around 0.9. Separate estimates of school-level efficiency were also made for individual country samples. While the estimates of inefficiency were generally lower -- a feature expected from using smaller samples -- there was little change in the relative rankings.

results, particularly for countries that are close to or define the efficiency frontier (see also below).[27]

- The scope for improving average PISA scores while holding inputs constant is much smaller: on average about 5% relative to the efficiency frontier, rising to around 10% for the least efficient (Figure 14, panel B), though this is in part driven by the assumption about the shape of the frontier.

- At the national level technical efficiency measures are broadly similar for the baseline case and the wider specification (Figure 15) when considering how much inputs could be cut while holding outputs constant.[28] There are some differences in ranking when the equity objective is also considered, though broadly speaking, countries that are inefficient when considering one output also tend to be relatively inefficient when considering two outputs. In general, countries which are the least efficient in terms of input efficiency also tend to be rather inefficient in terms of output efficiency.

- Some countries, particularly Ireland, are relatively more efficient at ensuring that educational attainment is evenly spread rather than in securing high levels of human capital alone. On the other hand, countries that remain close to the diagonal, such as Germany, appear relatively less efficient when the equity objective is included as an output in the specification.[29] Taking enrolment rates into account as a third output to assess broader equity concerns generally has a relatively modest effect on the estimates of efficiency (Annex 2, Tables A2.7 and A2.12).

34.   The potential gains to be made by eliminating inefficiency imply substantial possible improvements in performance or, alternatively, important resource savings. For example, the resource gains which would accrue if teacher-student ratios were reduced while holding outputs constant would largely be in the 0.2-0.4% of GDP range, but in some cases, notably Norway, savings could be substantially greater (Figure 16). For those countries where the frontier is subject to diminishing returns to scale, the marginal rates of return from investment in education would need to be quite high to justify maintaining resources while improving outcomes.

35.   However, in a small number of countries, notably Mexico, Turkey and, to a lesser extent, Portugal, estimated efficiency is relatively high -- partly as a result of few other countries with similar input and output mixes -- but the average performance in the PISA tests is comparatively poor (Figure 17). These countries, while apparently relatively efficient using country-level data, might nevertheless gain by improving their PISA performance further. Excluding Mexico, Turkey and Portugal from the analysis leaves most other estimates relatively unchanged, as their relevant "best practice" is defined by countries with different input mixes (Figure 18).

---

27.   While Turkey emerges as one of the most efficient countries, these results should be treated with caution. Turkey has very high drop-out rates for its students: only 54% of 15-year-olds are offered education compared with an OECD average of 95%.

28.   The results at the national level are for both public schools and government-dependent private schools; the results for public schools alone are very similar, but some countries, notably New Zealand and Spain, move closer to the efficiency frontier.

29.   If public schools only are taken into account, the average cross-country input and output oriented efficiency levels remain unchanged. But this hides some country differences, notably for Belgium (the Flemish community), where efficiency estimates deteriorate, and New Zealand, where efficiency estimates improve.

36.   A comparison of country- and school-level estimates of technical efficiency makes clear that the high efficiency scores from the country-level analysis are an artefact of the small sample size. Portugal and Turkey appear to be among the most efficient when comparing national averages but at the school level the efficiency of the median school is substantially worse relative to other countries. The comparison of schools can identify different schools both within Portugal and Turkey and across the rest of the sample that have similar input mixes that are also more efficient, whereas within the country sample there are few other countries with similar input mixes. This suggests that caution is required in assessing the relative efficiency using aggregate country-level data for those countries close to or defining the frontier.
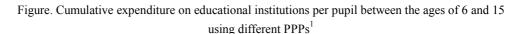
4.   Cost efficiency

37.   It is possible to estimate cost efficiency at the national level, complementing the estimates of technical efficiency. A country may be technically efficient but may be cost-inefficient if it fails to contain the relative costs of inputs. As noted above, higher wage costs may be associated with higher-quality teaching; alternatively, when governments exercise monopsonistic power in hiring teachers the result may be low cost teachers which attract poorer quality teachers. Technical efficiency may well be endogenous to costs. Subject to that caveat, this section examines measures of cost efficiency at the national level to complement the preceding analysis.

38.   Assessing cost efficiency is, in practice, more complicated than the analysis of technical efficiency. Both in constructing the relevant indicator of spending and comparing spending data across countries, particular care is needed in measuring inputs:

- Cost efficiency depends on cumulative spending on education per student between the ages of 6 and 15, this being calculated from total spending per full time equivalent student between 1993 and 2002, taking into account the progress of the student through pre-primary, primary and lower secondary education, depending on the education system of the country under consideration.

- Typically as *per capita* income increases so does the price of non-tradeables, such as education services. Failure to take account of this phenomenon will thus exaggerate the differences in spending between richer and poorer countries (Box 3). Here, the estimates of cumulative education spending per student have been converted into constant prices using the private consumption deflator and then into constant 2002 dollars using the PPP for private consumption, which may not eliminate this possible bias.

- In order to check the robustness of the purchasing power parity used in the analysis, additional estimates of cost efficiency were made (using spending on pupils in primary and secondary school in a single year) converted into constant dollar prices using PPPs for the service sector and total output (Table 4).

39.   The cost efficiency estimates use almost the same baseline and wider specification as those of technical efficiency at the national level; though the teacher-pupil variable is replaced with cumulative spending. Bearing in mind the caveats just described, the main results from the estimates of cost efficiency are:
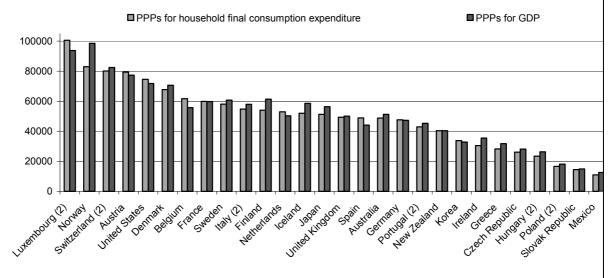
Box 3. **Comparing spending across countries**

The comparison of cost efficiency across countries is complicated by differences in relative prices in richer and poorer countries. The so-called "Balassa-Samuelson effect" suggests that the relative price of non-tradable goods (such as many services) tends to rise with *per capita* income, which is particularly pertinent for the comparison of spending on education. In this regard, the differences between high and lower income countries may be exaggerated when evaluated at market exchange rates or even when using purchasing power parities that include prices for tradable goods.

The translation of spending on education for each country into a common currency could use several different purchasing power parities (PPPs).

- PPPs that are education specific or relate to general government consumption (*i.e.* mainly education and health care) would be preferable. However, for the time being these PPPs are still largely input based rather than output based and therefore do not account for differences in productivity between countries. Furthermore, PPPs that include health care prices may be hard to interpret due to technical change in the sector and different regulatory regimes across countries.

- An alternative approach would be to use more aggregate PPPs. The PPP for the service sector is a possible candidate, though the weight of activities that bear little relation to education (for example, the rental sector, transport, water supply, etc.) suggests that little is to be gained over using the total economy PPP. However, the total economy PPP can also be problematic for some countries, such as Norway where a large share of total output is in the petroleum sector.

- In light of these complications, the approach adopted here was to use the PPPs for household final consumption, approach also used in the OECD's *Education at a Glance*. Comparisons are made on the basis of constant prices and constant purchasing power parity for private consumption. The estimates for cumulative expenditure on educational institutes are quite similar to those obtained with the PPPs for GDP (See Box Figure below).

Figure. Cumulative expenditure on educational institutions per pupil between the ages of 6 and 15 using different PPPs[1]

1. PPPs are expressed in national currencies per US dollar (US=1.00)
2. Public institutions only.
Source : Purchasing Power Parities and Real Expenditures, OECD-Eurostat; OECD (2005a); OECD calculations.

- Overall, the potential gains from eliminating cost inefficiencies are similar to those from eliminating technical inefficiency: inputs could be cut by one-sixth on average, with potential cuts of over one-quarter in Iceland, Luxembourg, Norway and the United States. Output improvements from the same level of inputs reach a maximum of approximately 15% (Figure 19, panels A and B).

- The uncertainty surrounding the cost efficiency estimates is quite similar to the technical efficiency measurement with greater uncertainty around the cost efficiency estimates for countries that are close to or define the frontier.

- The estimates of efficiency can be somewhat sensitive to the choice of deflator for cross-country comparison for some countries. For example, estimates of efficiency for the Czech Republic swing markedly when different purchasing power parties are used. Nevertheless, the estimates of efficiency are relatively stable in most countries (Table 4).

- As in the case of technical efficiency, some countries experience a relative deterioration in their performance when the outputs include the achievement of homogeneity in pupil test results (*e.g.* those remaining close to the diagonal line in Figures 20).[30] However, in general, the worst performers with respect to average PISA scores also generally emerge less favourably *vis-à-vis* the homogeneity of results.

- For some economies (notably the Czech Republic, Poland, the Slovak Republic and to a lesser extent Hungary), potential efficiency gains are noticeably lower than they are in terms of technical efficiency. This may be a reflection of teaching compensation (or quality) relative to other countries.

40.   Comparing cost efficiency estimates with PISA scores reveals only a weak tendency for better PISA performance to be related to better cost performance (Figure 21). This is not uniformly the case, however, and in some cases, notably Mexico and to a lesser extent Portugal, efficiency appears to be relatively high but PISA scores low by international comparison. In part this reflects the small sample bias discussed above, but the contrast also underscores the importance of judging the need for education spending not just in terms of the static cost savings which have been the subject of this paper but also the impact that extra education resources can have in terms of overall economic performance.

---

30.   When looking only at public schools, the country ranking on both the input and output scale is very similar. However, there are some differences, particularly Belgium (the Flemish community) and the Netherlands where public schools appear to be relatively less efficient.

BIBLIOGRAPHY

Afonso, A. and M. St. Aubyn (2005), "Cross-Country Efficiency of Secondary Education Provision: A Semi-Parametric Analysis with Non-discretionary Inputs", *ECB Working Paper*, No. 494.

Atkinson Review (2005), *Final Report: Measurement of Government Output and Productivity for the National Accounts*, Palgrave McMillan.

Banker, R., S. Janakraman and R. Natarajan (2002), "Evaluating the Adequacy of Parametric Functional Forms in Estimating Monotone and Concave Production Functions", *Journal of Productivity Analysis*, Vol. 17, Nos. 1-2.

Barnett, R., J. Glass, R. Snowdon and K. Stringer (2002), "Size, Performance and Effectiveness: Cost-Constrained Measures of Best-Practice Performance and Secondary-School Size", *Education Economics*, Vol. 10.

Barro, R. (2002), "Education as a Determinant of Economic Growth" in E. Lazear, (ed.) *Education in the Twenty-first Century*, Hoover Institution Press.

Bassanini, A. and S. Scarpetta (2001), "Does Human Capital Matter for Growth in OECD Countries? Evidence from Pooled Mean-Group Estimates", *OECD Economics Department Working Papers*, No. 282.

Bessent, A., W. Bessent, J. Kennington and B. Reagan (1982), "An Application of Mathematical Programming to Assess Productivity in the Houston Independent School District", *Management Science*, Vol. 28, No. 12.

Card, D. and A. Krueger (1992), "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States", *Journal of Political Economy*, University of Chicago Press, Vol. 100.

Coelli, T. (1996), "A Guide to DEAP Version 2.1: A Data Envelopment Analysis (Computer) Program", *Centre for Efficiency and Productivity Analysis Working Paper*, No. 96/8, Department of Econometrics, University of New England, Armidale NSW, Australia.

Coelli, T., A. Estache, S. Perelman, and L. Trujillo (2003), *A Primer on Efficiency Measurement for Utilities and Transport Regulators*, World Bank Publications.

Collesi, D. (2000), "Volume Measures and Productivity Analysis for the Non-Market Sector: The Italian Experience", Paper prepared for the 15th meeting of the Voorburg Group on Services Statistics, September 2000, Madrid, Spain.

Eurostat (2001), *Handbook on Price and Volume Measures in National Accounts*, European Commission, Luxembourg.

Figlio, D. (2006), "Testing, Crime and Punishment", *Journal of Public Economics*, Vol. 90.

Figlio, D. (1999), "Functional Form and the Estimated Effects of School Resources", *Economics of Education Review*, Vol. 8.

Fraumeni, B., M. Reinsdorf, B. Robinson, and M. Williams (2004), "Price and Real Output Measures for the Education Function of Government: Exploratory Estimates for Primary and Secondary Education", Presented at the Conference on Research in Income and Wealth, Conference on Price Index Concepts and Measurement, June 29, 2004, Vancouver, BC, Canada.

Gonand, F., I. Joumard and R. Price (2007), "Public Spending Efficiency: Institutional Indicators in Primary and Secondary Education" *OECD Economics Department Working Papers*, No. 543.

Greene, W. (1997), "Frontier Production Functions" in H. Pesaran and P. Schmidt, (eds.), *Handbook of Applied Econometrics.* Vol. 2*, Microeconomics,* Blackwell.

Grosskopf, S., K. Hayes, L. Taylor and W. Weber (1997), "Budget-Constrained Frontier Measures of Fiscal Equality and Efficiency in Schooling", *Review of Economics and Statistics*, 74.

Hampden-Thompson, G. and J. Johnston (2006), "Variation in the Relationship between Nonschool Factors and Student Achievement on International Assessments", *National Center for Education Statistics: Statistics in Brief*, NCES 2006-014.

Hanushek, E. (2003), "The Failure of Input-Based Schooling Polices", *The Economic Journal*, 113.

Hanushek, E. (2002), "Publicly Provided Education", *NBER Working Paper*, No. 8799.

Hanushek, E. (1986), "The Economics of Schooling: Production and Efficiency in Public Schools", *Journal of Economic Literature*, Vol. 24.

Murillo-Zamorano, L. (2004), "Economic Efficiency and Frontier Techniques", *Journal of Economic Surveys*, Vol. 18.

OECD (2004), *Learning for Tomorrow's World: First Results from PISA 2003*, Paris.

OECD (2005a), *Education at a Glance*, Paris.

OECD (2005b), *School Factors Related to Quality and Equity: Results from PISA 2000*, Paris.

OECD (2005c), *Problem Solving for Tomorrow's World: First Measure of Cross-Curricular Competencies from PISA 2003*, Paris.

OECD (2005d), *PISA 2003 Data Analysis Manual: SAS® Users*, Paris.

Pastor, J., J. Ruiz, and I. Sirvent (2002), "A Statistical Test for Nested Radial DEA Models", *Operations Research*, 50.

Rivkin, S., E. Hanushek, and J. Kain (2005), "Teachers, Schools, and Academic Achievement", *Econometrica*, Vol. 73.

Simar, L. and P. Wilson (2000), "Statistical Inference in Nonparametric Frontier Models: The State of the Art", *Journal of Productivity Analysis*, Vol. 13.

Wilson, P. (1993), "Detecting Outliers in Deterministic Nonparametric Frontier Models with Multiple Outputs", *Journal of Business & Economic Statistics,* 11.

Wilson, P. (2005a), "Efficiency in Education Production among PISA Countries, with an emphasis on Transition Economies", *mimeo*.

Wilson, P. (2005b), "FEAR: A Package for Frontier Efficiency Analysis with R", *Economics Department*, *University of Texas*, Austin.

Wößman, L. (2005), "The Effect Heterogeneity of Central Examinations: Evidence from TIMSS, TIMSS-Repeat and PISA", *Education Economics*, Vol. 13.

TABLES AND FIGURES

Table 1. Stochastic frontier specification

	Half normal distribution		Exponential distribution	
	Coefficient	z	Coefficient	z
Constant	5.064	205.54	5.041	209.97
Teachers student ratio	0.044	2.91	0.041	2.77
Computer availability	-0.008	-1.26	-0.007	-1.14
Socio-economic background	0.500	74.05	0.498	75.78
Language background	0.109	15.88	0.108	15.95
Teachers student ratio * teachers student ratio	-0.055	-6.93	-0.051	-6.50
Computer availability * computer availability	-0.012	-6.22	-0.011	-6.10
Teachers student ratio * computer availability	0.021	6.18	0.019	5.90
Lambda	1.958		0.990	

Note: The table presents the estimated coefficients and assoicated 2-scores for trans-log production funtions, assuming different inefficiency distributions. A positive coefficient signifies a positive relationship with average performance in PISA tests.

Table 2. Stochastic frontier estimates of technical efficiency at the school level

Preferred specification[1]

	Exponential		Half-normal	
	Median	Range 90%	Median	Range 90%
Australia	0.96	0.05	0.94	0.08
Austria	0.95	0.15	0.93	0.17
Belgium	0.97	0.11	0.95	0.14
Belgium Flemish c.	0.97	0.13	0.95	0.15
Belgium French c.	0.97	0.11	0.95	0.13
Canada	0.95	0.11	0.92	0.13
Czech Republic	0.96	0.09	0.93	0.12
Denmark	0.95	0.08	0.92	0.10
Finland	0.97	0.04	0.95	0.06
Germany	0.96	0.11	0.93	0.13
Greece	0.93	0.23	0.90	0.22
Hungary	0.95	0.21	0.92	0.21
Iceland	0.93	0.12	0.89	0.13
Ireland	0.96	0.05	0.94	0.07
Italy	0.96	0.17	0.94	0.18
Japan	0.97	0.09	0.96	0.13
Korea	0.97	0.05	0.96	0.09
Luxembourg	0.95	0.07	0.91	0.08
Netherlands	0.96	0.08	0.94	0.11
New Zealand	0.96	0.07	0.94	0.10
Norway	0.91	0.13	0.87	0.13
Poland	0.96	0.05	0.94	0.08
Portugal	0.97	0.09	0.95	0.12
Slovak Republic	0.95	0.10	0.93	0.12
Spain	0.94	0.11	0.94	0.10
Sweden	0.95	0.06	0.92	0.09
Switzerland	0.96	0.09	0.94	0.12
Turkey	0.96	0.15	0.93	0.17
United Kingdom	0.96	0.07	0.93	0.10
United States	0.93	0.18	0.89	0.17
Standard deviation	0.01	0.05	0.02	0.04
Median	0.96	0.10	0.94	0.12
Average	0.95	0.11	0.93	0.12

Note: The table gives the estimate of the level of efficiency for the median school and the range that encompasses 90% of schools in a particular country.
1. PISA score as output and teachers student ratio, computer availability, socio-economic and language backgrounds as inputs.

Table 3. Estimates of technical efficiency for the median school:
aggregate non-parametric results

	Input orientation	Output orientation	Number of schools
DEA, non-Increasing returns to scale	0.693	0.782	6204
Of which			
Public schools	0.689	0.777	4834
Government-dependent private schools	0.715	0.805	672
Independent private schools	0.684	0.799	194
Public funds > 50 %	0.693	0.780	5469
Public funds < 50%	0.693	0.803	397
Small schools	0.669	0.770	3102
Large schools	0.712	0.794	3102
DEA, variable returns to scale	0.692	0.782	6204
DEA, constant returns to scale	0.685	0.681	6204
Free Disposal Hull	0.920	0.875	6204

Note : DEA efficiency scores are robust estimates. Sub-categories may not sum up to 6204 due to missing data.
Source : Secretariat calculations based on PISA 2003 database.

Table 4. Estimates of potential resource savings and output gains using different PPPs

	Resource savings				Gains in PISA scores			
	PC over time (1)	PC (2)	CS (3)	GDP (4)	PC over time (5)	PC (6)	CS (7)	GDP (8)
Australia	15.8	15.4	15.5	15.4	4.6	4.7	4.8	4.7
Austria	15.2	15.1	15.1	15.1	6.0	5.9	6.0	5.9
Belgium Flemish c.	14.3	14.3	14.5	14.3	2.0	1.9	1.9	1.9
Belgium French c.	18.0	17.9	17.8	17.9	9.8	9.7	9.7	9.7
Canada	..	21.8	21.9	21.8	..	3.8	3.8	3.8
Czech Republic	14.9	10.0	15.4	7.4	4.3	2.5	4.6	2.1
Denmark	20.3	19.9	19.9	19.9	4.8	4.9	4.9	4.9
Finland	9.9	10.1	10.2	10.1	1.5	1.8	1.9	1.8
France	9.4	9.2	9.3	9.2	5.5	5.5	5.6	5.5
Germany	17.1	17.5	17.6	17.5	9.0	9.0	9.2	8.9
Greece	20.2	20.2	19.8	20.2	7.5	7.5	7.6	7.3
Hungary[1]	14.5	14.2	18.5	13.8	3.1	3.1	4.3	3.0
Iceland	30.2	29.7	29.7	29.7	4.8	4.8	4.7	4.8
Ireland	10.5	9.3	10.0	9.3	2.6	2.5	2.7	2.5
Italy[1]	18.3	18.2	18.2	18.2	7.1	7.1	7.1	7.1
Japan	6.5	6.4	7.0	6.4	3.8	3.7	3.9	3.7
Korea	9.9	9.8	10.1	9.8	3.6	3.5	3.6	3.5
Luxembourg[1]	22.2	22.2	22.1	22.2	7.3	7.2	7.3	7.2
Mexico	10.2	10.1	10.4	10.1	3.8	3.7	3.8	3.7
Netherlands	12.2	11.8	11.8	11.8	5.0	4.9	4.9	4.9
New Zealand	16.6	16.9	16.0	16.9	5.0	5.1	5.1	4.9
Norway	28.2	28.3	28.3	28.3	7.4	7.4	7.3	7.4
Poland[1]	7.2	9.4	10.0	10.0	2.3	2.6	2.6	2.8
Portugal[1]	6.1	6.0	6.0	6.0	2.1	2.1	2.2	2.1
Slovak Republic	9.5	10.4	9.6	10.4	3.7	3.8	3.7	3.8
Spain	10.3	10.1	10.1	10.1	3.5	3.4	3.5	3.4
Sweden	18.4	18.2	18.2	18.2	6.0	5.9	5.9	5.9
Switzerland[1]	10.7	10.6	10.6	10.6	5.4	5.4	5.4	5.4
United Kingdom	15.0	14.8	14.7	14.8	5.7	5.7	5.7	5.7
United States	23.1	23.0	23.1	23.0	8.3	8.3	8.3	8.3
Average	15.0	15.0	15.4	14.9	5.0	4.9	5.1	4.9

Note : DEA performed using the preferred specification for cost efficiency: 2 inputs (cumulative spending per student and student's socio-
economic background) and 2 outputs (average PISA scores and equity objective), assuming non-increasing returns to scale (NIRS).
Cumulative spending in columns 1 and 5 was estimated over the period 1993 to 2002 for a student aged 15 in 2002. In the remaining columns
it was estimated as the sum of per student spending for students between the age of 6 and 15 years in 2002.
Cumulative spending was converted into equivalent US dollars using PPP's for either private consumption (PC), consumer services (CS) or
gross domestic product (GDP).
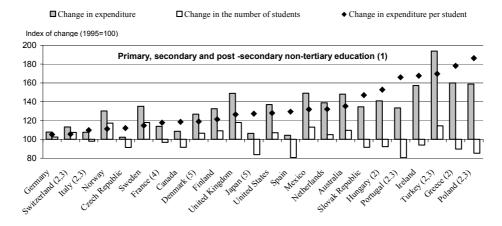1. Public institutions only.
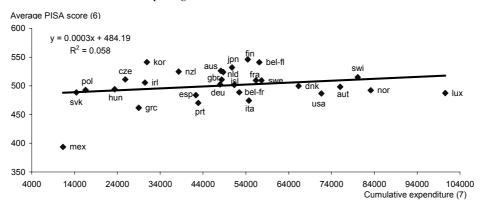Source: OECD (2005a); OECD (2005c); OECD calculations.

Figure 1. Level of resources and educational attainment

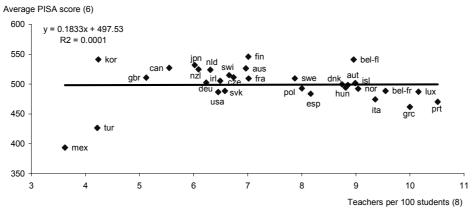**A. Changes in spending on educational institutions per student**

2003 constant prices

□ Change in expenditure □ Change in the number of students ◆ Change in expenditure per student

B. Spending on education and education attainment

C. Number of teachers per student and education attainment

1. Countries are ranked in ascending order of change in expenditure on educational institutions per student.
2. Public expenditure only.
3. Public institutions only.
4. Year of reference 2002.
5. Post-secondary non-tertiary included in both upper secondary and tertiary education.
6. Student performance on the combined reading, scientific, mathematical and problem solving scales in 2003.
7. Cumulative actual expenditure on educational institutions per pupil between the ages of 6 and 15 converted to 2002 US dollars using PPPs for household final consumption.
8. Number of teachers per 100 students in primary and secondary education in 2003 or latest year available: 2001 for Canada, 2002 for Korea, Portugal and Switzerland.

Source : OECD (2004, 2005a, 2006); OECD calculations.

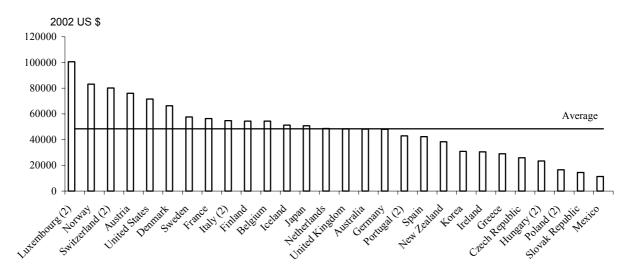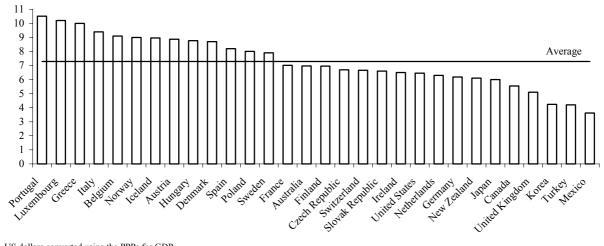**Figure 2. Human capital measures: educational attainment and equity, 2003**

1. Measured by the ratio of the average score of the 25th percentile to that of the 75th percentile (the higher the ratio, the greater the homogeneity in student performance).
Source: OECD (2004, 2005a).

Figure 3. Educational inputs per student

A. Cumulative expenditure on educational institutions per student between 6 and 15 years, 2002[1]

2002 US $

B. Number of teachers per 100 students in primary and secondary education, 2003[3]

1. US dollars converted using the PPPs for GDP.
2. Public institutions only.
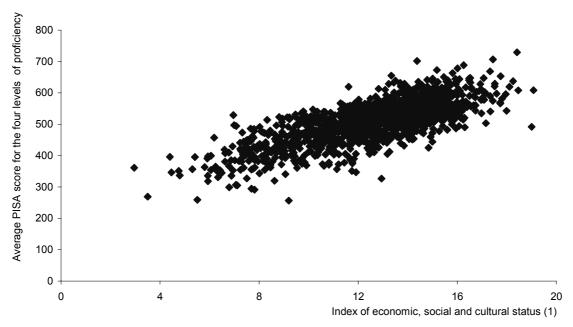3. Or latest year available: 2001 for Canada, 2002 for Korea, Portugal and Switzerland.
Source : OECD (2005a).

Figure 4. Socio-economic background and PISA scores, 2003

A: National level

y = 15.583x + 305.45
R^2 = 0.5008

B: School level

1. PISA index of economic, social and cultural status (ESCS) rebased on the sub-index of highest level of educational attainment of the parents. See footnote 16 in the core text.
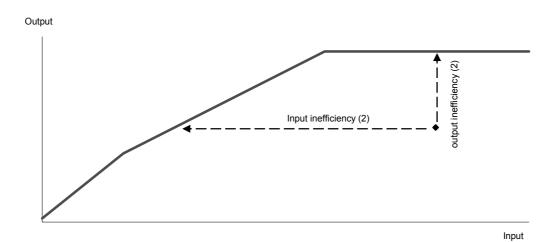   Source: OECD (2004,2005a).

Figure 5. The "efficiency frontier" and measurement of inefficiency[1]

1. The "efficiency frontier" has been designed under the assumption of non-increasing returns to scale.
2. Potential efficiency gains are derived by measuring the distance from the efficient frontier as a ratio of the distance from the axis to the efficiency frontier. They can be defined as the amount by which input could be reduced while holding constant the level of output (input inefficiency) or as the amount by which output could be increased while holding constant the level of input (output orientation).
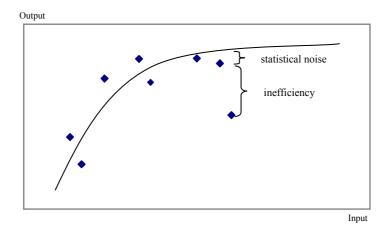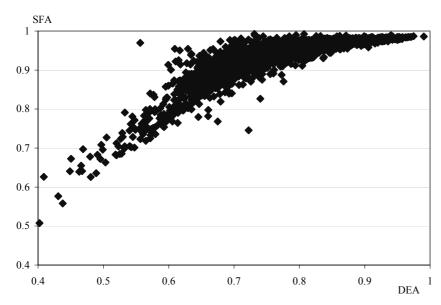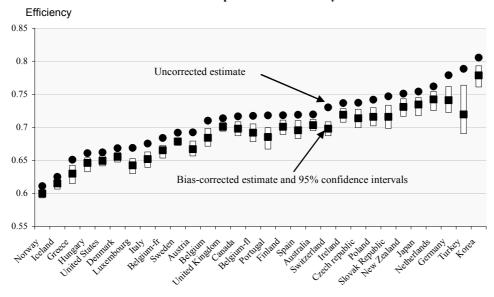
Figure 6. Stochastic frontier analysis

Figure 7. Comparison of DEA and SFA efficiency estimates
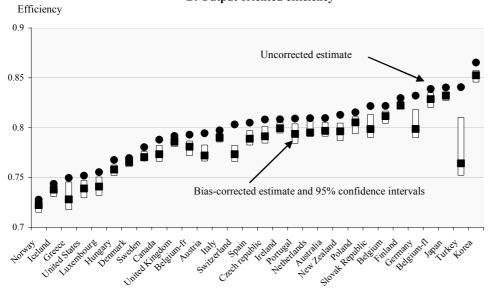
DEA output orientation, wider specification

Note: DEA is Data envelopement analysis and SFA is Stochastic Frontier Analysis.

Figure 8. Uncertainty surrounding technical efficiency estimates for the median school[1]
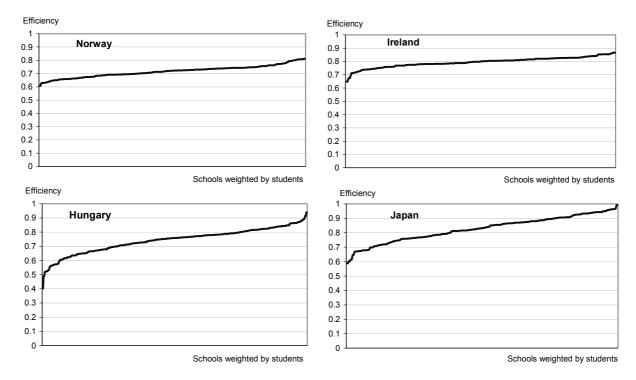
A. Input oriented efficiency

Efficiency

B. Output oriented efficiency

Efficiency

1. DEA performed with 4 inputs (teaching and computing resources, socio-economic status of students and language background) and one output (average PISA score).

Figure 9. Estimates of school-level efficiency within selected countries
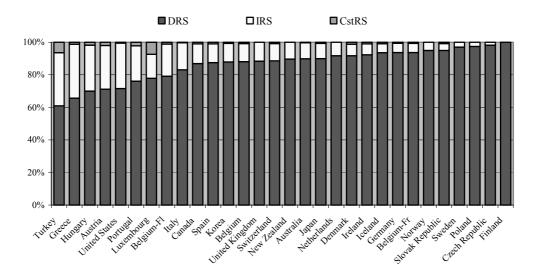
Bootstrap estimates, non-increasing return to scale, output efficiency

Note: These figures present the estimates of school-level efficiency for each country ordered by the level of efficiency (vertical axis), each school being weighted by the number of students. A fully efficient school would have a score of 1. DEA has been performed using PISA score as output and teaching and computing ressources, socio-economic status of students and language background as inputs.

Figure 10. Returns to scale on the efficiency frontier

Percentage of schools on the portion of the efficiency frontier
with diminishing, increasing or constant return to scale[1]

Note: DRS: Diminishing return to scale; IRS: Increasing return to scale; CstRS: Constant return to scale. DEA estimates with
PISA score as output and teachers student ratio, computer availability, socio-economic and language backgrounds as inputs.
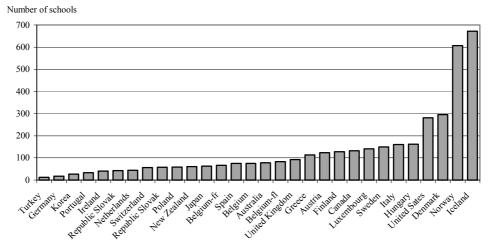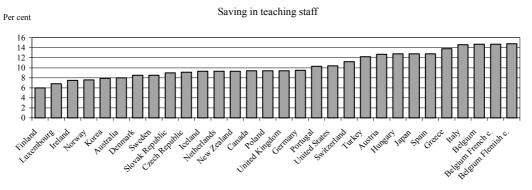1. Assuming that input inefficiency has been removed.

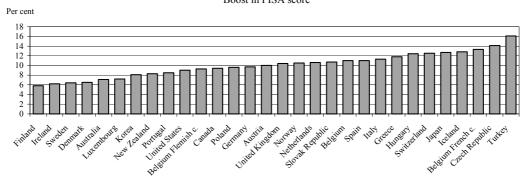**Figure 11. Average number of "reference" schools**

Number of schools

Note: For a given school, the "reference" schools are schools that produces at least as much output with the same level or fewer inputs.

Figure 12. Potential aggregate resource savings and gains in PISA score at the school level

A. Moving the efficiency level to at least the 95 percentile level

Saving in teaching staff

Per cent

Boost in PISA score

Per cent

B. Moving the efficiency level to at least the average level

Saving in teaching staff

Per cent

Per cent   Boost in PISA score

Pour les calculs voir le fichier de Douglas : S:\applic\...\savings&gains_50.xls

suppression des 5% en tête et enqueue sur l'échelle des performances.

suppression des 5% en queue sur l'échelle des performances et application du score du 5ème décile aux données entre ce décile et les % en queue.

Figure 13. Comparison of country and median school efficiency estimates
DEA bootstrap estimates[1]

A. Input oriented efficiency

B. Output oriented efficiency

1. Performed with 2 inputs (teachers per 100 students and students' socio-economic status) and 1 output (average PISA score) at the national and school levels.

37

Figure 14. Uncertainty surrounding technical efficiency estimates at the national level[1]

A. Input oriented efficiency

B. Output oriented efficiency

1. DEA performed with 2 inputs (teachers per 100 students and pupils' socio-economic background) and 2 outputs (average PISA scores and equity objective).

Figure 15. Technical effciency with and without an equity objective

Baseline specification

A. Input oriented efficiency

Baseline specification

B. Output oriented efficiency

Note: Baseline specification : 2 inputs (teachers per 100 students and pupil's socio-economic background) and one output (average PISA scores); preferred specification 2 inputs (teachers per 100 students and pupil's socio-economic background) and 2 outputs (average PISA scores and equity objective).

Figure 16. Potential resource savings at national level[1]

Per cent of GDP

1. Potential inputs cuts applied to compensation of all staff in primary, secondary and post-secondary non-tertiary education in year 2002.
Source : OECD calculations; Education at a glance, OECD indicators 2005.

Figure 17. Technical efficiency at the national level and PISA performance
Preferred specification

A. Input oriented efficiency

B. output oriented efficiency

Figure 18. Estimates of technical efficiency at the national level

A. Input oriented efficiency

☐ All available countries   ■ Mexico, Portugal and Turkey excluded

B. Output oriented efficiency

☐ All available countries   ■ Mexico, Portugal and Turkey excluded

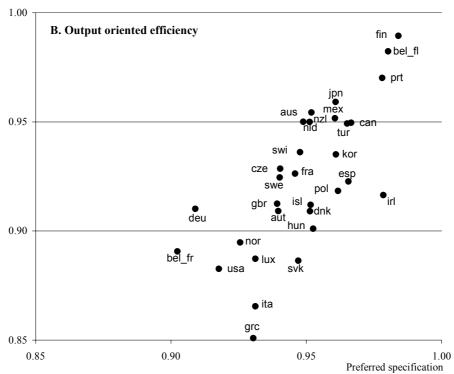Note: Preferred specification with 2 inputs (teachers per 100 students and pupil's socio-economic background) and 2 outputs (average PISA scores and equity objective).

Figure 19. Uncertainty surrounding cost efficiency estimates[1]

A. Input oriented efficiency

B. Output oriented efficiency

1. DEA performed with 2 inputs (cumulative expenditure per pupils and pupils' socio-economic background) and 1 output (average PISA score).
2. Data for these countries concern public institutions only.

Figure 20. Cost effciency with and without an equity objective

Note: Baseline specification : 2 inputs (cumulative expenditure per pupil and pupil's socio-economic background) and one output (average PISA scores); preferred specification 2 inputs (cumulative expenditure per pupil and pupil's socio-economic background) and 2 outputs (average PISA scores and equity objective).
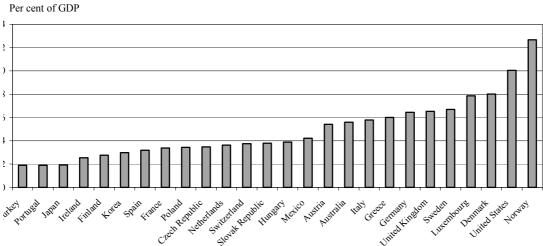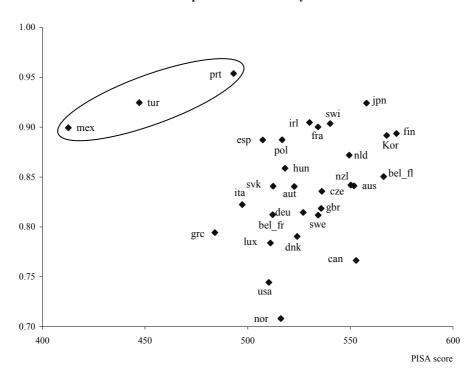
Figure 21. Cost efficiency at the national level and PISA performance
Preferred specification

A. Input oriented efficiency

B. output oriented efficiency

ANNEX 1: SUPPORTING TECHNICAL DOCUMENTATION

Non-parametric efficiency measures

41.   The Free Disposal Hull (FDH) and Data Envelopment Analysis (DEA) are non-parametric approaches to identifying efficiency relative to the observed most efficient observation. The measurement of inefficiency for units within the efficiency frontier assumes that the input or output mixes are held constant in moving towards the frontier (this feature is also called radial). The efficiency benchmark varies across observations due to the different mixes of inputs or outputs. By way of illustration, the following formulation gives the constant returns to scale versions in the input direction for the ith observation:

$$\text{Min}_{\theta,\gamma} \ \theta_i$$

$$\text{subject to, } \sum_{j=1}^{N} \gamma_{ij} \, x_j \le \theta_i \, x_i \, , \ \sum_{j=1}^{N} \gamma_{ij} \, y_j \ge y_i \, , \ \gamma_j \ge 0$$

where x and y are vectors of inputs and outputs, respectively, γ is a vector of constants (the weights given t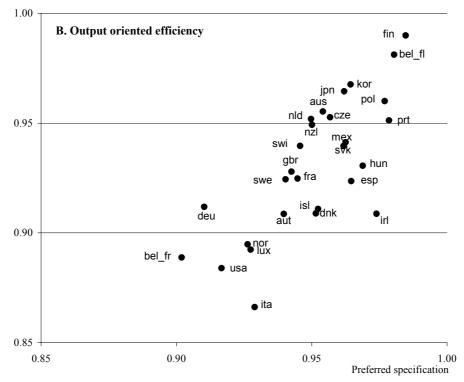o units that form the peer group for the measurement of efficiency) and θ is the distance measure of technical inefficiency. Thus, θ gives a measure of the proportion by which inputs can be reduced -- without reducing the output of y -- in relation to the best linear combination of other units in the sample.[31] In order to estimate variable (or non-increasing) returns to scale, the convexity constraint requires that the vector of constants, γ, multiplied by a row of ones is equal (or less than or equal) to 1.

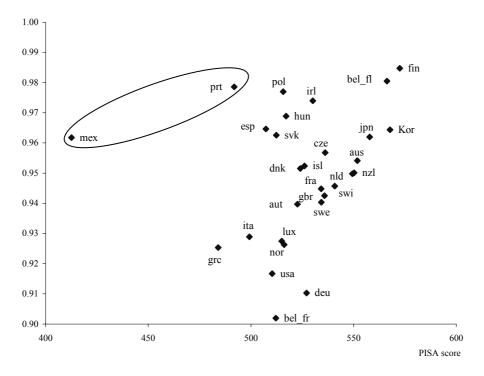42.   As noted in the main text, the non-parametric approach is not without drawbacks. The impact of measurement error, statistical noise and outliers that define the efficiency frontier will influence the estimates of efficiency for other observations within the frontier.[32] Furthermore, using variables with different distributions can have marked impacts on efficiency estimates.[33] Finally, small samples and relatively large numbers of inputs and/or outputs will bias estimates of efficiency (as an observation that defines a portion of the efficiency frontier may be excluded from the sample).

43.   Estimates of non-parametric efficiency need to account for these difficulties. With respect to statistical noise and measurement error, statistical re-sampling techniques (the jackknife and bootstrap described below) allow some testing of the robustness of results. Different methods, such as principal

---

31.   In the main text, results are reported using the convention that 1 is a fully efficient observation and that inefficient observations are less than 1. These are the Farrell measure of input orientated efficiency and the Shepard measures of output orientation efficiency.

32.   In addition, complex serial correlation will invalidate central assumptions to standard statistical tests used in any second stage analysis identifying determinants of differences in efficiency.

33.   This is the reason the percentile on percentile ratio is used as a measure of the homogeneity in PISA scores. The inverse of the standard deviation, for example, would lead to a non-normal distribution and at the extremes much larger estimates of inefficiency.

component analysis, can help reduce the dimensions of inputs or outputs and thus mitigate the problems associated with small sample bias.

Detecting outliers

44.   As a first step in examining the data, both summary statistics and graphical presentation can help identify observations that are *potential* outliers. A second approach is to use jackknife re-sampling to estimate whether the data point is an influential observation. This technique repeatedly recalculates the desired statistic (here the estimate of technical efficiency) omitting one observation sequentially. The comparison of the efficiency estimates for the whole sample with the estimates for the sample when an observation is left out helps identify those observations that make large differences in the efficiency estimates for other observations. The observations highlighted need not be a result of measurement error and "outliers" may describe a portion of the efficiency frontier where there are relatively few neighbouring observations. In this light, unless the observation is an obvious outlier, the approach adopted was to retain the observation. A drawback is that the technique will fail to detect groups of outliers that are close neighbours.[34]

Robustness of results

45.   Even though exploratory data analysis should help eliminate observations that are obvious outliers, the results are still sensitive to measurement error and statistical noise. Second, in a finite sample, particularly when the number of inputs and outputs rises, the efficiency estimate will be biased towards smaller estimates of inefficiency. Bootstrapping can help address these problems by producing confidence intervals around the point estimates as well as making a correction for the small sample bias. The bootstrap is based on the assumption that the observed data are representative of the population. Taking repeated samples that are the same size as observed data will mimic sampling from the total population. Roughly speaking, repeating this sampling generates different efficiency estimates and when the number of (re)-samples is large the standard errors of these estimates can be used to derive the confidence intervals. In addition, the bootstrapping procedure allows the small-sample bias to be estimated and a correction to be made to the efficiency estimate.[35]

Reducing the dimensions of the problem

46.   Given that finite sample bias is potentially aggravated by larger numbers of inputs and outputs in the non-parametric estimate (the so-called "curse of dimensionality"), techniques that can aggregate information, such as principal component analysis, are potentially useful. The objective of principal component analysis is to combine several variables with a minimal loss of information. The analysis finds a set of linear combinations of the variables -- the principal components -- which explain all the variance in the data. They are ordered such that the first principal component accounts for the majority of the variance, the second principal component -- which is uncorrelated with the first — accounts for the second most variance, and so forth. While there are as many principal components as original variables, it is often possible to explain most of the variance with a few principal components. The criteria for omitting or retaining a principal component can vary, though a benchmark of 90% of the cumulative variance explained is frequently used. The weight given to each variable (the factor loadings) gives an indication of the relative importance of the variables to the principal component and can give an intuition as to what underlying feature in the data the principal component is capturing.

---

34.   Wilson (1993) developed a test to examine possible groups of outliers, but this is computationally too expensive for large datasets.

35.   See Simar and Wilson (2000) for a full exposition.

47.   The PISA results for scores in the areas of Math, Problem Solving, Reading and Science are highly correlated and the first principal component accounts for almost all the variance. The first principal component explains just over 90% of the variation in the four separate PISA scores. Furthermore, the factor loadings -- or the weights given to each of the PISA scores -- are roughly equal (0.96, 0.96, 0.94, 0.95 for math, problem solving, reading and science, respectively), indicating that the first principal component alone is a good measure of overall performance in addition to capturing almost all the variance in the data (Figure A1.1).

Stochastic frontier analysis

48.   Stochastic frontier analysis is similar to standard regression techniques but differs by exploiting the one-sided nature of inefficiency to decompose the error term into a standard error term and an asymmetric component that measures inefficiency. Formally, the basic stochastic frontier model is given by:

$$y_i = f(x_i, \beta) + v_i - \upsilon_i$$

where, y_i is the output of school i, $f(.)$, is a measurable production function, x_i are exogenous variables, β is a vector of unknown parameters and $v_i\text{-}u_i$ is the composed error term consisting of v the symmetric disturbance and u the non-negative disturbance measuring the inefficiency of the school. There are a number of different types of assumptions on the distribution of the inefficiency, such as half-normal and exponential.[36]

49.   The principal advantage of this approach to measuring efficiency is that it addresses statistical noise explicitly. Furthermore, and in contrast to non-parametric approaches, standard statistical tests can be used to assess variables. The cost of specifying a specific functional form can be mitigated somewhat by the choice of a translog production function. This is a very flexible specification, also referred to as a second order approximation of an unknown functional form. However, estimating this production function in practice is often complicated by multicollinearity (the variables are too highly correlated with one another to allow their individual effects to be estimated with precision). In addition, the specification may not make economic sense for all observations and thus need to be checked. For example, the estimated coefficients may imply for a particular observation that increasing inputs would lead to fewer outputs.

The data

50.   In preparing the data, observations with missing data or zero entries for inputs or outputs were excluded, as well as observations which were clearly outliers. At the national level, empirical influence tests suggested Korea was an outlier. On this basis, results excluding Korea are reported in Annex 2 for comparison. The PISA 2003 and *Education at a* Glance databases also allowed the construction of separate indicators for the Flemish and French speaking communities in Belgium, though this means that a single observation for Belgium is excluded from the national-level data set. This gave 30 observations for the country level analysis of technical efficiency and 28 observations for the analysis of cost efficiency. At the school level the data cleaning procedures left a sample size of 6204 schools, but excluded all the schools in France and Mexico. Basic statistics about the data used in the estimations are reported in Table A1.1. In order to cover as many countries as possible in the analysis, the teacher per 100 student variable at the national level is taken from the *Education at a Glance* -- rather than the PISA 2003 -- database.

---

36.   Greene (1997).

Table A1.1. Basic features of the data set

	Country level				Median school level					
	PISA scores	Homogeneity of Pisa scores (25th/75th)	Cumulative expenditure per student (2002 US$)	Number of teachers per 100 students	Index of economic, social and cultural status	PISA scores	Number of teachers per 100 students	Number of computers per 100 students	Index of economic, social and cultural status	Language spoken at home	Number of schools
Australia	551.8	0.8	48170	7.0	13.2	549.5	7.3	26.1	13.4	0.96	316
Austria	522.7	0.8	76029	8.9	12.6	535.4	9.8	15.3	13.0	0.96	156
Belgium	542.9	0.7	54449	9.1	12.9	554.2	11.3	11.5	13.3	0.97	245
Belgium Flemish c.	566.3	0.8	57129	9.0	13.1	569.7	11.5	14.8	13.4	1.00	142
Belgium French c.	512.1	0.7	52491	9.5	12.7	530.8	10.7	8.2	13.2	0.97	91
Canada	552.8	0.8	..	5.5	14.1	543.1	6.2	21.7	13.8	1.00	923
Czech Republic	536.1	0.8	25852	6.7	13.0	536.2	6.6	7.8	13.2	1.00	228
Denmark	524.1	0.8	66323	8.7	13.1	526.9	8.9	15.4	13.3		170
Finland	572.5	0.8	54456	7.0	13.3	570.5	9.4	15.2	13.6	1.00	189
France	534.2	0.8	56361	7.0	12.1	..	..	..	..	..	..
Germany	527.1	0.7	47991	6.2	13.0	532.1	5.8	6.7	13.1	0.95	174
Greece	484.0	0.7	28982	10.0	11.8	442.4	10.8	6.7	11.0	1.00	166
Hungary	518.2	0.8	23362	8.8	12.1	480.5	8.9	11.0	11.6	1.00	223
Iceland	526.1	0.8	51315	9.0	15.0	528.6	10.3	19.1	14.2	1.00	109
Ireland	530.1	0.8	30456	6.5	12.1	534.7	7.1	9.4	12.5	1.00	104
Italy	497.4	0.7	54723	9.4	12.0	531.1	12.0	13.7	12.4	1.00	383
Japan	557.8	0.8	50763	6.0	12.1	555.4	7.0	13.3	12.4	1.00	140
Korea	567.7	0.8	30895	4.2	12.0	578.0	6.0	20.0	12.4	1.00	141
Luxembourg	511.0	0.8	100572	10.2	13.1	513.1	10.9	15.4	13.0	0.84	27
Mexico	412.7	0.7	11296	3.6	8.1	..	..	..	..	..	..
Netherlands	549.4	0.8	48711	6.3	12.7	549.4	6.4	13.2	13.2	1.00	121
New Zealand	550.3	0.8	38314	6.1	13.2	551.0	6.2	20.8	13.2	0.96	155
Norway	516.2	0.8	83165	9.0	14.7	519.6	9.8	15.4	14.7	1.00	139
Poland	516.8	0.8	16583	8.0	11.6	515.8	7.4	4.9	12.1	1.00	158
Portugal	493.2	0.8	42894	10.5	10.0	505.5	10.7	6.7	10.1	1.00	144
Slovak Republic	512.4	0.8	14420	6.6	12.1	510.1	6.7	4.6	12.4	1.00	259
Spain	507.3	0.8	42356	8.2	11.3	516.7	8.4	7.9	11.8	1.00	320
Sweden	534.3	0.8	57652	7.9	13.3	537.2	8.4	14.4	13.6	0.96	169
Switzerland	540.2	0.8	80123	6.7	12.2	522.3	8.8	13.2	12.2	0.94	386
Turkey	447.2	0.7	..	4.2	8.7	439.2	5.5	2.3	8.8	1.00	110
United Kingdom [1]	535.8	0.8	48355	5.1	12.8	537.0	6.7	21.9	12.8	1.00	345
United States	510.3	0.8	71626	6.5	13.5	519.1	6.9	26.4	13.7	1.00	204
Average	523.8	0.8	48860	7.4	12.4	527.8	8.4	13.4	12.7	0.98	

1. PISA 2003 results presented here must be treated with care. Uncertainties surrounding the sample are such that scores for the United Kingdom cannot reliably be compared with those of other countries.

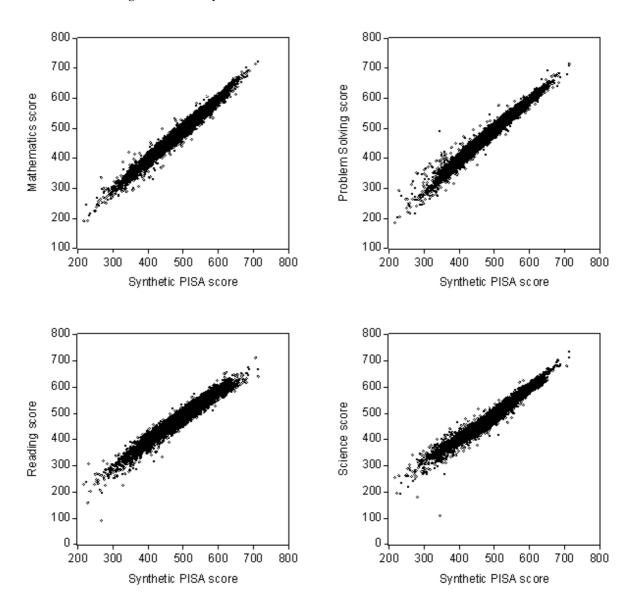Source: OECD Education Database and PISA 2003 Database

Figure A1.1. The synthetic PISA score and the individual PISA scores

ANNEX 2: SUPPLEMENTARY RESULTS

51.   The supplementary results present the underlying estimates of technical or cost efficiency for the preferred specification presented in the main text. In addition, details of additional tests or results are reported to support information presented in the main text.

School-level estimates of technical efficiency

52.   The preferred specification at the school level include the synthetic indicator or the PISA scores as an output and measures of teaching and computing inputs, student socio-economic and language background as inputs. Necessary data were unavailable for France and Mexico and as a result they are excluded from the school-level analysis. Table A2.1 presents the estimates for the median schools (the median school in the input orientation and the median school in the output orientation). The results of projecting schools onto the efficiency frontier to determine whether they would be on a portion of the frontier characterised by constant, diminishing or increasing returns to scale are presented in Table A2.2 for selected countries.

National estimates of technical and cost efficiency

53.   The preferred specification includes the synthetic indicator of the PISA scores and the ratio of the 25th to 75th percentile as measures of outputs and variables measuring teaching resources and the socio-economic background of parents as inputs. Using this specification, the main results at the school level are presented in Table A2.3 and Table A2.4 (with Korea excluded) and Table A2.5 (for public institutions only).[37] The estimates of alternative specifications for technical efficiency are reported in Table A2.6 (using the score on the PISA mathematics test as an output in place of the combined PISA score) and Table A2.7 (adding enrolment rates as a third output). The estimates of cost-efficiency at the national level (replacing the teaching input with the measure of cumulative spending) are reported in Table A2.8 and Table A2.9 (with Korea excluded) and Table A2.10 (for public institutions only). The estimates of alternative specifications for cost efficiency are reported in Table A2.11 (using the score on the PISA mathematics test as an output in place of the combined PISA score) Table A2.12 (adding enrolment rates as a third output), and Table A2.13 (using the all economy purchasing power parity (PPP) in place of the final consumption PPP).

---

37.   Additional specification tests included examining the importance of the choice of the output variable. One estimate used the mathematics score as the output score as this is often assumed to be less likely to be taught at home and therefore a better measure of the importance of the school. The sample was also split by size of schools. These specifications did not lead to markedly different results.

Table A2.1. DEA estimates of technical efficiency at the school level

Preferred specification[1]

	Input efficiency		Output efficiency	
	Median	Range 90%	Median	Range 90%
Australia	0.70	0.15	0.80	0.14
Austria	0.67	0.26	0.77	0.26
Belgium	0.68	0.32	0.81	0.28
Belgium Flemish c.	0.69	0.22	0.83	0.22
Belgium French c.	0.67	0.41	0.78	0.38
Canada	0.70	0.20	0.77	0.19
Czech Republic	0.71	0.18	0.79	0.24
Denmark	0.66	0.18	0.77	0.13
Finland	0.70	0.13	0.82	0.11
Germany	0.74	0.19	0.80	0.23
Greece	0.63	0.25	0.73	0.25
Hungary	0.65	0.26	0.76	0.29
Iceland	0.62	0.16	0.74	0.15
Ireland	0.72	0.17	0.80	0.13
Italy	0.65	0.30	0.79	0.28
Japan	0.73	0.30	0.83	0.28
Korea	0.78	0.21	0.85	0.20
Luxembourg	0.64	0.16	0.74	0.13
Netherlands	0.74	0.22	0.80	0.21
New Zealand	0.73	0.19	0.80	0.17
Norway	0.60	0.12	0.72	0.15
Poland	0.72	0.18	0.81	0.17
Portugal	0.69	0.22	0.79	0.22
Slovak Republic	0.72	0.20	0.80	0.23
Spain	0.70	0.23	0.79	0.20
Sweden	0.68	0.13	0.77	0.12
Switzerland	0.71	0.23	0.79	0.23
Turkey	0.72	0.28	0.76	0.30
United Kingdom	0.70	0.17	0.79	0.18
United States	0.65	0.19	0.74	0.20
Standard deviation	0.04	0.06	0.03	0.06
Median	0.70	0.20	0.79	0.21
Average	0.69	0.21	0.78	0.21

1. PISA score as output and teachers student ratio, computer availability, socio-economic and language backgrounds as inputs.

Table A2.2. Returns to scale on the efficiency frontier

Percentage of schools projected onto portions of the efficiency frontier
that are characterised by constant, diminishing or increasing returns to scale
(Preferred specification)

	Input orientation			Output orientation		
	Constant	Diminishing	Increasing	Constant	Diminishing	Increasing
Australia	0.3	89.9	9.8	0.3	99.7	0.0
Austria	1.9	71.2	26.9	0.6	99.4	0.0
Belgium	0.8	88.2	11.0	0.4	99.2	0.4
Belgium-Flemish c.	1.1	79.1	19.8	0.0	98.9	1.1
Belgium-French c.	0.7	93.7	5.6	0.7	99.3	0.0
Canada	1.0	86.9	12.1	0.4	99.5	0.1
Czech Republic	0.0	98.2	1.8	0.0	100	0.0
Denmark	1.2	91.8	7.1	0.6	99.4	0.0
Finland	0.0	100	0.0	0.0	100	0.0
France						
Germany	0.6	93.7	5.7	0.6	99.4	0.0
Greece	1.2	65.7	33.1	0.6	99.4	0.0
Hungary	1.8	70.0	28.3	0.4	99.1	0.4
Iceland	0.9	93.6	5.5	0.0	100	0.0
Ireland	1.0	92.3	6.7	1.0	99.0	0.0
Italy	0.3	83.0	16.7	0.0	99.7	0.3
Japan	0.7	90.0	9.3	0.0	100	0.0
Korea	0.7	87.9	11.3	0.0	100	0.0
Luxembourg	7.4	77.8	14.8	3.7	96.3	0.0
Mexico						
Netherlands	0.0	91.7	8.3	0.0	100	0.0
New Zealand	0.0	89.7	10.3	0.0	100	0.0
Norway	0.0	95.0	5.0	0.0	100	0.0
Poland	0.0	97.5	2.5	0.0	100	0.0
Portugal	2.1	76.1	21.8	2.1	97.9	0.0
Slovak Republic	0.8	95.0	4.2	0.4	99.6	0.0
Spain	0.9	87.5	11.6	0.3	99.7	0.0
Sweden	0.0	97.0	3.0	0.0	100	0.0
Switzerland	0.8	88.6	10.6	0.5	99.2	0.3
Turkey	6.4	60.9	32.7	7.3	87.3	5.5
United Kingdom	0.0	88.4	11.6	0.0	100	0.0
United States	0.5	71.6	27.9	1.0	98.5	0.5

Source: OECD Secretariat calculations.

Table A2.3. DEA estimates of technical efficiency at the national level

Preferred specification[1]

	Input efficiency			Output efficiency		
	VRS	NIRS	Cst RTS	VRS	NIRS	Cst RTS
Australia	0.84	0.84	0.78	0.95	0.95	0.78
Austria	0.84	0.84	0.79	0.94	0.94	0.78
Belgium Flemish c.	0.85	0.85	0.82	0.98	0.98	0.82
Belgium French c.	0.81	0.81	0.77	0.90	0.90	0.76
Canada	0.77	0.77	0.73	0.97	0.97	0.72
Czech Republic	0.84	0.84	0.77	0.94	0.94	0.77
Denmark	0.79	0.79	0.75	0.95	0.95	0.75
Finland	0.89	0.89	0.81	0.98	0.98	0.80
France	0.90	0.90	0.83	0.95	0.95	0.83
Germany	0.81	0.81	0.75	0.91	0.91	0.75
Greece	0.79	0.79	0.76	0.93	0.93	0.76
Hungary	0.86	0.86	0.81	0.95	0.95	0.81
Iceland	0.70	0.70	0.66	0.95	0.95	0.66
Ireland	0.90	0.90	0.81	0.98	0.98	0.81
Italy	0.82	0.82	0.78	0.93	0.93	0.78
Japan	0.92	0.92	0.86	0.96	0.96	0.85
Korea	0.89	0.89	0.88	0.96	0.96	0.84
Luxembourg	0.78	0.78	0.74	0.93	0.93	0.74
Mexico	0.90	0.90	0.87	0.96	0.96	0.81
Netherlands	0.87	0.87	0.80	0.95	0.95	0.80
New Zealand	0.84	0.84	0.77	0.95	0.95	0.77
Norway	0.71	0.71	0.66	0.93	0.93	0.66
Poland	0.89	0.89	0.84	0.96	0.96	0.83
Portugal	0.95	0.95	0.91	0.98	0.98	0.91
Slovak Republic	0.84	0.84	0.79	0.95	0.95	0.78
Spain	0.89	0.89	0.84	0.97	0.97	0.84
Sweden	0.81	0.81	0.76	0.94	0.94	0.75
Switzerland	0.90	0.90	0.84	0.95	0.95	0.83
Turkey	0.92	0.92	0.91	0.97	0.97	0.88
United Kingdom	0.82	0.82	0.78	0.94	0.94	0.76
United States	0.74	0.74	0.69	0.92	0.92	0.69
Standard deviation	0.06	0.06	0.06	0.02	0.02	0.06
Median	0.84	0.84	0.79	0.95	0.95	0.78
Average	0.84	0.84	0.79	0.95	0.95	0.78

Note: VRS = variable returns to scale; NIRS = non-increasing returns to scale; Cst RTS = constant returns to scale.
1. Bootstrap estimates with 2 inputs (teachers per 100 students and socio-economic background) and 2 outputs (average PISA score and homogeneity of PISA score).

Table A2.4. DEA estimates of technical efficiency at the national level without Korea

Preferred specification[1]

	Input efficiency			Output efficiency		
	VRS	NIRS	Cst RTS	VRS	NIRS	Cst RTS
Australia	0.88	0.88	0.78	0.95	0.95	0.78
Austria	0.85	0.85	0.78	0.94	0.94	0.78
Belgium Flemish c.	0.92	0.92	0.82	0.98	0.98	0.82
Belgium French c.	0.82	0.82	0.77	0.91	0.91	0.76
Canada	0.92	0.92	0.81	0.97	0.97	0.79
Czech Republic	0.86	0.86	0.77	0.94	0.94	0.76
Denmark	0.83	0.83	0.75	0.95	0.95	0.75
Finland	0.91	0.91	0.80	0.98	0.98	0.80
France	0.92	0.92	0.83	0.97	0.97	0.83
Germany	0.84	0.84	0.74	0.93	0.93	0.73
Greece	0.80	0.80	0.76	0.94	0.94	0.76
Hungary	0.88	0.88	0.81	0.96	0.96	0.80
Iceland	0.74	0.74	0.66	0.95	0.95	0.65
Ireland	0.93	0.93	0.81	0.99	0.99	0.81
Italy	0.83	0.83	0.78	0.94	0.94	0.78
Japan	0.93	0.93	0.85	0.97	0.97	0.85
Luxembourg	0.79	0.79	0.74	0.93	0.93	0.73
Mexico	0.91	0.92	0.86	0.97	0.97	0.80
Netherlands	0.91	0.91	0.80	0.97	0.97	0.79
New Zealand	0.89	0.89	0.76	0.97	0.97	0.75
Norway	0.72	0.72	0.66	0.93	0.93	0.66
Poland	0.92	0.92	0.84	0.97	0.97	0.83
Portugal	0.95	0.94	0.91	0.98	0.98	0.91
Slovak Republic	0.87	0.87	0.79	0.95	0.95	0.78
Spain	0.92	0.92	0.84	0.97	0.97	0.84
Sweden	0.85	0.85	0.75	0.94	0.94	0.75
Switzerland	0.94	0.94	0.83	0.97	0.97	0.83
Turkey	0.93	0.93	0.90	0.97	0.97	0.88
United Kingdom	0.94	0.93	0.84	0.97	0.97	0.83
United States	0.77	0.77	0.69	0.92	0.92	0.68
Standard deviation	0.06	0.06	0.06	0.02	0.02	0.06
Median	0.89	0.89	0.79	0.96	0.96	0.79
Average	0.87	0.87	0.79	0.96	0.96	0.78

Note: VRS = variable returns to scale; NIRS = non-increasing returns to scale; Cst RTS = constant returns to scale.
1. Bootstrap estimates with 2 inputs (teachers per 100 students and socio-economic background) and 2 outputs (average PISA score and homogeneity of PISA score).

Table A2.5. DEA estimates of technical efficiency at the national level, public institutions only

Preferred specification[1]

	Input efficiency			Output efficiency		
	VRS	NIRS	Cst RTS	VRS	NIRS	Cst RTS
Austria	0.83	0.83	0.80	0.93	0.93	0.80
Belgium Flemish c.	0.85	0.85	0.82	0.91	0.91	0.81
Belgium French c.	0.81	0.81	0.79	0.89	0.89	0.79
Canada	0.76	0.76	0.74	0.96	0.96	0.73
Czech Republic	0.82	0.82	0.79	0.93	0.93	0.79
Finland	0.89	0.89	0.83	0.98	0.98	0.83
Germany	0.80	0.80	0.76	0.91	0.91	0.76
Greece	0.80	0.81	0.78	0.92	0.92	0.78
Hungary	0.85	0.85	0.82	0.95	0.95	0.82
Iceland	0.70	0.70	0.68	0.95	0.95	0.68
Ireland	0.88	0.88	0.84	0.97	0.97	0.84
Italy	0.83	0.83	0.80	0.93	0.93	0.80
Japan	0.90	0.90	0.90	0.97	0.97	0.90
Korea	0.89	0.89	0.89	0.96	0.96	0.85
Luxembourg	0.78	0.78	0.75	0.92	0.92	0.75
Mexico	0.89	0.89	0.88	0.96	0.95	0.83
New Zealand	0.82	0.82	0.78	0.95	0.95	0.78
Norway	0.71	0.71	0.68	0.92	0.92	0.68
Poland	0.87	0.87	0.84	0.95	0.95	0.84
Portugal	0.94	0.94	0.92	0.97	0.97	0.91
Slovak Republic	0.84	0.84	0.80	0.93	0.93	0.80
Spain	0.90	0.90	0.88	0.96	0.96	0.88
Sweden	0.80	0.80	0.77	0.94	0.94	0.77
Switzerland	0.89	0.88	0.85	0.94	0.94	0.85
Turkey	0.93	0.93	0.90	0.97	0.97	0.88
United Kingdom	0.82	0.82	0.78	0.94	0.94	0.77
United States	0.75	0.75	0.72	0.92	0.92	0.71
Standard deviation	0.06	0.06	0.06	0.02	0.02	0.06
Median	0.83	0.83	0.80	0.94	0.94	0.80
Average	0.84	0.84	0.81	0.94	0.94	0.80

Note: VRS = variable returns to scale; NIRS = non-increasing returns to scale; Cst RTS = constant returns to scale.
1. Bootstrap estimates with 2 inputs (teachers per 100 students and socio-economic background) and 2 outputs (average PISA score and homogeneity of PISA score).

Table A2.6. DEA estimates of technical efficiency at the national level,

using the mathematics score[1]

	Input efficiency			Output efficiency		
	VRS	NIRS	Cst RTS	VRS	NIRS	Cst RTS
Australia	0.84	0.84	0.79	0.95	0.95	0.78
Austria	0.86	0.86	0.80	0.95	0.95	0.80
Belgium Flemish c.	0.90	0.90	0.85	0.98	0.98	0.84
Belgium French c.	0.83	0.83	0.79	0.92	0.92	0.79
Canada	0.91	0.91	0.75	0.99	0.99	0.73
Czech Republic	0.85	0.85	0.79	0.94	0.94	0.78
Denmark	0.83	0.83	0.78	0.96	0.96	0.78
Finland	0.90	0.90	0.81	0.99	0.99	0.81
France	0.89	0.89	0.84	0.97	0.97	0.84
Germany	0.82	0.82	0.76	0.92	0.92	0.75
Greece	0.76	0.76	0.73	0.94	0.94	0.72
Hungary	0.86	0.86	0.81	0.96	0.96	0.80
Iceland	0.73	0.73	0.69	0.96	0.96	0.68
Ireland	0.94	0.94	0.82	0.99	0.99	0.81
Italy	0.81	0.81	0.77	0.95	0.95	0.76
Japan	0.94	0.94	0.87	0.97	0.97	0.86
Korea	0.90	0.90	0.88	0.97	0.97	0.84
Luxembourg	0.80	0.80	0.75	0.95	0.95	0.75
Mexico	0.91	0.91	0.87	0.97	0.97	0.82
Netherlands	0.89	0.89	0.83	0.97	0.97	0.83
New Zealand	0.84	0.84	0.78	0.95	0.95	0.78
Norway	0.71	0.71	0.67	0.95	0.95	0.67
Poland	0.89	0.89	0.84	0.97	0.97	0.83
Portugal	0.95	0.95	0.91	0.98	0.98	0.90
Slovak Republic	0.87	0.87	0.81	0.96	0.96	0.81
Spain	0.93	0.93	0.85	0.98	0.98	0.85
Sweden	0.81	0.81	0.76	0.95	0.95	0.76
Switzerland	0.92	0.93	0.86	0.96	0.96	0.86
Turkey	0.93	0.93	0.91	0.97	0.97	0.89
United Kingdom	0.81	0.81	0.78	0.96	0.96	0.77
United States	0.74	0.74	0.69	0.93	0.93	0.68
Standard deviation	0.06	0.06	0.06	0.02	0.02	0.06
Median	0.86	0.86	0.80	0.96	0.96	0.80
Average	0.86	0.86	0.80	0.96	0.96	0.79

Note: VRS = variable returns to scale; NIRS = non-increasing returns to scale; Cst RTS = constant returns to scale.
1. Bootstrap estimates with 2 inputs (teachers per 100 students and socio-economic background) and 2 outputs (PISA score on the mathematics scale and homogeneity of this PISA score).

Table A2.7. DEA estimates of technical efficiency at the national level,

with enrolment as an output[1]

	Input efficiency			Output efficiency		
	VRS	NIRS	CST RTS	VRS	NIRS	CST RTS
Australia	0.88	0.88	0.88	0.98	0.98	0.88
Austria	0.87	0.87	0.87	0.95	0.95	0.87
Belgium Flemish c.	0.94	0.94	0.89	0.99	0.99	0.89
Belgium French c.	0.92	0.92	0.88	1.00	1.00	0.88
Czech Republic	0.90	0.90	0.91	0.99	0.99	0.91
Denmark	0.94	0.94	0.87	1.00	1.00	0.87
Finland	0.94	0.94	0.90	0.99	0.99	0.90
France	0.94	0.94	0.95	0.98	0.98	0.95
Germany	0.90	0.90	0.90	0.97	0.97	0.90
Greece	0.88	0.87	0.87	0.94	0.94	0.87
Hungary	0.96	0.96	0.92	0.99	0.99	0.92
Iceland	0.78	0.78	0.78	0.99	0.99	0.78
Ireland	0.95	0.95	0.98	0.99	0.99	0.98
Italy	0.89	0.89	0.89	0.95	0.95	0.89
Japan	0.94	0.94	0.98	0.99	0.99	0.97
Korea	0.94	0.94	0.93	0.98	0.98	0.91
Luxembourg	0.80	0.80	0.80	0.93	0.93	0.80
Mexico	0.94	0.94	0.93	0.98	0.98	0.92
Netherlands	0.94	0.94	0.93	0.99	0.99	0.93
New Zealand	0.86	0.86	0.86	0.96	0.96	0.86
Norway	0.79	0.79	0.79	0.99	0.99	0.79
Poland	0.95	0.95	0.96	0.98	0.98	0.96
Portugal	0.95	0.95	0.95	0.98	0.98	0.95
Slovak Republic	0.96	0.96	0.97	0.99	0.99	0.97
Spain	0.96	0.96	0.98	0.99	0.99	0.97
Sweden	0.86	0.86	0.87	0.99	0.99	0.87
Switzerland	0.94	0.94	0.94	0.98	0.98	0.94
Turkey	0.95	0.95	0.95	0.99	0.99	0.94
United Kingdom	0.94	0.94	0.94	0.99	0.99	0.94
United States	0.87	0.86	0.87	0.98	0.98	0.86
Standard deviation	0.05	0.05	0.05	0.02	0.02	0.05
Median	0.94	0.94	0.90	0.99	0.99	0.90
Average	0.91	0.91	0.90	0.98	0.98	0.90

Note: VRS = variable returns to scale; NIRS = non-increasing returns to scale; Cst RTS = constant returns to scale.
1. Bootstrap estimates with 2 inputs (teachers per 100 students and socio-economic background) and 3 outputs (average PISA score, homogeneity of PISA score and enrolment rates at age 15).

Table A2.8. DEA estimates of cost efficiency at the national level

Preferred specification[1]

	Input efficiency			Output efficiency		
	VRS	NIRS	CST RTS	VRS	NIRS	CST RTS
Australia	0.84	0.84	0.80	0.95	0.95	0.80
Austria	0.85	0.85	0.80	0.94	0.94	0.79
Belgium Flemish c.	0.86	0.86	0.83	0.98	0.98	0.83
Belgium French c.	0.82	0.82	0.77	0.90	0.90	0.77
Czech Republic	0.85	0.85	0.76	0.96	0.96	0.75
Denmark	0.80	0.80	0.76	0.95	0.95	0.76
Finland	0.90	0.90	0.83	0.98	0.98	0.82
France	0.91	0.91	0.85	0.94	0.94	0.85
Germany	0.83	0.83	0.78	0.91	0.91	0.78
Greece	0.80	0.80	0.76	0.92	0.93	0.75
Hungary[2]	0.86	0.86	0.76	0.97	0.97	0.75
Iceland	0.70	0.70	0.67	0.95	0.95	0.66
Ireland	0.90	0.89	0.82	0.97	0.97	0.81
Italy[2]	0.82	0.82	0.77	0.93	0.93	0.77
Japan	0.94	0.94	0.89	0.96	0.96	0.89
Korea	0.90	0.90	0.89	0.96	0.96	0.88
Luxembourg[2]	0.78	0.78	0.73	0.93	0.93	0.73
Mexico	0.90	0.90	0.86	0.96	0.96	0.80
Netherlands	0.88	0.88	0.83	0.95	0.95	0.83
New Zealand	0.83	0.83	0.79	0.95	0.95	0.79
Norway	0.72	0.72	0.67	0.93	0.93	0.67
Poland[2]	0.93	0.93	0.76	0.98	0.98	0.73
Portugal[2]	0.94	0.94	0.88	0.98	0.98	0.87
Slovak Republic	0.90	0.91	0.89	0.96	0.96	0.86
Spain	0.90	0.90	0.85	0.96	0.96	0.85
Sweden	0.82	0.82	0.77	0.94	0.94	0.77
Switzerland[2]	0.89	0.89	0.84	0.95	0.95	0.84
United Kingdom	0.85	0.85	0.80	0.94	0.94	0.80
United States	0.77	0.77	0.72	0.92	0.92	0.72
Standard deviation	0.06	0.06	0.06	0.02	0.02	0.06
Median	0.85	0.85	0.80	0.95	0.95	0.79
Average	0.85	0.85	0.80	0.95	0.95	0.79

Note: VRS = variable returns to scale; NIRS = non-increasing returns to scale; Cst RTS = constant returns to scale.
1. Bootstrap estimates with 2 inputs (cumulative expenditure per student and socio-economic background) and 2 outputs (average PISA score and homogeneity of PISA score).
2. Public institutions only.

Table A2.9 DEA estimates of cost efficiency at the national level without Korea

Preferred specification[1]

	Input efficiency			Output efficiency		
	VRS	NIRS	CST RTS	VRS	NIRS	CST RTS
Australia	0.89	0.89	0.80	0.97	0.97	0.79
Austria	0.87	0.87	0.79	0.94	0.94	0.79
Belgium Flemish c.	0.92	0.92	0.83	0.98	0.98	0.83
Belgium French c.	0.83	0.83	0.77	0.91	0.91	0.77
Czech Republic	0.92	0.92	0.76	0.98	0.98	0.75
Denmark	0.85	0.85	0.76	0.95	0.95	0.76
Finland	0.92	0.92	0.82	0.98	0.98	0.82
France	0.93	0.93	0.85	0.97	0.97	0.85
Germany	0.84	0.84	0.77	0.93	0.93	0.77
Greece	0.84	0.84	0.75	0.93	0.93	0.74
Hungary[2]	0.91	0.92	0.76	0.97	0.98	0.75
Iceland	0.75	0.75	0.66	0.95	0.95	0.66
Ireland	0.92	0.92	0.82	0.98	0.98	0.81
Italy[2]	0.83	0.83	0.77	0.94	0.94	0.77
Japan	0.95	0.95	0.88	0.97	0.97	0.88
Luxembourg[2]	0.80	0.80	0.73	0.93	0.93	0.73
Mexico	0.92	0.92	0.86	0.97	0.97	0.80
Netherlands	0.92	0.92	0.82	0.97	0.97	0.82
New Zealand	0.94	0.94	0.79	0.98	0.98	0.78
Norway	0.74	0.74	0.67	0.93	0.93	0.67
Poland[2]	0.93	0.93	0.76	0.98	0.98	0.73
Portugal[2]	0.96	0.95	0.88	0.98	0.98	0.87
Slovak Republic	0.92	0.92	0.88	0.97	0.97	0.86
Spain	0.94	0.94	0.85	0.97	0.97	0.84
Sweden	0.85	0.85	0.77	0.94	0.94	0.76
Switzerland[2]	0.92	0.92	0.84	0.97	0.97	0.84
United Kingdom	0.89	0.89	0.80	0.95	0.95	0.79
United States	0.78	0.78	0.72	0.92	0.92	0.72
Standard deviation	0.06	0.06	0.06	0.02	0.02	0.06
Median	0.92	0.92	0.79	0.97	0.97	0.79
Average	0.88	0.88	0.79	0.96	0.96	0.78

Note: VRS = variable returns to scale; NIRS = non-increasing returns to scale; Cst RTS = constant returns to scale.
1. Bootstrap estimates with 2 inputs (cumulative expenditure per student and socio-economic background) and 2 outputs (average PISA score and homogeneity of PISA score).
2. Public institutions only.

Table A2.10. DEA estimates of cost efficiency at the national level, public institutions only

Preferred specification[1]

	Input efficiency			Output efficiency		
	VRS	NIRS	Cst RTS	VRS	NIRS	Cst RTS
Austria	0.83	0.83	0.80	0.93	0.93	0.80
Belgium Flemish c.	0.85	0.85	0.82	0.91	0.91	0.82
Belgium French c.	0.81	0.81	0.79	0.89	0.89	0.79
Czech Republic	0.93	0.93	0.77	0.98	0.98	0.76
Denmark	0.79	0.79	0.77	0.95	0.95	0.77
Finland	0.89	0.90	0.83	0.98	0.98	0.83
Germany	0.81	0.81	0.78	0.90	0.90	0.78
Greece	0.78	0.78	0.76	0.93	0.93	0.75
Hungary	0.87	0.87	0.79	0.97	0.97	0.78
Iceland	0.70	0.70	0.68	0.95	0.95	0.68
Ireland	0.88	0.88	0.84	0.97	0.97	0.83
Italy	0.83	0.83	0.80	0.92	0.92	0.79
Japan	0.91	0.91	0.91	0.97	0.97	0.91
Korea	0.89	0.90	0.92	0.96	0.96	0.92
Luxembourg	0.78	0.78	0.75	0.93	0.93	0.75
Mexico	0.89	0.89	0.87	0.96	0.96	0.81
Netherlands	0.83	0.83	0.81	0.96	0.96	0.81
New Zealand	0.81	0.81	0.80	0.97	0.97	0.80
Norway	0.71	0.71	0.69	0.92	0.92	0.69
Poland	0.90	0.90	0.79	0.97	0.97	0.76
Portugal	0.94	0.94	0.91	0.97	0.97	0.91
Slovak Republic	0.90	0.90	0.88	0.96	0.96	0.86
Spain	0.91	0.91	0.88	0.96	0.96	0.87
Sweden	0.80	0.80	0.78	0.94	0.94	0.78
Switzerland	0.89	0.89	0.86	0.95	0.95	0.86
United Kingdom	0.83	0.83	0.80	0.94	0.94	0.80
United States	0.76	0.76	0.73	0.92	0.92	0.73
Standard deviation	0.06	0.06	0.06	0.02	0.02	0.06
Median	0.83	0.83	0.80	0.95	0.95	0.80
Average	0.84	0.84	0.81	0.95	0.95	0.80

Note: VRS = variable returns to scale; NIRS = non-increasing returns to scale; Cst RTS = constant returns to scale.

1. Bootstrap estimates with 2 inputs (teachers per 100 students and socio-economic background) and 2 outputs (average PISA score and homogeneity of PISA score).

Table A2.11. DEA estimates of cost efficiency at the national level,

using the mathematics score[1]

	Input efficiency			Output efficiency		
	VRS	NIRS	CST RTS	VRS	NIRS	CST RTS
Australia	0.84	0.85	0.82	0.96	0.96	0.82
Austria	0.87	0.87	0.83	0.95	0.95	0.83
Belgium Flemish c.	0.90	0.90	0.87	0.98	0.98	0.87
Belgium French c.	0.84	0.84	0.81	0.92	0.92	0.81
Czech Republic	0.86	0.86	0.79	0.96	0.96	0.79
Denmark	0.84	0.84	0.81	0.96	0.96	0.81
Finland	0.91	0.90	0.85	0.99	0.99	0.85
France	0.90	0.90	0.87	0.98	0.98	0.87
Germany	0.83	0.83	0.80	0.92	0.92	0.80
Greece	0.77	0.77	0.75	0.94	0.94	0.74
Hungary[2]	0.85	0.85	0.78	0.97	0.97	0.78
Iceland	0.74	0.74	0.71	0.96	0.96	0.71
Ireland	0.92	0.92	0.84	0.99	0.99	0.84
Italy[2]	0.80	0.80	0.77	0.94	0.94	0.77
Japan	0.94	0.94	0.92	0.97	0.97	0.92
Korea	0.90	0.91	0.92	0.97	0.97	0.92
Luxembourg[2]	0.78	0.78	0.75	0.95	0.95	0.75
Mexico	0.91	0.90	0.88	0.97	0.97	0.83
Netherlands	0.89	0.89	0.87	0.97	0.97	0.87
New Zealand	0.84	0.84	0.81	0.95	0.95	0.81
Norway	0.72	0.72	0.70	0.95	0.95	0.70
Poland[2]	0.93	0.93	0.79	0.98	0.98	0.77
Portugal[2]	0.96	0.96	0.90	0.99	0.99	0.90
Slovak Republic	0.91	0.91	0.92	0.97	0.97	0.90
Spain	0.95	0.95	0.88	0.99	0.99	0.87
Sweden	0.82	0.82	0.79	0.95	0.95	0.79
Switzerland[2]	0.91	0.91	0.88	0.96	0.96	0.88
United Kingdom	0.84	0.84	0.82	0.96	0.96	0.81
United States	0.77	0.77	0.74	0.93	0.93	0.73
Standard deviation	0.06	0.06	0.06	0.02	0.02	0.06
Median	0.86	0.86	0.82	0.96	0.96	0.81
Average	0.86	0.86	0.82	0.96	0.96	0.82

Note: VRS = variable returns to scale; NIRS = non-increasing returns to scale; Cst RTS = constant returns to scale.
1. Bootstrap estimates with 2 inputs (cumulative expenditure per student and socio-economic background) and 2 outputs (PISA score on the mathematics scale and homogeneity of this PISA score).
2. Public institutions only.

Table A2.12. DEA estimates of cost efficiency at the national level,

with enrolment as an output[1]

	Input efficiency			Output efficiency		
	VRS	NIRS	CST RTS	VRS	NIRS	CST RTS
Australia	0.88	0.88	0.88	0.98	0.98	0.88
Austria	0.88	0.88	0.88	0.95	0.95	0.88
Belgium Flemish c.	0.94	0.93	0.90	0.99	0.99	0.90
Belgium French c.	0.93	0.93	0.88	1.00	1.00	0.88
Czech Republic	0.94	0.93	0.91	0.99	0.99	0.91
Denmark	0.94	0.94	0.86	1.00	1.00	0.86
Finland	0.94	0.94	0.89	0.99	0.99	0.89
France	0.94	0.94	0.94	0.97	0.97	0.94
Germany	0.87	0.87	0.87	0.97	0.97	0.87
Greece	0.91	0.90	0.90	0.94	0.94	0.90
Hungary[2]	0.94	0.94	0.93	0.99	0.99	0.93
Iceland	0.78	0.78	0.76	0.99	0.99	0.76
Ireland	0.93	0.93	0.96	0.99	0.99	0.96
Italy[2]	0.87	0.87	0.87	0.95	0.95	0.87
Japan	0.94	0.94	0.98	0.99	0.99	0.98
Korea	0.94	0.94	0.95	0.98	0.98	0.95
Luxembourg[2]	0.80	0.80	0.79	0.93	0.93	0.79
Mexico	0.94	0.94	0.92	0.98	0.98	0.90
Netherlands	0.94	0.94	0.92	0.99	0.99	0.92
New Zealand	0.86	0.86	0.86	0.97	0.97	0.86
Norway	0.79	0.79	0.76	0.99	0.99	0.76
Poland[2]	0.96	0.96	0.95	0.99	0.99	0.94
Portugal[2]	0.96	0.96	0.97	0.99	0.99	0.97
Slovak Republic	0.94	0.94	0.92	0.98	0.98	0.90
Spain	0.96	0.96	0.97	0.98	0.98	0.97
Sweden	0.86	0.86	0.86	0.99	0.99	0.86
Switzerland[2]	0.92	0.92	0.92	0.97	0.97	0.92
United Kingdom	0.96	0.96	0.89	1.00	1.00	0.89
United States	0.81	0.81	0.82	0.98	0.98	0.82
Standard deviation	0.05	0.05	0.06	0.02	0.02	0.06
Median	0.94	0.93	0.90	0.99	0.99	0.90
Average	0.91	0.90	0.89	0.98	0.98	0.89

Note: VRS = variable returns to scale; NIRS = non-increasing returns to scale; Cst RTS = constant returns to scale.
1. Bootstrap estimates with 2 inputs (cumulative expenditure per student and socio-economic background) and 3 outputs
   (average PISA score, homogeneity of PISA score and enrolment rates at age 15).
2. Public institutions only.

Table A2.13. DEA estimates of cost efficiency at the national level

using the all economy PPP[1]

	Input efficiency			Output efficiency		
	VRS	NIRS	CST RTS	VRS	NIRS	CST RTS
Australia	0.84	0.84	0.80	0.95	0.95	0.80
Austria	0.85	0.85	0.80	0.94	0.94	0.79
Belgium Flemish c.	0.86	0.86	0.83	0.98	0.98	0.83
Belgium French c.	0.82	0.82	0.77	0.90	0.90	0.77
Czech Republic	0.85	0.84	0.76	0.96	0.96	0.75
Denmark	0.80	0.80	0.76	0.95	0.95	0.76
Finland	0.90	0.90	0.83	0.98	0.98	0.83
France	0.91	0.91	0.85	0.94	0.94	0.85
Germany	0.83	0.83	0.77	0.91	0.91	0.77
Greece	0.80	0.80	0.75	0.92	0.92	0.75
Hungary[2]	0.84	0.84	0.76	0.96	0.96	0.75
Iceland	0.70	0.70	0.67	0.95	0.95	0.66
Ireland	0.90	0.90	0.82	0.97	0.97	0.82
Italy[2]	0.82	0.82	0.77	0.93	0.93	0.77
Japan	0.94	0.94	0.89	0.96	0.96	0.89
Korea	0.90	0.90	0.89	0.96	0.96	0.88
Luxembourg[2]	0.78	0.78	0.73	0.93	0.93	0.73
Mexico	0.90	0.90	0.86	0.96	0.96	0.80
Netherlands	0.88	0.88	0.83	0.95	0.95	0.82
New Zealand	0.83	0.83	0.79	0.95	0.95	0.79
Norway	0.72	0.72	0.67	0.93	0.93	0.67
Poland[2]	0.93	0.93	0.77	0.98	0.98	0.75
Portugal[2]	0.94	0.94	0.88	0.98	0.98	0.87
Slovak Republic	0.90	0.90	0.90	0.96	0.96	0.86
Spain	0.90	0.90	0.85	0.96	0.96	0.85
Sweden	0.82	0.82	0.77	0.94	0.94	0.77
Switzerland[2]	0.89	0.89	0.84	0.95	0.95	0.84
United Kingdom	0.85	0.85	0.80	0.94	0.94	0.80
United States	0.77	0.77	0.72	0.92	0.92	0.72
Standard deviation	0.06	0.06	0.06	0.02	0.02	0.06
Median	0.85	0.85	0.80	0.95	0.95	0.79
Average	0.85	0.85	0.80	0.95	0.95	0.79

Note: VRS = variable returns to scale; NIRS = non-increasing returns to scale; Cst RTS = constant returns to scale.
1. Bootstrap estimates with 2 inputs (cumulative expenditure per student and socio-economic background) and 2 outputs (average PISA score and homogeneity of PISA score).
2. Public institutions only.

WORKING PAPERS

The full series of Economics Department Working Papers can be consulted at www.oecd.org/eco/Working_Papers/

545. *Monetary policy and macroeconomic stability in Latin America: the cases of Brazil, Chile, Colombia and Mexico*
(*February 2007*) Luiz de Mello and Diego Moccero

544. T*he Brazilian "tax war": the case of value-added tax competition among the states*
(*February 2007*) Luiz de Mello

543. *Public spending efficiency: institutional indicators in primary and secondary education*
(January 2007) Frédéric Gonand, Isabelle Joumard and Robert Price

542. *Enhancing turkey's growth prospects by improving formal sector business conditions*
(January 2007) Rauf Gönenç, Willi Leibfritz, Gökhan Yilmaz

541. *Fiscal relations across levels of government in Australia*
(January 2007) Vassiliki Koutsogeorgopoulou

540. *Russian manufacturing and the threat of 'Dutch Disease': A comparision of competitiveness developments in Russia and Ukrainian industry*
(January 2007) Rudiger Ahrend, Donato de Rosa and William Tompson

539. *Stimulating innovation in Russia: The role of institutions and policies*
(January 2007) Christian Gianella and William Tompson

538. *Healthcare reform in Russia: problems and prospects*
(January 2007) William Tompson

537. *A golden rule for Russia? How a rule-based fiscal policy can allow a smooth adjustment to the new terms of trade*
(January 2007) Christian Gianella

536. *From "clientelism" to a "client-centred orientation"? The challenge of public administration reform in Russia*
(January 2007) William Tompson

535. *Has the rise in debt made households more vulnerable?*
(December 2006) Nathalie Girouard, Mike Kennedy and Christophe André

534. *Social security reform in Brazil: Achievements and remaining challenges*
(December 2006) Fabio Giambiagi and Luiz de Mello

533. *Improving labour utilisation in Brazil*
(December 2006) Luiz de Mello, Naércio Menezes Filho and Luiz G. Scorzafave

532. *Boosting innovation performance in Brazil*
(December 2006) Carlos H. de Brito Cruz and Luiz de Mello

531. *Consolidating macroeconomic adjustment in Brazil*
(December 2006) Luiz de Mello and Diego Moccero

530. *Product market regulation in the non-manufacturing sectors of OECD countries: Measurement and highlights*
(December 2006) Paul Conway and Giuseppe Nicoletti

529. *The Turkish pension system: further reforms to help solve the informality problem*
(November 2006) Anne-Marie Brook and Edward Whitehouse

528. *Policies to improve Turkey's resilience to financial market shocks*
(November 2006) Anne-Marie Brook.

527. *Upgrading Japan's innovation system to sustain economic growth*
(November 2006 Randall S. Jones and Tadashi Yokoyama

526. *Strengthening the integration of Japan in the world economy to benefit more fully from globalisation*
(November 2006) Randall S. Jones and Taesik Yoon

525. *OECD's FDI regulatory restrictiveness index: Revision and extension to more economies*
(November 2006) Sven Blöndal and Alain de Serres

524. *Globalisation and inflation in the OECD economies*
(November 2006) Nigel Pain, Isabell Koske and Marte Sollie

523. *Identifying determinants of Germany's international price competitiveness – A structural VAR approach*
(November 2006) Martin Meurers

522. Short-term pain for long-term gain: the impact of structural reform on fiscal outcomes in EMU
(November 2006) Paul van den Noord and Boris Cournède

521. *Interactions between monetary and fiscal policy: How monetary conditions affect fiscal consolidation*
(November 2006) Rudiger Ahrend, Pietro Catte and Robert Price

520. *Restoring fiscal sustainability in the Euro Area: raise taxes or curb spending?*
(October 2006) Boris Cournède and Frédéric Gonand

519. *Should Measures of Fiscal Stance be Adjusted for Terms of Trade Effects*
(October 2006) David Turner

518. *Monetary policy and inflation expectations in Latin America: Long-run effects and volatility spillovers*
(October 2006) Luiz de Mello and Diego Moccero

517. *Social safety nets and structural adjustment*
(September 2006) Paul van den Noord, Nathalie Girouard and Christophe André

516. *Adapting the Icelandic education system to a changing environment*
(September 2006) Hannes Suppanz

515. *Forecasting monthly GDP for Canada*
(September 2006) Annabelle Mourougane

514. *Finland's housing market: reducing risks and improving policies*
(September 2006) Laura Vartia

513. *The Danish housing market: Less subsidy and more flexibility*
(September 2006) Espen Erlandsen, Jens Lundsgaard and Felix Huefner

512. Labour *market reform in Germany: How to improve effectiveness*
(September 2006) Eckhard Wurzel