39

The efficiency frontier as a method for gauging the performance of education at the national level

Bruno Eugène

National Bank of Belgium
bruno.eugene@nbb.be

The efficiency frontier as a method for gauging the performance of education at the national level

Bruno Eugène

National Bank of Belgium

bruno.eugene@nbb.be

This paper uses the free disposal hull framework in order to assess the degree of efficiency of 17 countries (14 'old' EU-15 Member States, the United States, Japan and Poland) in the field of education. Within this framework, a country with a high production value and limited costs is more efficient than a country with a lower production value and higher costs. This non-parametric method enables the construction of an efficiency frontier based on several indicators, which can constitute an objective to be attained. In order to do it, this paper aggregates a large number of outcome indicators, in addition to the traditional PISA indicators, and takes all costs into account.

Several of drawbacks indicate that results must be interpreted cautiously. These drawbacks aside, the analysis reveals that four countries are located on the efficiency frontier, namely Poland, Ireland, the Netherlands and Finland, while the USA or Italy are rather weak in this comparison. As a whole, the Belgian education system is more expensive but also leading to better results than the European average; Spain has just the opposite result, with low costs and rather weak results.

This working paper consists of a theoretical part and an applied part. The first section will concentrate on presenting the FDH framework, as well as the strengths and weaknesses of this approach. It will also set out how to proceed with the choice of inputs and outcomes, as well as their aggregation. The second section will look into the countries efficiency in the field of education.

This paper is based on a previously published working paper (Eugène B. (2008), The efficiency frontier as a method for gauging the performance of public expenditure: a Belgian case study, National Bank of Belgium Working Paper No. 138).

The views in this paper are those of the authors and do not necessarily reflect those of the National Bank of Belgium.

The author would like to thank Frieda Donkers, Hugues Famerée, Geert Langenus and Luc Van Meensel for their useful comments and discussions. The note also benefited from comments provided by Kris Van Cauter and Dries Dury.

All remaining errors are the authors' responsibility.

1   Introduction

The analysis of education's efficiency aims to compare the means employed on the one hand and the performance of the public education in terms of achieving its objectives on the other hand. Efficiency can thus be defined as when the best possible performance is achieved using as few inputs as possible.

There are several reasons for looking at general government efficiency. First, the high share of public expenditure in GDP generates distortive taxation. In that sense, any inefficient use of public means weighs on the economy as a whole. The second reason is also of a budgetary nature, as an ambitious fiscal policy is requested to be able to deal with the costs of an ageing population and the challenge of climate change in the longer term. Given the generally high share of public expenditure in Europe, therein lies a potentially large source of savings that needs to be addressed. Finally, a public sector that functions well is also important in the wider context of efforts to improve the competitiveness of the economy. According to some international rankings[1], a country's competitive position is largely influenced by characteristics linked to public sector performance as well as to the costs involved. While these arguments apply to all general government expenditure, including education, some additional elements concern education especially, such as the likely boost to future potential growth if the level of education is raised.

The objective of this working paper is to identify which European countries are the most efficient in their role as supplier of education. These countries can then be considered as benchmarks for the other countries. In this context, this paper is limited to the measurement of productive efficiency. The purpose is not to address directly equity considerations or macroeconomic considerations such as growth or employment objectives. Nor does it aim to explain the reasons for the relative efficiency or inefficiency of any country.

Efficiency as measured in this paper is a relative degree of efficiency, obtained by comparison with other countries' governments pursuing supposedly identical objectives. One of the main drawbacks of this analysis stems from it, as the quality of this measurement depends strongly on the quality of the selected countries used as references[2].

Amongst other weaknesses of this type of exercise, one must clearly be aware of the lack of ideal data, perfectly isolating the effects of general government action on the pursued objectives. This would notably need to also look at environmental factors, which has not been done here.

In addition, the various objectives need to be aggregated, which is an equally important problem. Consequently, even if all the desired information is available, aggregation still requires weights to be given to each of the indicators - reflecting the weights of the different objectives -,

---

[1] See, for example, World Economic Forum (2009).
[2] See Pestiau (2007).

something which cannot be done without a certain degree of value judgement. Given all these drawbacks, the results will have to be interpreted very carefully.

From a theoretical point of view, this working paper adopts the so-called Free Disposal Hull framework (FDH), a non-parametric method that enables the construction of an efficiency frontier based on several indicators, which can constitute an objective to be attained.

Compared with other studies on the subject that have used the same theoretical framework, the originality of this working paper lies in the importance that it attaches to the choice of appropriate outcome indicators, as well as the way in which these indicators are aggregated.

This working paper consists of a theoretical part and an applied part. The first section will concentrate on presenting the FDH framework, as well as the strengths and weaknesses of this approach. It will also set out how to proceed with the choice of inputs and outcomes, as well as their aggregation. The second section will look into the different countries' efficiency in the field of education, considering the 'old' EU-15 Member States, the United States, Japan and Poland, the largest of the new EU Member States. The last part presents the conclusions. All the data used in this paper are those available as at 29 April 2010 and can be found in the annexes.

2   Description of the theoretical approach

2.1   *The Free Disposal Hull (FDH) framework*

The FDH framework is a non-parametric method of constructing an efficiency frontier, along which the most efficient producers can be found, and underneath which producers that can improve their efficiency are located[3]. One of the greatest advantages of this framework is that it offers an all-encompassing, simple and easy-to-interpret view of a complex reality. This method makes it possible to express both outcome and costs or inputs borne by governments in a single synthetic indicator of efficiency.

As this method is rather intuitive, we will present it below with the help of a very brief theoretical example, based on a set of 5 producers, each using a different amount of inputs to each produce a certain volume of outcome.

---

[3] See Deprins D., L. Simar and Tulkens H. (1984).

Chart 1. FDH Framework: a theoretical example
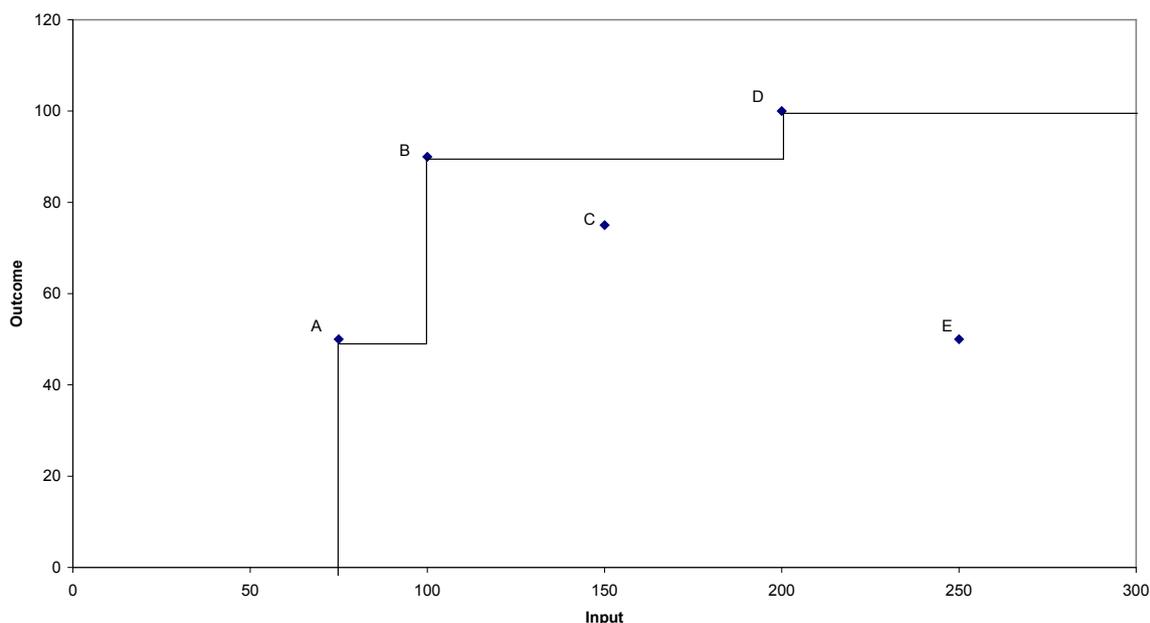
In this example, C and E are not efficient, while A, B and D are relatively efficient, since no other producer is both a user of fewer inputs and producer of more outcomes.

Efficiency can be improved either by reducing the use of inputs at constant output (horizontal leftward shift), or by increasing output at constant input (vertical upward shift), or by any combination of these two improvements. The distance to this frontier is an indicator of the degree of inefficiency.

Other methods follow this same line of intuitive reasoning and are therefore regularly used in literature on this type of subject[4]. They nevertheless differ from the FDH framework in that they require supplementary hypotheses concerning returns to scale. Data Envelopment Analysis (DEA) assumes a convex production function with either constant or decreasing returns to scale. The efficiency frontier thus encompasses the frontier as defined by the FDH framework and some countries judged to be efficient under the FDH framework are not according to the DEA method. This method is not used here for two reasons. On the one hand, it hardly affects the ranking of inefficient countries. On the other hand, the assumptions of constant or decreasing returns to scale are not necessarily established in the context of education, as it is possible for the returns to be locally decreasing. In this way, the FDH analysis seems to be justifiable and will be the only one used hereafter.

2.2   *Input measurements*

Generally speaking, an input is defined as what is used to produce a good or a service. It can be measured in physical units (number of teachers, books, etc.) or in monetary terms.

---

[4] See, in particular, Afonso and St. Aubyn (2005).

Since several inputs are implemented in the context of education, this working paper opts for the monetary measurement of inputs, which allows them to be easily aggregated. Moreover, from the point of view of public expenditure, it is more important to achieve the best performance at the lowest possible cost to the budget rather than reach a high technical degree of efficiency. Therefore, the choices made by public sector authorities must imperatively take account of relative factor cost, some countries having an advantage in being capital intensive and others labour intensive. Finally, the two measurements are actually closely related.

It is also important to ensure that the data from the various countries are comparable. This implies first of all a common monetary expression, consisting of expenditure either as a percentage of GDP or in absolute terms per capita, using the purchasing power parity method. Next, it is important to consider general government expenditure contributing to the outcome, such as compensation of teachers, intermediate consumption and gross fixed capital formation[5]. Lastly, private expenditure on education has also been taken into account in order to smooth out any differences there may be between the various countries in terms of modes of financing. Since it is not possible to distinguish between the different impact of public or private expenditure on the performance of a country, it is necessary to consider the total expenditure in question.

2.3   Outcome measurements

2.3.1   Output versus outcome

In the case of production of a marketable good, the measurement of output produced is directly linked to the value of this production on the market. But the value of non-marketable goods or services, such as education, is not set by the market and therefore has to be measured in a different way.

According to the national accounts, the value of government output is equal to the total supported costs, i.e. intermediate consumption, compensation of employees, other taxes less subsidies and consumption of fixed capital. This definition is not useful here, since efficiency seen as the relationship between total value produced and the cost of production would be close to a unit value by assumption. Therefore, other measurements of the objectives pursued need to be found.

The direct result of production is quite often quantifiable itself, too. It can include, for instance, the number of class-hours taught. But these "outputs", which are also occasionally regarded as inputs by some authors, do not reflect the final objective pursued by the policies in place.

These final objectives, yet to be identified, are called outcomes. The acquisition of knowledge and skills can be cited as a potential final objective of the educational system. Such

---

[5] In order to avoid taking into account a year in which investment was particularly high or low, it would have been better to take consumption of fixed capital into consideration but this was not possible due to data limitations.

objectives should be clearly identified, something which is all the more difficult since there are many different objectives pursued, and measured, which poses many problems, as will be seen later on in this paper.

2.3.2   Method used for the selection of outcome indicators

An initial selection of potentially suitable outcome indicators for measuring performance in education was carried out following a positive approach. The basic principle is to identify what the general government's objectives might be regardless of whether the indicators are available or not. On this basis, the second step is to look at the available indicators and to contemplate whether or not to take into account one or another of the available indicators. Since no single indicator is conceivable in the context of policies having various different goals, like education, it seemed useful to aggregate several indicators.

The selection process subsequently followed a few basic rules as a result of which some indicators were dropped. Certain indicators were automatically rejected for the simple reason that they did not cover Belgium, which was the focus of previous work. For similar reasons, indicators targeting too small a number of countries were also ruled out. In addition, if the definition of what was being measured by an indicator was not harmonised amongst countries, then performances could not be compared and the indicator would be devoid of interest. Moreover, relevant indicators for some parts of the world were less relevant in the context of a comparison of OECD member countries.

The end result of this selection was a set of performance indicators in education that then had to be aggregated in order to obtain a synthetic indicator of performance in education.

2.3.3   Method used for aggregating the different outcome indicators

In the theoretical example depicted in chart 1, one single input and one single output have been considered. In such a context, it is quite easy to determine the efficiency frontier. We have also seen earlier on in this paper that the choice of a monetary expression for inputs enables all inputs to be taken into consideration without major aggregation problems. For outcomes, however, there are several problems.

The aim is to obtain a synthetic outcome indicator that can be used in measuring the efficiency of education. Therefore, it is necessary to aggregate indicators that do not always have identical units (results for the PISA tests and percentage of the population with a secondary education, for example) or the same importance.

Standardisation of indicators

The objective here is to transform the indicator values into comparable units that can then be aggregated. That conveys an average value and an identical standard deviation for each of the

standardised indicators. The standardisation of indicators has therefore been systematically treated as follows:

$$OS_i = (O_i - \bar{O})/SD \qquad\qquad (1)$$

where OS_i is the standardised indicator for an outcome for country i,

O$_i$ is the indicator for an outcome for country i before standardisation,

Ō is the arithmetic mean of the different countries considered[6] for this outcome and

SD the standard deviation of the different countries considered for this outcome.

In this way, the average across the different countries for each indicator is 0 and the standard deviation for each indicator over the European countries is 1. This choice of standardisation makes it possible to aggregate the different indicators. It also gives them an equal weight, since the average deviations from the mean are identical for all the indicators. Without standardisation, the indicator with a large standard deviation would carry more weight than the one with a narrow standard deviation. The standardisation has thus removed the weighting differences of statistical origin.

Weighting the indicators

While some studies[7] refrain from giving different weights to the different outcome indicators, considering that this is a neutral stance, this paper chooses another option, because not weighting the indicators is tantamount to attaching the same importance to each one, which effectively means a choice that is far from neutral.

In this working paper, different weights have therefore been explicitly assigned to the different indicators. The weightings adopted in this way have thus been largely decided by the author's own choice[8]. Among the factors determining the weightings, it should be pointed out that the author wanted to:

- give preference to indicators that appeared to provide the best measurement of the objectives supposedly being pursued;
- give preference to any potential indicators that depend directly on public sector action;
- make sure that two similar indicators are not overweighted in total;
- give less weight to indicators derived from surveys with small samples.

For these last two reasons, using an endogenous and non-uniform weighting[9] is not suitable. Such a procedure effectively allows different weights to be given to different indicators for each country, since the preferences of individual states are, in principle, not the same. The suggested

---

[6] In the practical part of this working paper, the mean and the standard deviation will be calculated on the sample taken from the old EU-15 countries.
[7] See, for example, Sociaal en Cultureel Planbureau (2004).
[8] The choice of weightings was nevertheless influenced by deliberations on the matter with colleagues from the Research Department of the National Bank of Belgium.
[9] See Bowen and Moesen (2007).

procedure considers that countries' preferences are revealed when the various indicators are weighted so that the synthetic outcome indicator for each country is maximised. If the set of indicators is large, this could result in a high degree of efficiency for most countries, at least for each country leading one of the partial indicators' rankings. In addition, such a procedure does not allow any correction for the near redundancy of certain indicators nor for the relative robustness of some of these indicators. In order to do this, minimum and maximum weights should be determined for each indicator. In an extreme-case scenario, when many indicators are considered, this would come close to a fixed weight for each indicator, identical for all countries, which is what has been done here.

A correction had to be made for processing indicators that do not cover all countries. In the simplest cases, data were only missing for a year, in which case the nearest year available was used. In other cases, the indicators do not cover one or two countries, but this mainly concerns the United States and Japan. So as not to lose countries for which we might not have all the indicators, which would have led to a smaller sample and therefore to a truncated efficiency frontier, a simple correction procedure was used. This consists, for those countries, of constructing a synthetic indicator made up of the weighted average of all available indicators and then correcting for the weight of the indicators that are not available. In practice, if 20 p.c. of the weighted indicators are not available, the result obtained from the other indicators will be multiplied by 1.25. Any corrections made in this way have been very limited in both number and importance. These corrections will nevertheless be mentioned in due course in the practical part of this working paper.

2.4   *Shortcomings of the approach followed in this paper and attempts to limit them*

Two kinds of weakness need to be addressed before an efficiency frontier can be established. The first comes from the FDH framework itself, while the second is more a result of the imperfect indicators available and needed for the concrete application of this framework.

Simplification of reality

The simplicity of the methodology used in this paper is also one of its main weaknesses, since it reduces the education outcome to just a few parameters, while the situation is much more complex in practice. While some previous studies limited the measure of outcome to a few parameters, this paper takes more indicators into account, in order to get a better measure of the many objectives pursued.

Also, in many cases, the outcome indicators are no more than rough estimates of what one actually wants to measure. The ideal scenario for measuring the efficiency of education would be to be able to clearly identify the effect of the general government's action on it. In order to do this, one would have to be able to measure performance both with and without such intervention by the public authorities, something which is not possible in practice. Inherent in the indicators are in fact

some elements that are not directly linked to public-sector action. Thus, for instance, language skills are influenced notably by the quality of language education in a country, but also by its institutional context or the frequency of contacts between people speaking different languages. These so-called environmental or external factors are not accounted for here. In a following step, however, they should be introduced in an econometric work, considering the outcome as a dependent variable and the input, external factors and general government efficiency - the residual element - as explaining variables.

Finally, on the input side, there is no set rule either as to what should be taken into consideration. Identification of government spending has recently become harder by the increasing use of public-private partnerships, for example.

A relative measure of efficiency

As already mentioned, efficiency as measured here is only relative, as a function of general government efficiency in the other countries considered. Since there is nothing to indicate that these countries are efficient themselves, any potential efficiency gains identified here should be regarded as being a minimum possible.

Time lags

Policy results sometimes take years to materialise. In the field of education, expenditure made in one year will probably only yield results over a period of timeThis time lag also negatively influences the input efficiency, as there can be no immediate adjustment when objectives change.

Limitations stemming from the use of survey results

Problems related to the use of surveys must be pointed out, too. The partial indicators from the World Economic Forum Global Competitiveness Report[10] (WEF) or the IMD World Competitiveness Yearbook[11] (IMD) - references that are frequently used for measuring public sector efficiency – are, in fact, sometimes based on samples whose representativity is debatable. Most of the indicators published in the first case derive from surveys carried out among "business executives, (...) having also knowledge and experience of the global environment"[12]. In 2006, the sample covered 74 people in Belgium, but only 35 in Ireland or 51 in Germany, for example. In the second case, all the survey data came from a sample of about 4,000 people, which works out at an average of only 66 per country under consideration. Nevertheless, these surveys are mainly carried out among business executives in multinational corporations, so respondents are generally in a position to compare the situation in one country with their previous experience in other countries.

---

[10] World Economic Forum (2009).
[11] IMD World Competitiveness Yearbook (2009).
[12] World Economic Forum (2006), p. 125.

Rather than excluding survey results altogether, these indicators have generally been given a limited weighting. However, they still exert some influence, their weight being 25 p.c. in the synthetic outcome indicator.

In addition, the WEF or IMD indicators are more dedicated than others to measuring efficiency from the point of view of business life. Diversification of the indicators used should enable the best possible measurement of all the various objectives pursued, ranging from business profitability to individual well-being, as well as elements that are more closely linked to the notion of equity or life in society. Therefore, and despite the wide range of indicators used, it must be borne in mind that any bias in favour of certain objectives of improved business efficiency may distort the synthetic indicators used in the practical part of this working paper.

Lack of data harmonisation

Another weakness concerns the lack of harmonisation of the data available, as well as the fact that data for some countries are incomplete. This problem has been reduced on the one hand by using harmonised data availability as a criterion for indicator selection and, on the other hand, by interpolating the missing data for certain countries, if need be (see section 2.3.3). When a country subject to the interpolation procedure was found to be efficient, the whole efficiency frontier limited to the available indicators for this country was rebuilt. If its efficiency then remained valid - as was always the case here -, the country also continued to be considered as efficient for the whole set of indicators.

Multicollinearity

Within the range of available data, some are more or less closely linked. Therefore, the individual indicators should be tested for multicollinearity, which could lead to the rejection of some indicators. Rather than such a radical choice, the weight of obviously linked indicators has been reduced in order to avoid an over-representation of some factors in the global indicators.

In view of the above-mentioned limitations, the findings should be interpreted with great caution.

3   Presentation of the results

This part analyses the performance of several countries in one of their main functions in terms of expenditure, namely education. Relative efficiency is measured amongst 14 of the 15 'old' EU Member States. In addition, the United States, Japan and Poland – the biggest country among the new EU Member States - have also been considered. As far as performance is concerned, the European average, that excludes Luxembourg for which essential input data are missing, is constructed by weighting each country's result by its population size.

In what follows, the selected outcome indicators will be presented along with their weightings in the synthetic indicator. The individual figures from these indicators will be given in

the annex. New data frequently become available; those used here are statistics available as at 29 April 2010. Once aggregated, these indicators make it possible to construct the efficiency frontier, which will then be set out and commented on.

3.1   Outcome indicators

As regards education, the findings of the OECD's Programme for International Student Assessment (PISA) are frequently used in international comparisons. They are regularly updated and relate to skills of 15 year-old pupils – those who have almost reached the end of compulsory schooling in most countries - in reading, mathematics and science. Up to 2003, PISA also tested the capacity of pupils in problem-solving, which is still considered here. Altogether, these four indicators were given a total weight of 50 p.c. in the synthetic indicator for education. These indicators will often be referred to as a measure of achievement, in other words, the level of knowledge reached. One of the advantages of these indicators is that they cover a wide survey of pupils, which makes them more representative than the traditional survey findings. Moreover, these indicators are harmonised, since the same tests are carried out in all the countries studied. Finally, the results of the first three of these indicators - reading, mathematics, science – are also available separately for some regions[13].

Apart from the general standardisation applied to all indicators, the PISA indicators have been subject to a specific treatment. The scores obtained by pupils in the 5th, 25th, 50th, 75th and 95th percentiles were added together before being standardised. Taking these different percentiles into account – rather than just settling for the average or the median – is tantamount to considering that extreme results are also of importance[14].

One of the disadvantages of the PISA indicators, however, is the fact that their scope is limited to education up to the age of 15, on the one hand, and to certain subjects, on the other hand. They therefore need to be supplemented by other indicators that make it possible to measure acquisition of other competences and performance in education beyond that age.

Two additional indicators stem from hard data: number of students completing secondary and tertiary education. Unlike the PISA data, these indicators are not harmonised, since it is probably harder to obtain one of the qualifications in question in some countries than in others. Within the selected countries, however, the differences should be limited. In addition, the question can be raised whether these are correct outcome indicators. In this paper, they are considered as a

---

[13] For example for the Flemish and French speaking Communities, that are responsible for education in Belgium. This makes it possible to distinguish between the performance of the two main (language) Communities in Belgium. In Spain, this is also available for Andalusia, Basque Country, Cantabria, Galicia, La Rioja, Castilló and León, Navarre, Aragón, Catalonia and Asturias.

[14] Considering the standard deviation separately would have been preferable, but led to a new indicator that also had to be weighted. The five scores considered and aggregated allows to limit the number of indicators whithout losing much of the information. This way, the indicator will be positively influenced by a small standard deviation.

proxy of the contribution to labour force qualifications, and therefore as an outcome of the educational system. Moreover, they are exhaustive hard data indicators. These two indicators together count for 25 p.c. in the synthetic indicator.

Table - Outcome indicators for education

	Type	Source	Unit, question asked or reference group	Weight in synthetic indicator (in percentages)	Top performers
Reading skills	Student assessment	OECD (PISA)	15 year-old pupils	12.5	Finland, Ireland and Poland
Mathematics performances	Student assessment	OECD (PISA)	15 year-old pupils	12.5	Finland, Netherlands and Japan
Scientific literacy skills	Student assessment	OECD (PISA)	15 year-old pupils	12.5	Finland, Japan and the Netherlands
Problem-solving skills[15]	Student assessment	OECD (PISA)	15 year-old pupils	12.5	Finland, Japan and Denmark
Educational attainment: secondary education	Hard data	OECD	Percentage of 25-34 year-olds with an upper secondary education	12.5	Japan, Poland and Sweden
Educational attainment: tertiary education	Hard data	OECD	Percentage of 25-34 year-olds with a tertiary education	12.5	Japan, Ireland, Belgium and France
Language skills	Survey	IMD	"Language skills meet the needs of the society?"	10	Denmark, the Nehterlands and Sweden
Quality of educational system[16]	Survey	WEF+IMD	Average of 3 questions	10	Finland, Denmark and Belgium
Availability of skilled labour	Survey	IMD	"Skilled labour is readily available?"	5	Ireland, Japan and Sweden

---

[15] This indicator does not cover the United Kingdom.
[16] This indicator is made up of 3 sub-indicators: "the educational system meets the needs of a competitive economy?", according to the IMD and WEF, and "university education meets the needs of a competitive economy?".

Lastly, to round off coverage of the objectives that should be assigned to education as a whole, four indicators derived from surveys have also been taken into account. The first concerns knowledge of foreign languages and thus extends the scope of school achievement. It was given a weighting of 10 p.c. As for confidence in the quality of the education system, its perceived quality and the availability of an appropriate supply of workers on the labour market, their statistical weakness – related to the small size of the survey sample – and question formulation to a large extent geared towards the competitiveness of the economy led to them being given a smaller weighting of 5 p.c. each.

Among the indicators that have not been retained, that concerning the survival rate in tertiary education seemed to have no different statistical content for the number of students completing secondary and tertiary education.

3.2   Expenditure on education

Expenditure on education can either be expressed as a percentage of GDP or as a cost per pupil. In the first case, the advantage is that the cost of living – and thus the wage - differences among countries are corrected by using GDP as a reference. However, this does not make it possible to take into consideration one of the main factors determining levels of expenditure on education, i.e. the number of pupils taught. Therefore, the basic unit used below is cost per pupil, adjusted for purchasing power parity. This last correction is not fully satisfactory, but seemed less detrimental to the measurement of costs than an expression in percentage of GDP, as the proportion of pupils in each country's population is different. However, it must be kept in mind that this measure is detrimental (advantageous) to countries with a relatively high (low) GDP.
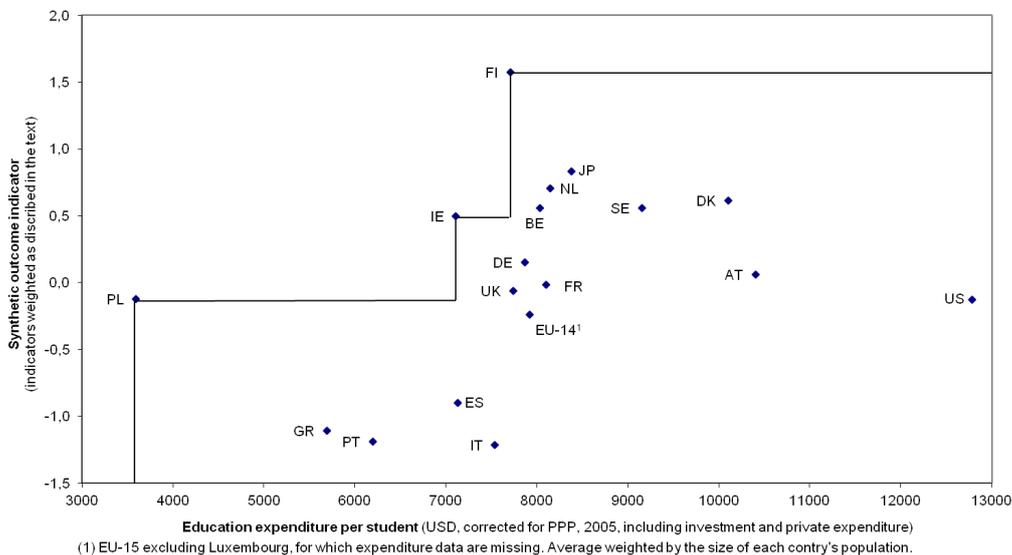
Costs under consideration include all types of teaching, all levels of education, all sources of funding – public and private – and all types of expenditure, i.e. investment costs related to education too.

3.3   Efficiency in education

Aggregating the nine indicators selected and the total of the costs taken into consideration puts three countries on the efficiency frontier, namely Poland, Ireland and Finland. These countries' efficiency levels in the field of education are particularly high in relation to the resources employed in this sector. The next chart makes it possible to identify some relationships amongst countries. So, for example, Ireland is the only country leading Spain in the efficiency ranking. It should also be pointed out that Portugal and Italy are inefficient compared to Spain. The Spanish option, for example, is to spend less money than the European average, which results in a lower performance. However, it cannot be said that one is more efficient than the other. This also applies to the European average compared to the United States: the Europeans spent far less, but also have a lower outcome than the US. However, the marginal gain to reach the American level seems to

cost quite a lot. On the contrary, Japan reaches a much higher outcome while spending only marginally more.

Chart 2 - Efficiency frontier of the education system (whole system)

Education expenditure per student (USD, corrected for PPP, 2005, including investment and private expenditure)
(1) EU-15 excluding Luxembourg, for which expenditure data are missing. Average weighted by the size of each contry's population.

In order to join the most efficient group of countries, Spain could try to be as efficient as Ireland. These countries spend about the same on education while performing much better in terms of outcome. Such a conclusion, however, is strongly limited by the weaknesses of the framework, as explained above.

4 Conclusion

The aim of this working paper was to try and determine to what extent European governments efficiently fulfil their role in the area of education.

The analysis presented here is based on the Free Disposal Hull framework, which has simple principles and is easy to interpret. Efficiency is established in relation to other countries, comparing resources deployed and the value of production. A country with a high production value and limited costs is thus more efficient than a country with a lower production value and higher costs. Taking all the efficient countries together enables an efficiency frontier to be established as an efficiency target for the other countries to meet.

The limits to this analysis are mainly to be found in measuring the value of production. Because the value of education is not generally determined by market forces, it has to be estimated with the help of other elements. These elements, referred to as outcome, should be a measure of the extent to which public authorities meet their targets. This implies being able to identify multiple objectives and to measure the extent to which they are achieved.

A first group of limitations to this framework has to do with the aggregation of the various outcome measurements. Indeed, the plethora of objectives pursued has to be aggregated into one single outcome indicator, as the costs of meeting the different objectives are not divisible. We have shown that aggregating the various sub-indicators into one single synthetic outcome indicator could not avoid some degree of subjectivity, reflected in the weight given to each of the sub-indicators. There is nothing wrong with the idea of giving an identical weighting, as has often been done in earlier research work, but it is certainly not a guarantee of objectivity.

The second series of limitations is related to the indicators themselves. Particular attention has been paid to the choice of indicators in this working paper, with a wide range of indicators being used. However, the measurements made in this way are still not perfect and should therefore be treated with caution.

These words of caution aside, the analysis reveals that three of the countries taken into account in this paper can be considered as efficient: Poland, Ireland and Finland. On average, the European education system is less expensive but also produces lower results than the United States. Compared to Japan, Europe appears to have lower results, while paying only marginally less for the education of its pupils.

Although this working paper provides some indication of the efficiency of education in an international perspective, there is still a lot of research work to be done in this area, for instance to map efficiency developments over time and to see why some countries are more efficient than others.

5   References

Afonso A. and M. St. Aubyn (2005), *Non-parametrical approaches to education and health efficiency in OECD countries*, Journal of Applied Economics, vol. VIII, No. 2, pp. 227-246.

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

Bowen H. and W. Moesen (2007), *Benchmarking the competitiveness of nations: benevolence versus equal treatment*, presented at CESifo Venice Summer Institute, Workshop on "The Many Dimensions of Competitiveness", Venice International University, San Servolo, 20-21 July 2007.

Deprins D., L. Simar and H. Tulkens (1984), *Measuring Labor-Efficiency in Post Offices*, in Marchand, M., P. Pestiau and H. Tulkens (eds.), The Performance of Public Enterprises: Concepts and Measurement, North-Holland Publishing Co., Amsterdam, pp. 243-268.

Eugène B. (2008), *The efficiency frontier as a method for gauging the performance of public expenditure : a Belgian case study*, National Bank of Belgium Working Paper No. 138.

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

IMD World Competitiveness Yearbook (2009), Lausanne.

OECD (2005), Education at a glance, 2005.

OECD (2006), Education at a glance, 2006.

OECD (2007), Education at a glance, 2007.

OECD (2009), Education at a glance, 2009.

Pestiau P. (2007), *Assessing the performance of the public sector*, CREPP Working paper 2007/03.

Sociaal en Cultureel Planbureau (2004), *Prestaties van de publieke sector: Een internationale vergelijking van onderwijs, gezondheidszorg, politie/justitie en openbaar bestuur*, Den Haag.

Sutherland D., R. Price, I. Joumard and C. Nicq (2007), *Performance indicators for public spending efficiency in primary and secondary education,* OECD Economics Department Working Papers No. 546.

World Economic Forum (2006), *The Global Competitiveness Report 2006-2007*, Palgrave Macmillan, Geneva.

World Economic Forum (2009), *The Global Competitiveness Report 2009-2010*, Palgrave Macmillan, Geneva.

6 Appendix: Database

	Outcome indicators											Expenditure	
	Reading skills 15 year-olds (sum of P-5, P-25, mean, P-75 and P-95 values)	Mathematics performances 15 year-olds (sum of P-5, P-25, mean, P-75 and P-95 values)	Scientific literacy skills 15 year-olds (sum of P-5, P-25, mean, P-75 and P-95 values)	Problem solving skills 15 year-olds (sum of P-5, P-25, mean, P-75 and P-95 values)	Language skills are (not) meeting the needs of the enterprises	Educational attainment secondary education (percentage of 25-34 year-olds with an upper secondary education)	Educational attainment tertiary education (percentage of 25-34 year-olds with a tertiary education)	Quality of the educational system in your country (1= does not meet the needs of a competitive economy; 7=meets the needs of a competitive economy?") [3]	Quality of the educational system ("the educational system (does not) meet(s) the needs of a competitive economy?") [3]	Quality of the educational system ("the University education (does not) meet(s) the needs of a competitive economy?") [3]	Availability of skilled labor ("skilled labor is (not) readily available")	Annual expenditure on educational institutions per student (in equivalent US dollars converted using PPPs for GDP, by level of education, based on full-time equivalents)	Total population (thousands)
	2006	2006	2006	2003	2009	2007	2007	2009	2009	2009	2009	2002	2008
BE	2468	2584	2533	2604	7.8	82	41	5.6	6.6	7.0	6.5	7933	10622.0
DK	2461	2561	2478	2628	8.7	85	40	5.7	7.5	7.9	6.9	9261	5492.0
DE	2452	2518	2567	2552	6.9	85	23	4.7	6.1	6.8	6.5	7129	82120.0
GR	2273	2290	2359	2238	6.5	75	28	3.2	3.1	3.2	5.0	4136	11242.1
ES	2287	2397	2438	2401	2.6	65	39	3.4	3.3	3.4	4.4	5914	45593.4
FR	2409	2469	2462	2584	3.4	83	41	4.8	5.9	5.4	6.1	7467	64120.2
IE	2575	2505	2538	2487	5.2	83	44	5.6	6.9	7.2	7.9	5711	4443.0
IT	2320	2309	2375	2331	3.6	68	19	3.3	4.2	4.2	5.3	7708	59832.2
NL	2512	2647	2614	2597	8.5	83	37	5.2	6.9	7.1	6.0	7241	16440.0
AT	2429	2515	2540	2526	6.6	87	19	4.9	6.7	7.2	6.6	8943	8343.9
PT	2344	2327	2369	2338	7.0	44	21	3.6	4.3	5.6	5.2	6080	10622.4
FI	2728	2737	2810	2733	8.1	90	39	5.9	8.5	8.1	6.9	7304	5313.0
SE	2521	2512	2512	2538	8.3	91	40	5.3	5.9	6.9	6.9	8520	9219.0
UK	2462	2481	2569	n.a.	3.8	75	37	4.6	5.1	6.2	5.6	6691	61350.0
US	2462[1]	2376	2448	2382	4.6	87	40	4.8	4.8	7.1	6.7	11152	304530.0
JP	2471	2610	2640	2716	3.2	94[1]	54	4.5	5.5	4.8	7.2	7438	127842.0
PL	2526	2478	2490	2433	5.5	92	30	4.1	4.0	4.6	4.3	2962	38116.0

[1] 2003.
[2] 2002.
[3] These three indicators have been aggregated.
[+] Weighted average of Flemish Community (60 %) and French Community (40 %)