Eugène, Bruno

Working Paper

The efficiency frontier as a method for gauging the performance of public expenditure: a Belgian case study

NBB Working Paper, No. 138

Provided in Cooperation with:
National Bank of Belgium, Brussels


This Version is available at:
http://hdl.handle.net/10419/144351

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
The efficiency frontier as a method for gauging the performance of public expenditure: a Belgian case study

by Bruno Eugène

September 2008  No 138
Statement of purpose:
The purpose of these working papers is to promote the circulation of research results (Research Series) and analytical studies (Documents Series) made within the National Bank of Belgium or presented by external economists in seminars, conferences and conventions organised by the Bank. The aim is therefore to provide a platform for discussion. The opinions expressed are strictly those of the authors and do not necessarily reflect the views of the National Bank of Belgium.

Orders
For orders and information on subscriptions and reductions: National Bank of Belgium, Documentation - Publications service, boulevard de Berlaimont 14, 1000 Brussels.
Tel +32 2 221 20 33 - Fax +32 2 21 30 42


© National Bank of Belgium, Brussels
All rights reserved.
Reproduction for educational and non-commercial purposes is permitted provided that the source is acknowledged.

ISSN: 1375-680X (print)
ISSN: 1784-2476 (online)
Abstract

This paper uses the Free Disposal Hull framework in order to assess the relative efficiency of Belgian general government in the field of health care, education and public order and safety. In order to do so, this paper aggregates a large number of outcome indicators.

Several drawbacks indicate that results must be interpreted cautiously. These drawbacks aside, the analysis reveals that Belgium is relatively efficient in the field of health care. As a whole, the Belgian education system is more expensive but also produces better results than the European average. However, an analysis based on a limited set of indicators reveals that the French-speaking education sector is very inefficient while the Flemish Community’s efficiency is markedly better. As far as public order and safety are concerned, major improvements could and should be made, either to improve service or cut costs.

Key Words: Public spending efficiency, FDH.
JEL Classification: H11, H51, H52, H59.

Corresponding author:

Bruno Eugène, NBB, Research Department, e-mail: bruno.eugene@nbb.be.

A preliminary version of this paper was presented at the ECB Public Finance Workshop on "Challenges for government spending in the EU", held on 6 December 2007.

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, Dries Dury and Ivan Matalík.

The views in this paper are those of the author and do not necessarily reflect those of the National Bank of Belgium. All remaining errors are the author’s responsibility.
# TABLE OF CONTENTS

1. Introduction .............................................................................................................................1

2. Description of the theoretical approach ................................................................................2
   2.1. The Free Disposal Hull (FDH) framework ................................................................. 2
   2.2. Input measurements ........................................................................................................4
   2.3. Outcome measurements ............................................................................................... 5
       2.3.1. Output versus outcome.......................................................................................... 5
       2.3.2. Method used for the selection of outcome indicators .............................................. 5
       2.3.3. Method used for aggregating the different outcome indicators............................ 6
   2.4. Shortcomings of the approach followed in this paper and attempts to limit them .......... 8

3. Presentation of the results ...................................................................................................11
   3.1. Health care ......................................................................................................................11
       3.1.1. Outcome indicators ...............................................................................................11
       3.1.2. Expenditure on health care ..................................................................................15
       3.1.3. Efficiency in the field of health care ....................................................................16
   3.2. Education .......................................................................................................................19
       3.2.1. Outcome indicators ...............................................................................................19
       3.2.2. Expenditure on education .....................................................................................21
       3.2.3. Efficiency in education .........................................................................................22
   3.3. Public order and safety ...................................................................................................24
       3.2.1. Outcome indicators ...............................................................................................24
       3.2.2. Expenditure on public order and safety .................................................................26
       3.2.3. Efficiency in public order and safety ......................................................................26

4. Conclusion ............................................................................................................................27

References ..................................................................................................................................29
Appendix: Database.....................................................................................................................31
National Bank of Belgium - Working papers series .....................................................................35
1. **Introduction**

The analysis of general government efficiency aims to compare the means employed by the general government on the one hand and the performance of the public services in terms of achieving their 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 high share of public expenditure in Belgium, 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.

The objective of this working paper is to measure to what extent the Belgian general government is efficient or not in its role as supplier of certain community services. 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 Belgian general government.

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.

---

\(^1\) See, for example, World Economic Forum (2006).

\(^2\) See Pestiaux (2007).
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 -, 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, on the one hand, in the importance that it attaches to the choice of appropriate outcome indicators, as well as the way in which these indicators are aggregated. On the other hand, the accent is clearly on the efficiency of the Belgian public sector. Moreover, unlike some previous studies, the various functions of general government will not be aggregated, with health care, education and public order and safety being analysed separately. Research has also been done on general public services, but is not published here due to the weakness of the available indicators and the difficulty in identifying the objectives pursued by general government in this field.

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 Belgium’s efficiency in the field of health care, education and public order and safety, taking as a benchmark 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 2 June 2008 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

---

3 These functional categories are based on the United Nations Classification of the Functions of Government (COFOG).
efficiency are located⁴. 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 produce a certain volume of outcome.

In this example, C and E are not efficient, while A, B and D 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.

Chart 1. FDH Framework: a theoretical example

Other methods follow this same line of intuitive reasoning and are therefore regularly used in literature on this type of subject⁵. They nevertheless differ from the FDH framework in that they

---

⁵ See, in particular, Afonso and St. Aubyn (2005).
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, from a practical point of view, 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 present context, 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, doctors, etc.) or in monetary terms.

Since several inputs are implemented in this context, 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 costs, some countries having an advantage in being capital intensive and others labour intensive. Finally, the two measurements are actually closely related; the monetary one includes the physical one, as a first step in the production process is to get physical inputs by spending money and the second step is to reach some degree of outcome with those physical inputs.

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 employees, intermediate consumption and gross fixed capital formation\(^6\). Lastly, private expenditure has also been taken into account sometimes 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 influence of public or private expenditure on the performance of a country, it is necessary to consider the total expenditure in question.

\(^6\) 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.
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 those considered in this study, 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. These are, for instance, number of class-hours taught in education or the number of hospital beds available for the population in health care. 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 are called outcomes. Improvement in health, increase in life expectancy or the acquisition of knowledge and skills can be cited as potential final objectives of the health care and educational system. These objectives should be clearly identified, something which is all the more difficult since there are often 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 each area under examination 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, it seemed useful to aggregate several indicators, for each of the four public sector functions examined.

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, the main focus of this working paper. 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. Some indicators have also been rejected for more specific reasons, as will be stated in the empirical part of this working paper.

The end result of this selection was a set of performance indicators that then had to be aggregated.

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 the various public sector functions studied. Therefore, it is necessary to aggregate indicators that do not always have identical units (life expectancy in years and infant mortality in percentage for health care, 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:

\[
\text{OS}_i = \frac{(O_i - \bar{O})}{\text{SD}}
\]

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,
- \( \bar{O} \) is the arithmetic mean of the different countries considered\(^7\) for this outcome and
- \( \text{SD} \) 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 same sample is 1. This choice of standardisation makes it

\(^7\) 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.
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\(^8\) 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\(^9\). 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\(^10\) 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 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.

---

8 See, for example, Sociaal en Cultureel Planbureau (2004).
9 The choice of weightings was nevertheless influenced by deliberations on the matter with colleagues from the Research Department of the National Bank of Belgium.
10 See Bowen and Moesen (2007).
A correction had to be made for processing indicators that do not cover all countries\textsuperscript{11}. In the simplest cases, data were only missing for a year, in which case the nearest year available was used (or the average over several years). In other cases, the indicators do not cover one or several countries, but this mainly concerns the United States, Japan or Poland. 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.

The approach followed here enables indicators to be aggregated to the desired level. Thus, the remainder of this working paper will look separately into three public sector functions, making it possible to compare what is comparable, without pushing the degree of aggregation too far. So, an overall performance indicator for general government as a whole would lead to aggregation of elements that have nothing in common. An overall indicator has therefore not been established, although some exploratory studies\textsuperscript{12} have done this.

\textbf{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.

\textit{Simplification of reality}

The simplicity of the methodology used in this paper is also one of its main weaknesses, since it reduces general government outcome to just a few parameters, while the situation is often 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 general government would be

\textsuperscript{11} As mentioned in the choice of indicators, Belgium is never concerned here.

\textsuperscript{12} See Afonso \textit{et al} (2003).
to be able to clearly identify the effect of its action. 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, life expectancy is influenced notably by the quality of the health care system in a country, but also by diet or environmental factors. 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. Whether it is a question of performance in the field of education or health care, expenditure made in one year will probably only yield results over a period of time, even up to the entire life span of an individual. This, for instance, would be the case if the infant mortality rate were to be brought down, with the impact being not only direct but also indirect in terms of the resultant increase in average life expectancy.

This 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\textsuperscript{13} (WEF) or the IMD World Competitiveness Yearbook\textsuperscript{14} (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

\textsuperscript{13} World Economic Forum (2007).
\textsuperscript{14} IMD World Competitiveness Yearbook (2008).
published in the first case derive from surveys carried out among "business executives, (...) having also knowledge and experience of the global environment"\textsuperscript{15}. 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 come 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.

Rather than excluding survey results altogether, these indicators have generally been given a limited weighting. However, they still exert some influence, particularly in the case of the outcome indicator in the field of public order and safety. The total weight of survey data in the outcome indicator in these areas is about 80 p.c.

During the exploratory work for this working paper, a trial to measure the outcome for general public services was also carried out. However, as outcome indicators relied only upon survey data, they were found to be too weak and therefore are not mentioned here.

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.

\textit{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 the Belgian general government sector in three of its main functions, namely health care, education and public order and safety. Efficiency is measured taking as a benchmark the 14 other ‘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 for health care and education. As far as performance is concerned, the European average, that sometimes excludes certain countries for which essential data are missing, is constructed by weighting each country’s result by either its GDP or population size.

For each of the three general government sector functions analysed, the selected outcome indicators will be presented along with their weightings in the synthetic indicators. 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 2 June 2008. Once aggregated, these indicators make it possible to construct the efficiency frontier, which will then be set out and commented on.

3.1. **Health care**

3.1.1. **Outcome indicators**

The main objectives in the field of health care should be an improvement in the general standard of public health compared with a situation where there is no intervention by the public authorities. Eight indicators have been selected in order to measure performance in the health care sector as well as possible.

As far as an improvement in the standard of public health is concerned, three indicators based on hard data have been considered, together accounting for 50 p.c. of the synthetic health care performance indicator. Life expectancy and healthy life expectancy together make up 30 p.c. of the synthetic indicator. This weighting, which may seem rather small, is largely a result of their partial analogy. Taken together, they nevertheless make up one of the most important indicators. Infant mortality, which counts for 20 p.c., is also a key element.
Table 1 - Outcome indicators for health care

<table>
<thead>
<tr>
<th>Type</th>
<th>Source</th>
<th>Unit, question asked or reference group</th>
<th>Weight in synthetic indicator (in percentages)</th>
<th>Belgium’s position (out of 18 countries)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Life expectancy</td>
<td>Hard data</td>
<td>WHO</td>
<td>Total population at birth</td>
<td>10</td>
</tr>
<tr>
<td>Healthy life expectancy</td>
<td>Hard data</td>
<td>OECD/WHO</td>
<td>Total population at birth</td>
<td>20</td>
</tr>
<tr>
<td>Infant mortality rate</td>
<td>Hard data</td>
<td>OECD</td>
<td>Deaths per 1,000 live births (under the age of 5 years)</td>
<td>20</td>
</tr>
<tr>
<td>Average waiting time (non-urgent treatment)</td>
<td>Assessment on the basis of OECD data</td>
<td>SCP</td>
<td>4 groups of countries with the same results, on the basis of national data</td>
<td>20</td>
</tr>
<tr>
<td>Perceived health status</td>
<td>Large survey</td>
<td>OECD on the basis of national surveys</td>
<td>Percentage of the population, of all ages, reporting at least good health</td>
<td>7.5</td>
</tr>
<tr>
<td>Confidence in health care</td>
<td>Survey</td>
<td>European values survey</td>
<td>Share of the population questioned expressing confidence (&quot;a great deal&quot; and &quot;quite a lot&quot;)</td>
<td>7.5</td>
</tr>
<tr>
<td>Public's satisfaction with the health care system</td>
<td>Survey</td>
<td>OECD</td>
<td>Share of the population questioned expressing satisfaction (&quot;runs well&quot; and &quot;minor changes needed&quot;)</td>
<td>7.5</td>
</tr>
<tr>
<td>Health infrastructure</td>
<td>Survey</td>
<td>IMD</td>
<td>&quot;Health infrastructure meets the needs of society?&quot;</td>
<td>7.5</td>
</tr>
</tbody>
</table>

The indicator concerning the average waiting time for non-urgent treatment has been given a weight of 20 p.c., because it does better than the others at isolating public authorities’ action from the influence of other factors. It would have had an even higher weighting if it had been based on more robust data.

---

16 Japan is not covered by this indicator.
17 For this indicator, for which data are only available for a few years, the last available year (2003, 2004 or 2005) was considered, except for Austria (1999). Greece is not covered by this indicator, however.
18 This indicator does not cover the United States or Japan.
19 This indicator does not cover the United States, Japan or Poland.
Lastly, four indicators relate to survey findings drawn from fairly wide samples. The indicators for individual perceived health status, confidence in the health care system, public satisfaction with this system and assessment of existing health infrastructure together get a weighting of 30 p.c. in the synthetic indicator, or 7.5 p.c. each.

Among the available indicators, some pertaining to health policy have not been selected. An important one is an indicator based on the residual effect of the health care system on life expectancy, once external factors have been taken into consideration (see box 1). While giving similar results to the "classical" life expectancy indicator, this indicator has not been included here, as explained in the box below. Alongside it, numerous indicators linked to the prevention policy that are almost exclusively related to direct action by the public authorities were not considered as they measure intermediate objectives and not outcomes. Furthermore, while important in some countries, some indicators are less relevant when it comes to the sample countries used here. Lastly, for want of recent data, other indicators have also been rejected.

Box 1 - Alternative measure of life expectancy at birth, corrected for environmental factors

One argument very detrimental to the simplified presentation made here is that the external factors which affect (healthy) life expectancy at birth, like food habits, pollution or welfare, are not accounted for in the efficiency measurement of the health care system. To do so, one should consider the synthetic outcome indicator and use it in a classical econometric regression, including expenditure but also those external factors. This has not been done here given the focus of this working paper on the outcome measurement. However, an alternative exercise within the FDH framework can be briefly presented here.

The OECD is currently working on spending efficiency in the health care sector. A very interesting data analysis led to the isolation of a country-specific effect on life expectancy at birth from other effects like health care expenditure, education, tobacco and alcohol consumption, diet, as well as the impact of pollution and GDP. The OECD uses the life expectancy at birth as a proxy for the full health care outcome. The residual country-specific factor is then considered as an indicator of the efficiency of the health care system. Given that the aim of the current working paper is to calculate efficiency by relating expenditure to outcome, the effect of health care expenditure on life

---

20 Number of mammograms carried out, immunisation rates against certain diseases (diphtheria, tetanus, pertussis, measles, influenza, etc.) or the extent of cervical cancer screening, for example.

21 Malaria prevalence, for example.

22 Infant health or the work absenteeism rate, for example.

23 Joumard, I. et al. (2008)
expectancy is also included in the country-specific effect here, contrary to the approach followed in Joumard et al. (2008).

It can be shown that the impact of using one or the other standardised results of the classical life expectancy indicator or the OECD indicator described above is broadly limited. Poland and Greece are the two countries benefiting the most from a correction for the external factors, but Belgium and Ireland would also be slightly better off after a correction. The most negatively affected by the correction is Sweden, as the external factors contribute positively to life expectancy in that country. Also negatively affected by the correction are Germany, the United Kingdom, Austria and the United States.

![Graph: "Classical" life expectancy vs. OECD country specific effect](image)

Sources: OECD, NBB.

Whether the sole indicator is the "classical" life expectancy or the OECD country-specific indicator does not change much at the efficiency frontier. Poland, Ireland and Sweden are on the efficiency frontier in both cases. However, Greece joins the group of efficient countries on the one hand and Finland falls out of it. Belgium is close to the efficiency frontier, but is relatively better off when considering the OECD country-specific effect, benefiting from corrections for the level of education but also for drinking and smoking habits. The USA is the worst-performing country in both cases (not shown on the following chart).
Even if conceptually better than the "classical" life expectancy indicator, the OECD country-specific effect indicator has nevertheless not been taken into consideration in what follows, for several reasons. First, it is a better indicator of life expectancy but does not cover a full set of countries, as only 13 countries in our sample are covered, and not Japan, Spain, Italy, Luxembourg or Portugal. Second, this OECD indicator could replace the "classical" life expectancy one, but the latter weights only 10 p.c. of the synthetic outcome indicator for the health care system. At least for Belgium, the difference would then be only marginal. It could have been better if the correction had also been applied to healthy life expectancy, a far more relevant indicator, but this was not done by the OECD. Finally, other external factors should also be corrected for in order to improve this OECD indicator, like accidental deaths and suicide.

3.1.2. Expenditure on health care

As regards expenditure on health care measurement, it is a question of determining what should be taken into consideration. This paper relies on the definition set by the OECD, where much of the data comes from: "the total expenditure on health is defined as the sum of expenditure on activities that (...) has the goals of: promoting health and preventing disease, curing illness and reducing premature mortality, caring for persons affected by chronic illness who require nursing care, caring for persons with health-related impairments, disability, and handicaps who require nursing care, assisting patients to die with dignity, providing and administering public health, providing and administering health programmes, health insurance and other funding arrangements. With this boundary, general public safety measures such as technical standards monitoring and road safety..."
are not considered as part of expenditure on health. Activities such as food and hygiene control and health research and development are considered health-related, but are not included in total health expenditure". Moreover, it appears useful to consider private expenditure too, because funding methods differ considerably from one country to the other. Both current expenditure and investment costs have been considered, so as to avoid introducing any bias into the technological choice made.

Expenditure on health care thus defined is expressed in GDP terms. An expression per capita could have been advocated, since the population is a determining factor in total expenditure on health. However, the age structure of the population determines spending more than the total population and the impact of a country’s demography on health care expenditure cannot be easily neutralised. In addition, an expression per capita would need to be adjusted for purchasing power parity, the robustness of which is regularly called into question.

3.1.3. Efficiency in the field of health care

By relating costs to the synthetic outcome indicator for health care - an aggregation of the partial indicators -, the efficiency frontier can be established, as shown in chart 2.

---

In 2005, the share of private expenditure in total expenditure on health ranged from 9.4 p.c. in Luxembourg (2004) to 54.9 p.c. in the United States, with Belgium in between with 27.7 p.c.
This reveals that three countries in the sample, namely Poland, Finland and Japan are relatively efficient. Ireland is not efficient, as its expenditure on health care is equal to those in Finland but its outcome is lower. Since three indicators are missing for Japan, a synthetic performance indicator limited to the five partial indicators that are available for all countries was rebuilt. As Japan was then still on the efficiency frontier and preserved its relative position, it was not eliminated from the sample\textsuperscript{25}. However, there is no doubt that this country stands out sharply from the others, especially in terms of lifestyle, which can have a determining influence on health without that implying any greater general government sector efficiency in this field. If Japan is left out of the equation for this reason, three other countries can be considered to be relatively efficient: Spain, Luxembourg and Sweden. Also noteworthy is the exceptionally inefficient situation of the United States, where total costs are particularly high without those bringing good results in terms of outcome. Here, the difference in lifestyle such as eating habits is also a potential source of divergence from the other countries in the sample.

Although Belgium is not on the efficiency frontier, it is nevertheless not far away from it. Apart from Japan, three countries are more efficient, namely Sweden, Luxembourg and Austria. These latter three countries can serve as a benchmark for Belgium: Sweden, for example, reaches a higher outcome with costs almost 12 p.c. lower. Luxembourg also reaches a higher level of outcome than Belgium, but its costs are almost 20 p.c. lower. In addition, in this last case, the gap cannot be attributed to a difference in lifestyle. All in all, Belgium seems to have opted for a more costly policy than the European average but one that offers a better outcome.

**Box 2 - Influence of the selected weightings**

This box shows that the choices involved when composing synthetic indicators, in other words the choice of weighting allocated to the partial indicators, can greatly influence the relative efficiency of the different countries. In order to do this, two complementary cases were chosen. In the first case, only the partial indicators based on hard data make up the synthetic outcome indicator\textsuperscript{26}. In the second case, the other five indicators make up the synthetic outcome indicator. The relative efficiency of Belgium in those two cases appears to be very different.

In the first case, the predominance in the synthetic indicator of factors that are most unfavourable to Belgium would put it behind as many as 8 of the 17 other sample countries and behind the average of the EU-15.

\textsuperscript{25} Note that, in that case, Sweden would also be on the efficiency frontier.

\textsuperscript{26} The indicators on (healthy) life expectancy and infant mortality are considered as hard data.
In the second case, on the contrary, Belgium is very close to the efficiency frontier. The limitation of partial indicators to the survey data that are more favourable to Belgium actually leads to a better synthetic outcome.

In both cases, however, some indicators are de facto omitted and, consequently, all the assumed outcomes no longer covered. Nevertheless, the importance of the choice of indicator weightings or the risks posed by excessive subjectivity on the part of those setting the weightings has been emphasised in this way.
3.2. Education

3.2.1. 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, science and problem-solving. 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 the Flemish and French Communities, that are responsible for education in Belgium, which makes it possible to distinguish between the performance of the two main (language) communities in Belgium27.

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 importance28.

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 proxy of the contribution to labour force qualifications, and therefore as an outcome of the educational

27 The results for the German Community are not included here for the sake of simplicity.
28 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 without losing much of the information. This way, the indicator will be positively influenced by a small standard deviation.
Moreover, they are exhaustive hard data indicators. These two indicators together count for 25 p.c. in the synthetic indicator.

### Table 2 - Outcome indicators for education

<table>
<thead>
<tr>
<th>Type</th>
<th>Source</th>
<th>Unit, question asked or reference group</th>
<th>Weight in synthetic indicator (in percentages)</th>
<th>Belgium’s position (out of 17 countries)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading skills</td>
<td>Student assessment OECD (PISA)</td>
<td>15 year-old pupils</td>
<td>12.5</td>
<td>7 (FL: 3; FR: 15)</td>
</tr>
<tr>
<td>Mathematics performances</td>
<td>Student assessment OECD (PISA)</td>
<td>15 year-old pupils</td>
<td>12.5</td>
<td>4 (FL: 2; FR: 13)</td>
</tr>
<tr>
<td>Scientific literacy skills</td>
<td>Student assessment OECD (PISA)</td>
<td>15 year-old pupils</td>
<td>12.5</td>
<td>8 (FL: 3; FR: 15)</td>
</tr>
<tr>
<td>Problem-solving skills</td>
<td>Student assessment OECD (PISA)</td>
<td>15 year-old pupils</td>
<td>12.5</td>
<td>4 (out of 16)</td>
</tr>
<tr>
<td>Educational attainment: secondary education</td>
<td>Hard data OECD</td>
<td>Percentage of 25-34-year-olds with an upper secondary education</td>
<td>12.5</td>
<td>7</td>
</tr>
<tr>
<td>Educational attainment: tertiary education</td>
<td>Hard data OECD</td>
<td>Percentage of 25-34-year-olds with a tertiary education</td>
<td>12.5</td>
<td>2</td>
</tr>
<tr>
<td>Language skills</td>
<td>Survey IMD</td>
<td>“Language skills meet the needs of the society?”</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>Confidence in the education system</td>
<td>Survey European Values Study</td>
<td>Share of the population surveyed expressing confidence (“a great deal” + “quite a lot”)</td>
<td>5</td>
<td>5 (out of 15)</td>
</tr>
<tr>
<td>Quality of educational system</td>
<td>Survey WEF+IMD</td>
<td>Average of 3 questions</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Availability of skilled labour</td>
<td>Survey IMD</td>
<td>“Skilled labour is readily available?”</td>
<td>5</td>
<td>10</td>
</tr>
</tbody>
</table>

29 Luxembourg is not considered in this table. When giving the position of the Flemish and the French Communities, Belgium is not considered as a whole, but just the two Communities, i.e. 18 “countries”.
30 This indicator does not cover the United Kingdom.
31 The United States and Japan are not covered by this indicator.
32 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.

3.2.2. 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.

In a partial efficiency frontier, limited to attainment among 15 year-old pupils in certain subjects, which enables a distinction to be made between the performance of the French and the Flemish Communities, the cost of education should be limited to cumulated education expenditure up to this age. Based on data supplied by OECD member countries\(^{33}\), the Organisation’s Secretariat has worked out cumulative expenditure on educational institutions per pupil between the age of 6 and 15 by adding up actual total spending each year (from 1993 to 2002) for a student (based on full-time equivalents) aged 15 in 2002, calculated according to the corresponding level of education for

\(^{33}\) OECD (2005).
each country and converted to 2002 US dollars using PPPs for household final consumption\textsuperscript{34}. This reflects accurately expenditure on educating a specific pupils cohort, but does not fully reflect the changes in spending patterns since 1993.

3.2.3. Efficiency in education
Two efficiency frontiers can be established for education. The first covers efficiency of the whole education system and the second frontier is limited to education up to the age of 15 years.

Global efficiency frontier: Belgium as a whole
Aggregating the ten indicators selected and the total of the costs taken into consideration puts four countries on the efficiency frontier, namely Poland, Ireland, the Netherlands and Finland. These countries’ efficiency levels in the field of education are particularly high in relation to the resources employed in this sector. Apart from Finland and the Netherlands, Japan also leads Belgium in the efficiency ranking. It should also be pointed out that Austria and the United States are inefficient compared to Belgium. Lastly, the Belgian option is to spend more money than the European average in order to attain a better performance, but it cannot be said that one is more efficient than the other.

In order to join the most efficient group of countries, Belgium could try to be as efficient as the Netherlands or Finland. These spend almost 9 p.c. less on education while performing better than

\textsuperscript{34} Sutherland et al. (2007).
Belgium in terms of outcome. Such a conclusion, however, is strongly limited by the weaknesses of the framework, as explained above.

Partial efficiency frontier: findings for the French and Flemish Communities

Since organisation of education in Belgium is the responsibility of the communities, it is also interesting to be able to judge their efficiency levels separately. However, since the indicators available at the level of the communities are limited to the PISA test scores for reading, mathematics and science, the results turn out to be partial.

The PISA survey findings have thus been compared with the financial resources deployed by the different countries to obtain these results. They show Poland, Ireland, the Netherlands and Finland on the efficiency frontier. Only Finland, Ireland, the Netherlands and Japan are ahead of Belgium taken as a whole, while Belgium does better than France, Sweden, Denmark, Austria, Luxembourg and the United States in the efficiency stakes. In Ireland, expenditure on primary and lower secondary education is nearly 45 p.c. lower than in Belgium as a whole, for example, while this country has a better outcome.

Also to be noted is the extreme position of countries with high GDP - Luxembourg - and low GDP - Poland. This is partially due to the input measurement, and would have been different if relating expenditure to GDP.
Amongst the possible reasons for this huge difference between the Flemish and the French Communities, some have been listed by the Belgian High Council of Employment\textsuperscript{35} as hypotheses. It seems that backward pupils are far more numerous in the French Community, with a direct impact on the level of education at the age of 15 and indirect impact on the motivation of students. The various types of education system are also considered differently: while technical and professional education have a higher status in the Flemish Community, this is far less the case in the French Community, where it is more of a relegation. Next, the degree of autonomy of schools seems larger in the Flemish Community, where there are more free schools. The socio-economic environment, somewhat weaker in the French Community, probably contributes to a lower level of education here, too. Finally, factors like the importance of “renovated” education, which is more developed in the French Community, could be detrimental to its efficiency.

According to this synthetic indicator, the main observation is that the Flemish Community’s outcome is clearly better than the French Community’s, at scarcely any higher cost per pupil. The French Community therefore has wide potential margins for improvement here. A reduction in expenditure is probably not appropriate, while a very strong improvement in outcome, up to the level reached by the Netherlands for example, should be a top priority for the French Community. In the case of the Flemish Community, efficiency gains are also still possible, either by limiting costs or by improving outcome.

3.3. Public order and safety

Public order and safety, here, is taken to mean joint action taken by the forces of law and order (police and the law courts). There are fewer indicators available and they are less robust than for the preceding fields of activity. The results should therefore be interpreted with even greater caution.

3.3.1. Outcome indicators

The performance indicators are compiled from survey data only. Two indicators nevertheless come from larger-sample surveys than the WEF or IMD surveys, comprising around 2,000 people per country. The first is an indicator of the risk of falling victim to a crime such as theft, burglary, intimidation or rape. This indicator, which excludes motoring offences, drug trafficking or fraud, counts for 15 p.c. in the synthetic indicator. The second one is an indicator of how those surveyed are satisfied with the law-and-order mission of the police in their immediate proximoty; this indicator accounts for 10 p.c. of the synthetic indicator. It is close to a police reliability indicator, derived from a smaller-sample survey which also counts for 10 p.c. The business cost of crime and violence

\textsuperscript{35} Conseil Supérieur de l’emploi (2003).
indicator covers all action by the forces of law and order. Its 15 p.c. weighting in the synthetic indicator could have been higher if it had covered society as a whole rather than being restricted to its business segment. An indicator based on survey results whether personal security and private property are adequately protected has been given a weight of 15 p.c. Lastly, one single indicator refers specifically to justice, to see whether it is fairly administered, which explains its 35 p.c. weight in the synthetic indicator.

Table 3 - Indicators for public order and safety

<table>
<thead>
<tr>
<th>Type</th>
<th>Source</th>
<th>Unit, question asked or reference group</th>
<th>Weight in synthetic indicator (in percentages)</th>
<th>Belgium's position (out of 15 countries)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Burden of crime</td>
<td>Large survey</td>
<td>Risk of falling victim of a crime during the last 12 months</td>
<td>15</td>
<td>11</td>
</tr>
<tr>
<td>Police satisfaction</td>
<td>Large survey</td>
<td>Percentage satisfied with police controlling crime in local area</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>Business cost of crime and violence</td>
<td>Survey</td>
<td>WEF</td>
<td>The incidence of common crime and violence (e.g. street muggings, firm being looted) imposes significant costs on businesses</td>
<td>15</td>
</tr>
<tr>
<td>Police reliability</td>
<td>Survey</td>
<td>Police services can be relied upon to protect businesses from criminals</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>Personal security and private property</td>
<td>Survey</td>
<td>IMD</td>
<td>Personal security and private property are adequately protected</td>
<td>15</td>
</tr>
<tr>
<td>Justice</td>
<td>Survey</td>
<td>Justice is fairly administered</td>
<td>35</td>
<td>11</td>
</tr>
</tbody>
</table>

Many figures from the "European Sourcebook of Crime and Criminal Justice Statistics"\(^{38}\), such as the number of offences per 100,000 inhabitants or the prosecution statistics and conviction numbers, have not been used as indicators. First, the data on the subject are still far from being harmonised, so the statistics are often incomparable. Secondly, most of these statistics are potentially biased in favour of the less efficient countries. It is quite possible that the proportion – and therefore the number - of offences reported to the police by victims increases with the efficiency of the police force in tracking down criminals. To get round this bias, a figure indicating the proportion of cases resolved – and even crimes punished – would be more convincing, but this

\(^{36}\) The sample is limited to the oldest 15 EU countries.

\(^{37}\) About 2,000 people surveyed in each country, i.e. much more than WEF or IMD samples (see section 2.4).

\(^{38}\) EUICS (2005).
does not exist. Other interesting statistics could be the proportion of crime solved or the time required before the start of a trial or the final judgement. Unfortunately, such statistics are not widely available on a harmonised basis.

3.3.2. **Expenditure on public order and safety**

Figures for spending on public order and safety come from COFOG data reported by EU Member States. The allocation of each public expenditure component to one function only is again a simplification of reality. As there is an element of uncertainty, for example when one expenditure has several goals, again, these figures must be treated with care.

Expenditure on public order and safety is expressed as a percentage of GDP. Spending on this general government function accounts only for about 1.8 p.c. of GDP on average in the EU-15, considerably less than the 9 percent-plus spent on health care and the 5 p.c. spent on education. Poland, Japan and the United States are taken out of the sample of countries analysed here because costs in the field of public order and safety are not available in the first of those countries and not available according to the same source for the other two.

3.3.3. **Efficiency in public order and safety**

The findings indicate that Denmark and Austria are relatively efficient as far as public order and safety are concerned. Belgium is in a unfavourable position, as it is lagging behind no less than seven other countries. A country like Denmark pays out 36 p.c. less than Belgium, but still has a
better outcome. The best performing country in terms of outcome, Austria, spends around 10 p.c. less than Belgium.

Compared with the EU-15 average, Belgium seems to have opted for lower expenditure together with lower outcome.

4. CONCLUSION

The aim of this working paper was to try and determine to what extent the Belgian general government sector efficiently fulfils its role in the areas of health care, education and public order and safety. Each of these general government sector services has been analysed separately, as there is little point in treating them as a whole.

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’ general government sectors, 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 general government sector production 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 Belgium is relatively efficient in the field of health care when compared with the other countries considered - the EU-15, Japan, the United States and Poland. Only four European countries – Spain, Sweden, Luxembourg and Austria - and Japan are more efficient than Belgium. The latter has opted for relatively high expenditure in order to gain an equally high outcome. Limiting the synthetic outcome indicator to hard data only - (healthy) life expectancy and infant mortality - would give a less favourable picture for Belgium. There is some margin for improvement in the field of health care - which is somewhat more expensive than on average in the EU-15 - but not as much as in the other general government sector services analysed. In Belgium, appreciation of the medical infrastructure, confidence in and public's satisfaction with the health care system are amongst the highest, and the average waiting time for a non-urgent treatment is the shortest.

As a whole, the Belgian education system is more expensive but also leading to better results than the European average. However, a clear distinction appears between the efficiency of the education system of the French Community and that of the Flemish Community. An analysis based on a limited set of indicators reveals that the French-speaking education sector is very inefficient. French Community pupils' performances in the PISA indicators in reading, mathematics and sciences are weak. The French Community therefore has wide potential margins for improvement and this needs to be a key policy objective. The Flemish Community's efficiency in education is markedly better, without however reaching the efficiency frontier.

As far as public order and safety are concerned, major improvements could and should be made, either to improve service or cut costs.

Although this working paper provides some indication of the efficiency of Belgian general government in an international perspective, there is still a lot of research work to be done in this area, for instance to widen the range of indicators covering performance, extend research to other general government sector services - such as general public services or even certain public enterprises – and map efficiency developments over time.
REFERENCES


AMECO Database, European Commission.


European Values Study Group and World Values Survey Association (2005), European and world values surveys integrated data file, 1999-2002, release I.


IMD World Competitiveness Yearbook (2006), Lausanne.

IMD World Competitiveness Yearbook (2007), Lausanne.

IMD World Competitiveness Yearbook (2008), Lausanne.


OECD Health Data 2006.


## Appendix: Database

### Health care

<table>
<thead>
<tr>
<th>Outcome indicators</th>
<th>Expenditure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Life expectancy at birth, in years</td>
<td>Health infrastructure meets the needs of the society?</td>
</tr>
<tr>
<td>Healthy life expectancy at birth, in years</td>
<td>Expenditure on prevention and public health, public and private (percentages of GDP)</td>
</tr>
<tr>
<td>Infant mortality rate, deaths per 1,000 live births (under the age of 5 years)</td>
<td>GDP (millions of euro)</td>
</tr>
<tr>
<td>Average waiting time (non-urgent treatment)</td>
<td>Total population (thousands)</td>
</tr>
<tr>
<td>Perceived health status (percentage of population reporting good health)</td>
<td></td>
</tr>
<tr>
<td>Confidence in health care (percentage of answers &quot;a great deal&quot; + &quot;quite a lot&quot;)</td>
<td></td>
</tr>
<tr>
<td>Public's satisfaction with the health care system (percentage of answers &quot;runs well&quot; + &quot;minor changes needed&quot;)</td>
<td></td>
</tr>
<tr>
<td>Expenditure on prevention and public health, public and private (percentages of GDP)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>DE</td>
<td>79.0</td>
<td>71.8</td>
<td>4.0</td>
<td>72.6</td>
<td>53.0</td>
<td>47.1</td>
<td>8.1</td>
</tr>
<tr>
<td>AT</td>
<td>79.5</td>
<td>71.4</td>
<td>4.0</td>
<td>73.5</td>
<td>86.3</td>
<td>67.2</td>
<td>8.9</td>
</tr>
<tr>
<td>BE</td>
<td>78.7</td>
<td>71.1</td>
<td>4.0</td>
<td>76.9</td>
<td>82.6</td>
<td>65.1</td>
<td>8.7</td>
</tr>
<tr>
<td>DK</td>
<td>77.9</td>
<td>69.8</td>
<td>2.0</td>
<td>79.4</td>
<td>69.6</td>
<td>51.6</td>
<td>7.5</td>
</tr>
<tr>
<td>ES</td>
<td>80.7</td>
<td>72.6</td>
<td>4.1</td>
<td>68.3</td>
<td>65.6</td>
<td>46.5</td>
<td>6.7</td>
</tr>
<tr>
<td>FI</td>
<td>78.9</td>
<td>71.1</td>
<td>1.0</td>
<td>66.4</td>
<td>84.4</td>
<td>72.6</td>
<td>7.1</td>
</tr>
<tr>
<td>FR</td>
<td>80.3</td>
<td>72.0</td>
<td>4.0</td>
<td>79.5</td>
<td>77.4</td>
<td>63.9</td>
<td>7.8</td>
</tr>
<tr>
<td>GR</td>
<td>79.3</td>
<td>71.0</td>
<td>3.0</td>
<td>n.a.</td>
<td>26.8</td>
<td>18.8</td>
<td>4.4</td>
</tr>
<tr>
<td>IE</td>
<td>79.5</td>
<td>69.8</td>
<td>2.5</td>
<td>82.9</td>
<td>57.5</td>
<td>20.4</td>
<td>3.2</td>
</tr>
<tr>
<td>IT</td>
<td>80.4</td>
<td>72.7</td>
<td>2.5</td>
<td>58.9</td>
<td>36.7</td>
<td>30.9</td>
<td>5.8</td>
</tr>
<tr>
<td>LU</td>
<td>79.3</td>
<td>71.5</td>
<td>2.6</td>
<td>74.0</td>
<td>77.9</td>
<td>67.7</td>
<td>8.0</td>
</tr>
<tr>
<td>NL</td>
<td>79.4</td>
<td>71.9</td>
<td>2.0</td>
<td>76.6</td>
<td>75.1</td>
<td>45.6</td>
<td>7.3</td>
</tr>
<tr>
<td>PT</td>
<td>78.2</td>
<td>69.2</td>
<td>3.5</td>
<td>39.0</td>
<td>44.0</td>
<td>14.3</td>
<td>4.6</td>
</tr>
<tr>
<td>UK</td>
<td>79.0</td>
<td>70.6</td>
<td>5.1</td>
<td>73.9</td>
<td>58.7</td>
<td>31.2</td>
<td>4.9</td>
</tr>
<tr>
<td>SE</td>
<td>80.6</td>
<td>73.3</td>
<td>2.4</td>
<td>74.4</td>
<td>76.3</td>
<td>47.7</td>
<td>7.4</td>
</tr>
<tr>
<td>US</td>
<td>77.8</td>
<td>69.3</td>
<td>3.0</td>
<td>88.7</td>
<td>n.a.</td>
<td>4.9</td>
<td>15.3</td>
</tr>
<tr>
<td>JP</td>
<td>82.1</td>
<td>75.0</td>
<td>2.8</td>
<td>38.7</td>
<td>n.a.</td>
<td>6.4</td>
<td>8.0</td>
</tr>
<tr>
<td>PL</td>
<td>75.1</td>
<td>65.8</td>
<td>3.0</td>
<td>54.5</td>
<td>56.6</td>
<td>n.a.</td>
<td>1.9</td>
</tr>
<tr>
<td>EU-15</td>
<td>79.6</td>
<td>71.7</td>
<td>2.8</td>
<td>72.6</td>
<td>59.4</td>
<td>44.0</td>
<td>6.75</td>
</tr>
</tbody>
</table>

1 Average weighted by the population for outcome indicators, by the GDP for expenditure.
2 2004.
3 1=very problematic, 2=moderately problematic, 3=slightly problematic, 4=no problem.
4 1999.
| 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) | 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) | Confidence in the educational system ("a great deal" and "quite a lot") | Quality of the educational system ("the educational system in your country (1= does not meet the needs of a competitive economy; 7=meets the needs of a competitive economy") | Quality of the educational system ("the University education (does not) meet(s) the needs of a competitive economy") | 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) | Cumulative expenditure per student between 6 and 15 years (US dollars converted using PPPs) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| BE Flemish Community | 2468 | 2584 | 2533 | 2604 | 7.6 | 78 | 39 | 78 | 5.7 | 6.2 | 6.9 | 5.3 | 7933 | 55273 |
| BE French Community | 2570 | 2691 | 2624 | n.a. | n.a. | n.a. | n.a. | n.a. | n.a. | n.a. | n.a. | n.a. | 57129 |
| DK | 2461 | 2561 | 2478 | 2628 | 8.5 | 86 | 35 | 75 | 5.8 | 6.7 | 7.1 | 5.6 | 9261 | 66323 |
| DE | 2452 | 2518 | 2567 | 2552 | 6.5 | 85 | 22 | 73 | 4.9 | 4.8 | 5.9 | 6.5 | 7129 | 47991 |
| GR | 2273 | 2290 | 2359 | 2238 | 7.3 | 72 | 24 | 37 | 3.3 | 3.4 | 3.1 | 5.2 | 4136 | 28982 |
| ES | 2287 | 2397 | 2438 | 2401 | 2.9 | 60 | 38 | 68 | 3.8 | 3.5 | 4.1 | 4.9 | 5914 | 42356 |
| FR | 2409 | 2469 | 2462 | 2584 | 3.9 | 80 | 37 | 68 | 4.8 | 4.9 | 4.9 | 6.2 | 7467 | 56361 |
| IE | 2575 | 2505 | 2538 | 2487 | 5.4 | 78 | 37 | 86 | 5.6 | 7.1 | 7.7 | 6.2 | 5711 | 30456 |
| IT | 2320 | 2309 | 2375 | 2331 | 3.7 | 60 | 12 | 53 | 3.4 | 4.0 | 4.1 | 5.4 | 7708 | 54723 |
| LU | 2378 | 2444 | 2423 | 2463 | 8.7 | 68 | 19 | 68 | 4.2 | 5.2 | 4.8 | 5.0 | n.a. | 100572 |
| NL | 2512 | 2647 | 2614 | 2597 | 7.8 | 76 | 22 | 73 | 5.2 | 6.1 | 6.4 | 5.7 | 7241 | 48711 |
| AT | 2429 | 2515 | 2540 | 2526 | 7.1 | 85 | 15 | 86 | 5.2 | 6.8 | 7.4 | 7.4 | 8943 | 76029 |
| PT | 2344 | 2327 | 2369 | 2338 | 6.6 | 37 | 16 | 60 | 3.5 | 3.3 | 4.4 | 4.8 | 6080 | 42894 |
| FI | 2728 | 2737 | 2810 | 2733 | 6.7 | 89 | 40 | 89 | 6.0 | 7.2 | 6.8 | 4.8 | 7304 | 54456 |
| SE | 2521 | 2512 | 2512 | 2538 | 7.9 | 91 | 40 | 68 | 5.2 | 5.1 | 6.1 | 6.2 | 8520 | 57652 |
| UK | 2462 | 2481 | 2569 | n.a. | 3.3 | 71 | 33 | 66 | 4.6 | 4.2 | 5.4 | 5.0 | 6691 | 48355 |
| US | 2462 | 2376 | 2448 | 2382 | 4.4 | 87 | 39 | n.a. | 5.1 | 5.3 | 7.2 | 6.7 | 11152 | 71626 |
| JP | 2471 | 2610 | 2640 | 2716 | 3.5 | 94 | 52 | n.a. | 4.7 | 4.8 | 4.4 | 6.4 | 7438 | 50763 |
| PL | 2526 | 2478 | 2490 | 2433 | 4.0 | 57 | 20 | 81 | 4.0 | 3.7 | 4.6 | 4.1 | 2962 | 16583 |

2. 2002.
3. These three indicators have been aggregated.
4. Weighted average of Flemish Community (60 %) and French Community (40 %)
Public order and safety

<table>
<thead>
<tr>
<th>Outcome indicators</th>
<th>Expenditure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Burden of crime (being or not victim of a crime during the last 12 months, percentage)</td>
<td>Expenditure on public order and safety (percentages of GDP) and GDP (millions of euro)</td>
</tr>
<tr>
<td>Business cost of crime and violence (&quot;the incidence of common crime and violence (e.g., street muggings, firms being looted) (1=imposes significant costs on businesses; 7=does not imposes significant costs on businesses)?&quot;)</td>
<td>2005 2007 2005 2007 2007 2007 2005 2005</td>
</tr>
<tr>
<td>Police satisfaction (percentage satisfied with police controlling crime in local area)</td>
<td></td>
</tr>
<tr>
<td>Police reliability (&quot;police services (1=cannot be relied upon to protect businesses from criminals; 7=can be relied upon to protect businesses from criminals)?&quot;)</td>
<td></td>
</tr>
<tr>
<td>Personal security and private property (&quot;personal security and private property are (not) adequately protected&quot;)</td>
<td></td>
</tr>
<tr>
<td>&quot;Justice is (not) fairly administered?&quot;</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>BE</td>
<td>17.8</td>
<td>5.5</td>
<td>71</td>
<td>5.4</td>
<td>7.2</td>
<td>6.1</td>
<td>1.5</td>
<td>302.0</td>
</tr>
<tr>
<td>DK</td>
<td>19.3</td>
<td>6.5</td>
<td>82</td>
<td>6.6</td>
<td>9.1</td>
<td>9.0</td>
<td>1.0</td>
<td>207.8</td>
</tr>
<tr>
<td>DE</td>
<td>13.1</td>
<td>6.4</td>
<td>74</td>
<td>6.6</td>
<td>8.4</td>
<td>8.2</td>
<td>1.6</td>
<td>2244.6</td>
</tr>
<tr>
<td>GR</td>
<td>12.3</td>
<td>5.9</td>
<td>57</td>
<td>4.7</td>
<td>7.0</td>
<td>5.9</td>
<td>1.2</td>
<td>198.6</td>
</tr>
<tr>
<td>ES</td>
<td>9.0</td>
<td>5</td>
<td>58</td>
<td>5.6</td>
<td>5.7</td>
<td>4.6</td>
<td>1.8</td>
<td>908.5</td>
</tr>
<tr>
<td>FR</td>
<td>12.0</td>
<td>5.1</td>
<td>60</td>
<td>5.8</td>
<td>7.6</td>
<td>6.5</td>
<td>1.3</td>
<td>1726.1</td>
</tr>
<tr>
<td>IE</td>
<td>22.1</td>
<td>5.2</td>
<td>78</td>
<td>5.6</td>
<td>7.8</td>
<td>7.7</td>
<td>1.4</td>
<td>161.5</td>
</tr>
<tr>
<td>IT</td>
<td>12.6</td>
<td>4.6</td>
<td>65</td>
<td>4.8</td>
<td>5.5</td>
<td>3.6</td>
<td>1.8</td>
<td>1428.4</td>
</tr>
<tr>
<td>LU</td>
<td>12.7</td>
<td>5.9</td>
<td>62</td>
<td>5.8</td>
<td>6.7</td>
<td>6.4</td>
<td>1.8</td>
<td>149.1</td>
</tr>
<tr>
<td>NL</td>
<td>19.8</td>
<td>5</td>
<td>70</td>
<td>5.9</td>
<td>6.9</td>
<td>9.4</td>
<td>1.8</td>
<td>509.0</td>
</tr>
<tr>
<td>AT</td>
<td>12.2</td>
<td>6.4</td>
<td>81</td>
<td>6.2</td>
<td>9.4</td>
<td>9.0</td>
<td>1.4</td>
<td>245.3</td>
</tr>
<tr>
<td>PT</td>
<td>10.4</td>
<td>5.9</td>
<td>67</td>
<td>5.4</td>
<td>7.5</td>
<td>7.2</td>
<td>1.3</td>
<td>149.1</td>
</tr>
<tr>
<td>FI</td>
<td>12.7</td>
<td>6.7</td>
<td>90</td>
<td>6.7</td>
<td>8.7</td>
<td>8.2</td>
<td>1.4</td>
<td>157.3</td>
</tr>
<tr>
<td>SE</td>
<td>16.2</td>
<td>5.9</td>
<td>65</td>
<td>5.7</td>
<td>7.9</td>
<td>8.4</td>
<td>1.3</td>
<td>294.7</td>
</tr>
<tr>
<td>UK</td>
<td>21.0</td>
<td>4.7</td>
<td>75</td>
<td>5.4</td>
<td>6.4</td>
<td>6.9</td>
<td>2.6</td>
<td>1804.6</td>
</tr>
<tr>
<td>EU-15</td>
<td>14.6</td>
<td>5.4</td>
<td>68.9</td>
<td>5.8</td>
<td>7.2</td>
<td>6.6</td>
<td>1.8</td>
<td>10367.4</td>
</tr>
</tbody>
</table>

1 Average weighted by the GDP.
98. "Dynamics on monetary policy in a fair wage model of the business cycle", by D. De la Croix, G. de Walque and R. Wouters, Research series, October 2006.
102. "Fiscal sustainability indicators and policy design in the face of ageing", by G. Langenus, Research series, October 2006.
104. "Exploring the CDS-Bond Basis" by J. De Wit, Research series, November 2006.