

ISSN 1725-3187 (online) ISSN 1016-8060 (print)

## EUROPEAN ECONOMY

## Economic Papers 549 | June 2015

# Efficiency estimates of health care systems

## João Medeiros and Christoph Schwierz



**Economic Papers** are written by the staff of the Directorate-General for Economic and Financial Affairs, or by experts working in association with them. The Papers are intended to increase awareness of the technical work being done by staff and to seek comments and suggestions for further analysis. The views expressed are the author's alone and do not necessarily correspond to those of the European Commission.

Comments and enquiries should be addressed to:

European Commission Directorate-General for Economic and Financial Affairs Unit Communication B-1049 Brussels Belgium E-mail: <u>ecfin-info@ec.europa.eu</u>

#### LEGAL NOTICE

Neither the European Commission nor any person acting on its behalf may be held responsible for the use which may be made of the information contained in this publication, or for any errors which, despite careful preparation and checking, may appear.

This paper exists in English only and can be downloaded from <u>http://ec.europa.eu/economy\_finance/publications/</u>.

#### Europe Direct is a service to help you find answers to your questions about the European Union.

#### Freephone number (\*): 00 800 6 7 8 9 10 11

(\*) The information given is free, as are most calls (though some operators, phone boxes or hotels may charge you).

More information on the European Union is available on <u>http://europa.eu</u>.

Luxembourg: Publications Office of the European Union, 2015

KC-AI-15-549-EN-N (online) ISBN 978-92-79-44820-1 (online) doi:10.2765/49924 (online) KC-AI-15-549-EN-C (print) ISBN 978-92-79-44821-8 (print) doi:10.2765/82810 (print)

© European Union, 2015 Reproduction is authorised provided the source is acknowledged. European Commission Directorate-General for Economic and Financial Affairs

## **Efficiency estimates of health care systems in the EU**

João Medeiros and Christoph Schwierz

#### Abstract

There is ample evidence of widespread inefficiency in health care systems. This paper aims to estimate relative efficiency of health care systems across all EU countries. The paper uses a comprehensive battery of models with different combinations of input and output variables. Outputs are the commonly reported health outcome indicators, such as life expectancy, healthy life expectancy and amenable mortality rates. Inputs include (per capita) expenditure on health care, physical inputs and environmental variables. Results obtained in this paper are in line with previous empirical research. On average in the EU, life expectancy at birth could be increased by 2.3% or 1.8 years, when moving from current positions to the efficiency frontier. Specifically, the Czech Republic, Lithuania and Slovakia have the lowest efficiency scores in most of the models used. Hungary, Latvia, Poland and Estonia, although scoring marginally better than the previous group are also underperformers. Belgium, Cyprus, Spain, France, Italy, Sweden and the Netherlands consistently score among the top seven performers in most of the models.

JEL Classification: 112, 118, C51. Keywords: healthcare system, efficiency, data envelopment analysis, Europe.

**Contact**: Christoph Schwierz, Directorate-General for Economic and Financial Affairs, European Commission, office CHAR 12/122, B-1049 Brussels, Rue de la Loi/Wetstraat 170, 1040 Bruxelles/Brussel, tel. +32-2-296.07.52, <u>christoph.schwierz@ec.europa.eu</u>

We thank various colleagues from the European Commission for very useful discussions and comments. Any remaining errors are our exclusive responsibility. The views in the paper are those of the authors and should not be attributed to the European Commission.

## TABLE OF CONTENTS

1.	Introduction	6
2.	Overview of empirical evidence on health efficiency	7
3.	Measuring efficiency	10
4.	Estimation results	12
5.	Efficiency scores used to group/cluster countries	17
6.	Estimation of potential economic savings from improving efficiency	22
7.	Conclusions	24
ANN	EX 1: Definition of variables	29
ANN	ex 2: dataset used in the dea analysis	30
ANN	EX 3 Data envelopment analysis (DEA)	31
ANN	EX 4 Some auxiliary diagnosis elements for DEA	34
ANN	EX 5 DEA efficiency scores for all models	37
ANN	EX 6 Spearman rank order correlations for DEA models	38
ANN	EX 7 Stochastic frontier analysis (SFA)	39
ANN	EX 8 Potential economic savings of improving efficiency	42
ANN	EX 9 Estimating Malmquist indexes	44
ANN	EX 10 Country profiles -low efficiency scores	49

## LIST OF TABLES

Table 1 – DEA models used	12
Table 2 – Analysis of DEA efficiency scores and country clusters	18
Table 3 – Health outcomes for countries with low DEA efficiency scores	21
Table 4 – Kolmogorov-Smirnov tests (p-values)	36
Table 5: SFA, Error components frontier (ECF) model	40
Table 6: Efficient effects frontier (EEF) model	40
Table 7: Spearman correlations between DEA vs. SFA scores	41
Table 8: Panel regression estimates (2013-2010) of life expectancy	42
Table 9: Cumulated productivity, efficiency, and technological changes: 2003-2010	45
Table 10: Cumulated productivity, efficiency, and technological change: 2003-2010	46

## LIST OF GRAPHS

Graph 1 – Health outcomes and health care spending	10
Graph 2 – Estimated (bias corrected) production function and country observations	13
Graph 3 – Bias-corrected DEA and confidence interval (95% CI)	14
Graph 4 – Potential gains in health outcomes	15
Graph 5 – Potential improvement in health outcomes based on DEA results	16
Graph 6 – Average efficiency scores using 2 or 3 clusters	19
Graph 7 – Agglomerative hierarchical clustering	19
Graph 8 – Average efficiency scores versus life expectancy and per capita health expenditure	20
Graph 9 – Input- versus output-orientation DEA scores	22
Graph 10 – Projected increase in public health expenditure in the EU until 2060 (Ageing Report, 2012) taking into account potential efficiency gains	23
Graph 11 – DEA production possibility frontier	31
Graph 12 – Isoquant of a DEA	34
Graph 13 – Isoquant of a DEA	35
Graph 14 – The output distance function and the production possibility set	47

## 1. Introduction

In all EU countries, during most of the second half of the 20<sup>th</sup> century, health expenditure has been growing faster than national income (Ageing Report, 2012; Maisonneuve and Martins, 2013; Medeiros and Schwierz, 2013). This strong growth can be attributed to demand and supply side factors, such as population ageing and medical innovation. Besides these factors, there is ample evidence of pervasive inefficiency in the process of transforming resources into health outcomes, generating economic waste and being a contributory factor for the excessive health expenditure growth. This paper aims to estimate relative efficiency of health care systems in all EU countries.

The application of efficiency concepts to health care systems is challenging, raising both theoretical and practical problems.<sup>1</sup> As an example, health care activities such as hospital discharges, are often seen as intermediate outputs (Jacobs et al., 2006), because health care activities do not necessarily have an immediate impact on improving health outcomes, which is what patients and practitioners are looking for. In practice, the relation between inputs, (intermediate) outputs and health outcomes is complex and multifaceted. Inputs and outputs differ in often inadequately measured dimensions, such as on quantity and quality, while health outcomes are also affected by past and current lifestyle behaviour and environmental factors outside the immediate control of the health system. Also, data availability is rather limited over time and across country, restricting the use of different models, and thereby making the assessment of relative health efficiency challenging.

Despite the empirical difficulties in applying efficiency concepts to health systems, there is a considerable body of evidence at both the macro and micro levels on the pervasiveness of inefficiency in the health sector. Many findings of wasteful use of resources have been reported in the empirical literature, inter alia: i) sub-optimal setups for delivery of care; ii) inefficient provision of acute hospital care; iii) fraud and corruption in health care systems; and iv) a sub-optimal mix of preventative versus curative care (see Section 2).

Consequently, reducing inefficiencies can lead to substantial gains. OECD (2010) estimates that average life expectancy could increase by about 2 years for the OECD as a whole, if resources were used more efficiently. Conversely, holding health outcomes at current levels, while increasing efficiency to the level of the best performing countries, would free-up a considerable amount of resources.<sup>2</sup> This could help reducing the long-term growth rate of health expenditure without compromising access to (quality) care, which is a major concern for European policy makers.

This paper applies well-established methods, updating previous studies undertaken at the OECD (Joumard et al., 2010) and WHO (2000), and takes stock of results.<sup>3</sup>

In this paper, the term efficiency means technical efficiency,<sup>4</sup> implying the maximisation of outputs for a given level of inputs (or the minimisation of inputs for a given level of outputs).<sup>5</sup> Outputs are commonly reported

<sup>&</sup>lt;sup>1</sup> The Indicator Subgroup of the Social Policy Committee (SPC) has stated that measuring performance (in particular: efficiency) is a very complex task that requires appropriate and valid outcome indicators that need further elaboration. The OECD recently proposed to develop health care efficiency indicators. This work will also be supported by a joint European Commission-OECD project on Efficiency of health systems. See "Progress Report on the review of the Joint Assessment Framework in the area of health" at: http://ec.europa.eu/social/BlobServlet?docId=13724&langId=en

 $<sup>^2</sup>$  For the United States, Berwick and Hackbarth (2012) estimated that health sector inefficiencies amount to around 1/3 of total expenditure. These authors calculate that gradually improving health sector efficiency would allow stabilising the health expenditure to GDP ratio, which otherwise is projected to increase from 17.7% in 2011 to over 20% of GDP in 2020. This is a conservative estimate compared with a IBM study, which surveyed a large number of economists that ranked the health sector as the economic sector with the highest percentage of inefficiency (estimated to be above 40%).

<sup>&</sup>lt;sup>3</sup> In the empirical literature, a large number of efficiency analyses has been published. For example, comparing DEA indices of healthcare efficiency in 191 countries (WHO, 2000; Hollingsworth and Wildman, 2003), with SFA estimates (Green, 2004). The Google Scholar search engine returned 16,600 articles after a search on "efficiency", "dea", "sfa", and "health" for the period 1990-2014 on 6 September 2014. In addition, there is a considerable literature comparing the results of both methods (Hjalmarsson et al., 1996; De Cos and Moral-Benito, 2011). There were also studies on primary care, physicians, pharmacies, nursing homes and purchasers of care. Many studies used DEA but, in more recent years, an increasing number of studies use Stochastic Frontier Analysis (see e.g. Hollingsworth and Street, 2006; Hollingsworth, 2008; Hollingsworth and Peacock, 2008).

health outcome indicators, such as life expectancy, healthy life expectancy<sup>6</sup> and amenable mortality rates.<sup>7</sup> Inputs may include expenditure on health care, physical inputs and environmental variables. The aim is to assess whether efficiency scores are robust (i.e. within a relatively narrow interval) across a comprehensive battery of models. This is important, as previous studies have shown that results may differ significantly depending on the models chosen; thereby, we give considerable emphasis to the robustness of our results, by testing their consistency across different model specifications.

In order to assess (relative) technical efficiency in the health sector, this paper mainly uses non-parametric frontier methods based on Data Envelopment Analysis (DEA). DEA is a non-parametric<sup>8</sup> technique, where all deviations between observed values and an estimated production possibility frontier are attributed to inefficiency. In addition, Stochastic Frontier Analysis (SFA) is used as a sensitivity analysis. SFA methods require assuming a particular functional form for the production function, which allows for the presence of both stochastic errors and inefficiency.

This paper is organised as follows. In Section 2, we make the case that the empirical literature has found ample evidence of resource waste in health sectors. In Section 3, we apply a battery of DEA models to obtain efficiency scores. In Section 4, and based on these scores, we provide robustness checks through an evaluation of their internal consistency. Section 5 identifies a number of underperforming countries. Section 6 provides a ballpark estimate for the reduction in health expenditure resulting from the adoption of the more efficiency practices across countries. In a technical Annex, a few SFA models are also used to estimate efficiency scores, together with an estimation of productivity growth in health care based on the calculation of Malmquist indexes. Finally, Section 7 concludes.

## 2. Overview of empirical evidence on health efficiency

There is a considerable body of evidence at both the macro and micro levels on the pervasiveness of inefficiency in the health sector. Many findings of wasteful use of resources have been reported in the empirical literature, inter alia: i) sub-optimal setups for delivery of care; ii) inefficient provision of acute hospital care; iii) fraud and corruption in health care systems; iv) large unexplained variation in the quantity and quality of care across and within countries; and, v) a sub-optimal mix of preventative versus curative care. In this Section, some major empirical results are summarised.

Joumard et al. (2010) argue that institutional characteristics can have a significant impact on measured efficiency, suggesting that a reconfiguration of current policies, together with appropriate institutional reform could improve overall efficiency. Efficiency scores seem to be closely related to a number of institutional features of health systems, inter alia: i) the allocation of resources between in- and out-patient care; and, ii) the payment schemes or incentives for care providers. The 2010 EPC/EC Joint Report on Health Systems (European Commission, 2010) and a study on fiscal sustainability challenges (European Commission, 2014), including the health and long-term care areas, provide a vast number of concrete country-specific examples of potential inefficiencies, listing possible remedies. Examples of potential inefficiencies relate to: i) suboptimal mix between private and public funding; ii) mismatch of staff skills; iii) suboptimal provision of primary health care services; iv) unnecessary use of specialist and hospital care; v) too few day-case surgeries and missing concentration of hospital services; vi) deficiencies in general governance of health systems and lack of managerial skills; vii) insufficient data collection, IT use, and health technology assessment to improve decision-making processes; and, viii) inadequate access to more effective health promotion and disease prevention.

<sup>&</sup>lt;sup>4</sup> Allocative efficiency means the ability of a decision unit to use inputs in optimal proportions, given their prices and the production set. Productivity efficiency implies both technical and allocative efficiency.

<sup>&</sup>lt;sup>5</sup> In DEA jargon, the former uses output orientation, while the latter uses input orientation.

<sup>&</sup>lt;sup>6</sup> Health life expectancy is a measure of disability-free life expectancy which indicates how long people can expect to live without disability.

<sup>&</sup>lt;sup>7</sup> Amenable mortality rate is the percentage of deaths that could be avoided through timely access to adequate medical intervention.

<sup>&</sup>lt;sup>8</sup> This does not require assuming a particular functional form for the production function.

OECD (2014) points to significant inefficiencies in the use of resources, based on major geographical variations in medical practice across and within OECD countries. Excess variability raises questions about the quality, equity, and efficiency of resource allocation and use in general. As an example, across countries, cardiac procedure rates vary more than three-fold, while hospital medical admission rates vary twofold. Variations are also pronounced within countries. However, no conclusive evidence is given on whether unnecessary care is being delivered in areas of high activity, or alternatively needs are left unmet in regions of low activity. However, it seems quite clear that health systems are underperforming overall. While this study does not determine precisely how much of these variations are unwarranted, the OECD concludes that they are too large to be explained solely by patient needs and/or preferences.

In the hospital sector, numerous studies have documented the large degree of inefficiency at the level of acute care. In an exploratory study, OECD (2008) assesses cross-country differences<sup>9</sup> in health care performance at the hospital level. Unit costs for studied interventions differ significantly. Ignoring quality differences, calculations suggest that costs could be dramatically reduced by between 5% to 48% relatively to the best performer for each type of intervention. Comparisons among Nordic hospitals show that even within countries with relatively similar institutional features, there seems to be a large cost-saving potential, ranging from 23% to 44%. There is also evidence of a large within-country dispersion of costs, indicating that there is scope for cost savings if underperforming hospitals adopt best national practices.

In "Health at a Glance" (2014), the OECD points towards inefficiencies in different areas of care. Shorter hospital stays and growing use of generic drugs have saved costs, but persistent large variations in medical practice point towards further potential inefficiencies. There are wide variations in the use of diagnostic and surgical procedures, which cannot be explained by differences in clinical needs.<sup>10</sup> Despite recorded improvements in quality of acute care and primary care, further improvements are warranted given current resource use. This is the case for avoidable hospital admissions for chronic diseases such as asthma and diabetes. Also, countries should improve primary care to further reduce costly hospital admissions for these conditions.

Based on the variation in the number of interventions both at the regional- and hospital level, EuroHOPE<sup>11</sup> results suggest that efficiency could be improved. In Finland, Hungary, Italy, the Netherlands, Norway, Scotland, and Sweden, this project followed patients with specific conditions<sup>12</sup> for one-year after onset of the disease. There were differences in the performance of health care systems in all of the health problems or diseases considered. In addition, in all the countries there were wide regional- and hospital-level differences.

The EuroDRG<sup>13</sup> project explored inefficiencies related to payment systems based on diagnosis-related groups (DRG) in 12 European countries, including potential for improvement. Unintended consequences of DRG based hospital payments included cherry picking, dumping, upcoding, overtreatment, and frequent readmissions.<sup>14</sup> The project states that intentional upcoding and overtreatment are substantial problems in France and Germany. In Germany, out of 12% of hospital cases reviewed (about two million cases), about 40% of those contained coding errors or overtreatment, mostly unnecessary admissions or excessive length of stay.<sup>15</sup> A related study found

<sup>&</sup>lt;sup>9</sup> Involving the following countries: Denmark, Finland, France, Germany, Iceland, Norway, Sweden, the United Kingdom and the United States.
<sup>10</sup> For example, in 2011, caesarean sections made up more close to 40% of all births in Italy, and 17% in the Netherlands, possibly

<sup>&</sup>lt;sup>10</sup> For example, in 2011, caesarean sections made up more close to 40% of all births in Italy, and 17% in the Netherlands, possibly suggesting overuse. <sup>11</sup> EuroHOPE is a project funded under the European Commission's 7<sup>th</sup> Framework programme with the objective of evaluating the

<sup>&</sup>lt;sup>11</sup> EuroHOPE is a project funded under the European Commission's 7<sup>th</sup> Framework programme with the objective of evaluating the performance of European Health Care systems in terms of outcomes, quality, use of resources and costs. http://www.eurohope.info/doc/summary.pdf

<sup>&</sup>lt;sup>12</sup> Acute myocardial infarction, ischaemic stroke, hip fracture, breast cancer, very low birth weight (VLBW), and very low gestational age (VLGA). <sup>13</sup> From DBC are a superconductive for did by the Francesco Completion  $Z^{th}$  Francesco I for a superconductive for a superconductive for the formula term of a superconductive formula term.

<sup>&</sup>lt;sup>13</sup> EuroDRG was a research project funded by the European Commission's 7<sup>th</sup> Framework programme. It formed a team of researchers from twelve European countries (Austria, England, Estonia, Finland, France, Germany, Ireland, the Netherlands, Poland, Portugal, Spain and Sweden). They analysed the national DRG-based hospital payment systems by using qualitative and quantitative research methods.

<sup>&</sup>lt;sup>14</sup> Cherry picking occurs if hospitals exploit payment incentives to select the less costly, more profitable patients and/or to "dump" them, i.e. transfer or avoid the unprofitable patients. Upcoding refers to coding additional diagnoses on patients to achieve higher payment. Hospitals may also re-admit patients for unnecessary services or misplaced services (e.g. those better placed at outpatient settings.

<sup>&</sup>lt;sup>15</sup> According to the article: "Examples of upcoding included newborns with a secondary diagnosis of "need for assistance with personal care" (ICD-10:Z74.1), patients with an asymptomatic urinary tract infection coded as acute cystitis, and "miscounting" the number of hours for patients with artificial ventilation (leading to substantially higher payments). Some hospitals were found to use procedure codes for "geriatric

interestingly that within countries cost-differences were not related to differences in quality of care. This shows a potential for improving performance by containing cost or improving quality/outcome.<sup>16</sup>

Fraud and corruption<sup>17</sup> in health care systems are another source of waste (European Commission, 2013). This study concludes that corruption in the health sector occurs in all EU MSs, and that both the nature and the prevalence of its typologies differ across the EU.

Empirical evidence on health inefficiency relates it also to a suboptimal policy-mix between preventive and curative care. It is universally acknowledged that lifestyle factors, such as tobacco smoking, obesity, wrong diet and lack of physical activity have a significant impact on health outcomes, increasing demand for health services. Major chronic diseases can often be prevented through lifestyle changes. Prevention policies may lead to a longer period of life without diseases and reduce costs. However, the health benefits of prevention may also increase the overall life span in such a way that especially older people can live longer but with chronic diseases. This increases health care demand and ultimately costs. Thus, in net terms it is not certain that cost savings associated with better outcomes outweigh the possible higher costs associated with bad health at older ages.

According to the OECD (2010),<sup>18</sup> although having access to a number of preventive health interventions may generate a meagre reduction in the order of 1% in total expenditure for major chronic diseases, what seems to matter is the adoption of an optimal mix of spending on prevention and treatment. For example, moving resources from treatment to prevention of cardiovascular diseases or diabetes will increase the cost-effectiveness of spending, while relying on treatment alone will be suboptimal (AcademyHealth, 2012). It may be safely assumed, that an optimal mix of these policies is yet to be achieved in many EU countries.

The Healthbasket<sup>19</sup> project tried to gather information on the basket of services offered in different Member States, how they are defined, how often they are used for particular patients, what are their costs and what prices are paid for them. Overall, while differences in average costs were significant between countries, within-country variation was also unexpectedly large - in some cases, larger than between-countries. These differences are partly due to different accountancy standards, but also due to prices and, most importantly, due to large and apparently real differences in practice (and therefore differences in actual coverage of services).<sup>20</sup> The Healthbasket project suggests that intra-country variation may be larger and more significant for many medical services than inter-country variation.

Finally, Carone et al. (2013) suggest that there are obvious gains to be made in pharmaceutical spending without decreasing patient access to high-quality pharmaceuticals. Often, decisions made to pay for a medicine with public money are not fully transparent, based on relevant criteria and are difficult to change. Prices are not revised on regular basis and taking into account new evidence of cost-effectiveness of drugs already on the reimbursement list. Incentives for rational use of medicines are often not in place, especially policies favouring generic substitution appear suboptimal in several EU countries. Tendering for purchasing pharmaceuticals in a hospital setting is not used to its full potential and tools for improving prescribing behaviour of doctors could be further expanded.

http://www.oecd.org/health/economics-of-prevention.htm <sup>19</sup> http://www.oecd.org/health/health-systems/38680411.pdf

early rehabilitation," although they did not have geriatric specialists. Others admitted patients without proper justification for procedures that should, in general, be performed on a day case basis; and a large number of hospitals were found to discharge patients later than necessary."

<sup>&</sup>lt;sup>16</sup> See also special edition: "Diagnosis-Related Groups in Europe (EuroDRG): Do they explain variation in hospital costs and length of stay across patients and hospitals?", in Busse R, Geissler A, Mason A, Or Z, Scheller-Kreinsen D, Street A (2012) Health Economics, Volume 21 (Supplement 2)

Corruption relates to bribery in medical service delivery, procurement corruption, improper marketing relations, misuse of (high) level positions, undue reimbursement claims and fraud and embezzlement of medicines and medical devices.

<sup>&</sup>lt;sup>20</sup> Other explaining factors include data recording, cost-shifting to patients, exchange rates, demarcation of service to other sectors, etc.

## 3. Measuring efficiency

Health expenditure in percentage of  $\text{GDP}^{21}$  has continued to rise in all EU countries over the past decades, despite sustained policy efforts to arrest this trend. In 2011, total spending on health amounted to about 10.2% of GDP in the EU. The cross-country variation is wide, however, ranging from 5.7% in Romania to between 11% to 12% in Denmark, Germany, France and the Netherlands. Health spending is projected to continue rising at a faster pace than income, driven mainly by technological innovation, relative price developments, as well as demographic factors (Maisonneuve and Oliveira, 2013).

Increased spending was accompanied in the past by improved health outcomes. However, the degree of improvement in health outcomes varies considerably across countries (Joumard et al., 2010; Heijink R. et al., 2015). High spenders do not necessary rank high in terms of health outcomes (Graph 1). For instance, Spain records the highest life expectancy, but is a median spender compared to other EU countries. Conversely, Belgium and Denmark rank among the high spenders, but reach only average levels of health outcomes. Also, indicators which are more closely related to health system performance, such as healthy life years at age 65 and amenable mortality rates show larger differences in outcomes for the same level of spending than life expectancy, probably because the former are more affected by variables outside the control of health systems. This cross-country variation in outcomes is often interpreted as an indication of potential health system inefficiency.

Empirical evidence suggests a non-linear relationship between health spending and outcomes, reflecting the impact of other factors, inter alia, historical expenditure patterns on health and other welfare policies, socioeconomic variables, lifestyle behaviour, and environmental factors. Thus, any methodology attempting to estimate the efficiency of health spending needs to take into account a wide range of relevant variables in the functioning of health systems to obtain unbiased and efficient estimates. However, in practice due to issues of causality and lack of data, this can be achieved only to a limited degree.

Similarly to previous analyses, we mainly use Data Envelopment Analysis (DEA) to derive estimates (called scores) of relative technical efficiency in transforming inputs into outputs. We apply a wide range of different models to measure both potential inputs/expenditure savings<sup>22</sup> (for given outputs), and potential increases of outputs (for given inputs)<sup>23</sup>. We test a wide range of models in order the provide robustness checks. The Estimated DEA models are listed in Table 1.<sup>24</sup> For a description of the datasets and DEA estimation techniques see Annexes 1, 2 and 3.





<sup>&</sup>lt;sup>21</sup> In this paper, health expenditure includes both public and private spending on current goods and services and capital investment.

<sup>&</sup>lt;sup>22</sup> Input orientated models.

<sup>&</sup>lt;sup>23</sup> Output oriented models.

<sup>&</sup>lt;sup>24</sup> Given the (internal) validation of the various model estimates of efficiency, and in line with the principle of parsimony (Occam's razor), we prefer using simple DEA models with only one or two outputs and one or two inputs.



*Notes*: For a detailed description of the variables see Annexes 1 and 2.

Health outputs are measured mainly in terms of outcomes. The following 7 health outcomes are used: (adjusted) life expectancy at birth and at age 65, (adjusted) healthy life expectancy at birth and at age 65, and standardised amenable mortality.<sup>25</sup> Thus, this paper uses both broad and narrow indicators, in terms of their relation to health system performance. Life expectancy can be viewed as a broad indicator that is influenced by many factors, besides the provision of health services, namely a wide range of socio-economic variables, such as education, income, working conditions, and lifestyle behaviour. Life expectancy has the added advantage of being easily calculated and widely available both across time and space. Healthy life years and standardised amenable mortality are mode difficult to define and calculate, but conversely represent narrow indicators more directly linked to health institutions and policies, thereby being potentially more relevant for policy recommendation.

 $<sup>^{25}</sup>$  For a description of the variables see Annex 1. For a presentation of the dataset see Annex 2.

#### Table 1 – DEA models used

Output variations			Input variations		
Health outcomes	Model 1		Model 2		Model 3
Life expectancy at birth	Health	Health	A composite indicator of the	Indicators of	A composite indicator of the
	expenditure	expenditure per	socio-economic environment	physical inputs	socio-economic environment
Life expectancy at age 65	per capita, in	capita, in PPP	(GDP per capita, educational	(hospital beds,	(GDP per capita, educational
Healthy life expectancy at age 65	PPP		attainment) and lifestyle factors	nurses,	attainment) and lifestyle
Healthy life expectancy at birth			(lagged consumption of alcohol	physicians per	factors (lagged consumption of
Amenable mortality			and tobacco, obesity)	capita)	alcohol and tobacco, obesity)
	1	Health	A composite indicator of the		A composite indicator of the
Life expectancy at birth - adjusted for		expenditure per	socio-economic environment		socio-economic environment
lifestyle (according to Heijink R. 2015)		capita - adjusted	(GDP per capita, educational		(GDP per capita, educational
Healthy life expectancy at birth -		for life-style, in	attainment)		attainment)
adjusted for lifestyle (according to		РРР			
Heijink R. 2015)					

Notes: For a detailed description of the variables and DEA estimation techniques used see Annexes 1, 2 and 3. We use both "physical" and monetary variables as inputs. The former have the advantage of not requiring the use of dedicated price deflators, thereby not being affected by potential large price distortions present in health sector outlays. Conversely, monetary variables, particularly total health expenditure, can give a more comprehensive measure of efficiency that potentially encompasses all resources.

We also use adjusted measures of life expectancy and healthy life expectancy at birth based on Heijink R. et al. (2015).<sup>26</sup> They simulate the impact of lifestyle differences on health outcomes across EU countries based on a micro-simulation model. In their model, health outcomes and health spending are based on individual level data, which allows for a more accurate adjustment of health outcomes and health spending to lifestyle differences across individuals' life-courses than it is possible to make using macro data.

We also assess intermediate outputs in terms of hospital discharges and outpatient consultations to measure subsector efficiency, which is particularly relevant for in- and out-patient care. However, for a number of methodological reasons, this approach is problematic (Journard 2010).<sup>27</sup> Thus we do not emphasise these results, but present them for the sake of comparability with those obtained using health outcomes.

We define inputs in terms of: i) total per capita health expenditure in purchasing power parities  $(PPP)^{28}$ ; ii) per capita physical inputs, such as hospital beds, and the number of physicians and nurses, and; 3) environmental or lifestyle variables, such as smoking, alcohol consumption, diet, education, and income, which can be important explanatory factors of health outcomes.

## 4. Estimation results

Results derived from 21<sup>29</sup> DEA models<sup>30</sup> suggest that there might be significant scope to increase efficiency across the EU. Graphical presentation of DEA analysis can be very helpful in interpreting results.<sup>31</sup> Graph 2 draws the production function for the DEA model with life expectancy as the outcome variable and per capita health expenditure (using health PPP) as the input variable. An intuitive grasp of (relative) efficiency rankings,

<sup>&</sup>lt;sup>26</sup> This study was funded by the European Commission and carried out by the Dutch National Institute for Public Health and the Environment (RIVM). http://ec.europa.eu/health/systems\_performance\_assessment/docs/2015\_maceli\_report\_en.pdf

Journard et al. (2010) give two main reasons for preferring using outcome variables in frontier analyses, such as DEA and SFA models, rather than output variables of the healthcare system, namely i) lack of fully consistent output data across countries; and ii) although individual (medical) outputs may be produced efficiently, they might still have a very limited impact on the health status of the population if the overall mix of medical treatments/procedures is unbalanced. Specific, Eurostat 2010 PPPs for the health sector are used.

<sup>&</sup>lt;sup>29</sup> 21=7 output variables \* 3 types of models (Table 1).

<sup>&</sup>lt;sup>30</sup> For some technical details concerning DEA methodology see Annex 3.

<sup>&</sup>lt;sup>31</sup> Graphical analysis is commonly carried out in the following two cases: i) 1 output and 1 input variables (production function); or, ii) two input variables (isoquants, see Annex 4).

either using output or input orientation,<sup>32</sup> is immediate from observation of country positions relative to the production function/frontier.





Notes: Horizontal/vertical distances to the production function measure the degree of inefficiency (using input or output orientation, respectively). The various production functions correspond to no bias correction (DEA), bias correction (Boot), and a confidence interval of bias correction values (CI97.5)

It should be highlighted that results vary not only across DEA models (Section 5), but the bootstrapping method used (Simar and Wilson, 1998) provides confidence intervals for bias corrected estimates. Graph 3 suggests that for some countries, bias corrected DEA scores can be subject to considerable uncertainty (i.e. confidence intervals are large), which can potentially change both the assessment on efficiency and actual country rankings.

Sources: Own calculations.

<sup>&</sup>lt;sup>32</sup> Inverse of the vertical or horizontal distances from the observation to the bias corrected frontier, using output or input orientation measures of efficiency, respectively.





Sources: Own calculations. Notes: Countries are ranked in increasing order of (bias-corrected) output efficiency. The 95% confidence interval of bias-corrected DEA scores is added to the Graph, providing an estimation of uncertainty.

Graph 4 plots potential gains per country for the 7 health outcome variables, where for each outcome variable the scores of 3 DEA models have been averaged. Potential gains are calculated as the improvement resulting from moving from country averages to the efficiency frontier. Results suggest that all 7 measures of efficiency tend to be positively correlated (see Annex 4).

In the EU on average, life expectancy at birth could be increased by 2.3% or 1.8 years (1.2 years at age 65) by moving from current positions to the efficiency frontier.<sup>33</sup> For the worst performer, Slovakia, the rise could amount to 6.4 years for life expectancy at birth and 3.2 at age 65. Average healthy life expectancy in the EU could increase on average by 6.1 years at birth and 2.9 at age 65. Average amenable mortality rates could nearly be halved in the EU by moving to the efficiency frontier.

<sup>&</sup>lt;sup>33</sup> Changes in the outcome variable resulting from moving from the cross-country median of the average of the scores for the three models to the efficiency frontier. As a measure of location, the median is more robust to outliers than the average.

#### Graph 4 – Potential gains in health outcomes



*Sources:* Own calculations. Countries ranked in increasing order of life expectancy at birth. *Notes:* i) Potential gains are estimated by measuring the number of years of life that could be added (or the proportion of amenable mortality that could be reduced), if a country moves from inside to the efficiency frontier, while holding inputs constant (output orientation in DEA). ii) Models are described in **Table** 1. iii) Graph shows averaged gains in health outcomes across the three models available for each health outcome.

Graph 5 shows DEA results for the 7 outcome variables. The corresponding efficiency scores are presented in Annex 5. General comments on the results are the following:

- Sensitivity to changes in inputs appears to be high for low performers, such as Hungary and Poland. This is particular evident when using life expectancy outcomes. However, uncertainty surrounding efficiency scores (i.e. confidence interval) is higher for countries close to the median in the sample, when using amenable mortality as outcome variable (Panel E).
- Although country rankings vary according to the outcome variable/DEA model being used, overall we obtain a clear and consistent picture in terms of country rankings (Section 5).
- Efficiency scores seem to have a higher sensitivity to changes in outcome/output than to input variables.
- Relative efficiency scores are not significantly affected by using either adjusted life expectancy or healthy life expectancy measures calculated by Heijink R. et al. (2015).

Additional sensitivity analysis suggests the following conclusions:

- In certain circumstances, namely for countries that appear to be close to best practice (Annex 5), confidence intervals are wide, making challenging the interpretation of results.
- When using intermediate output measures, such as hospital discharges or outpatient consultations, instead of one of the seven outcome measures previously employed, country efficiency rankings can change dramatically. For example, while the Netherlands and Cyprus appear efficient according to outcome measures, both countries score poorly when using intermediate outputs.





Sources: Own calculations. Countries ranked in decreasing order of potential gains.

*Notes*: i) Potential gains are estimated by measuring the number of years of life that could be added (or the proportion of amenable mortality that could be reduced), if a country moves from inside to the efficiency frontier, while holding inputs constant (output orientation in DEA). ii) Models are described in **Table 1** 

## 5. Efficiency scores used to group/cluster countries

The sensitivity of DEA scores to the model used is not surprising. In this section, we investigate whether consistent patterns emerge across models, allowing for a meaningful country classification (or clustering). In Table 2 and Annex 5, efficiency scores are colour coded: "red" meaning below the 25th percentile (worst performing countries); "green" above the 75<sup>th</sup> percentile (best performing countries); and "white" for countries laying in the interquartile range IQR (i.e. between the 25<sup>th</sup> and the 75<sup>th</sup> percentile). Results of our analysis suggest that often countries consistently fall in the same category across various models.<sup>34</sup>

For a better interpretation of results, efficiency scores are grouped in Table 2, according to:

- 1. Mean score by health outcome measure (results in part (1) of Table 2);<sup>35</sup>
- 2. Overall mean score across the 21 models (results in part (2) of Table 2);
- 3. Count the number of times a country scores in the lowest 25<sup>th</sup> percentile, in the interquartile range, and in the highest 75<sup>th</sup> percentile across all models (results in part (3) of Table 2);
- 4. A cluster analysis performed to classify countries (results in part (4) of Table 2).<sup>36</sup>

The main features that emerge can be described as:

- The **Czech Republic, Lithuania** and **Slovakia** are consistently among the lowest scorers. Overall, these countries are among the seven worst performers, and are clustered in the bottom group of countries.<sup>37</sup> For instance, Lithuania's efficiency scores are consistently among the lowest, according both to individual outcomes and the overall mean. Lithuania scores in the lowest 25<sup>th</sup> percentile in 19 models, being in the IQR group in only 2 models. Finally, cluster analysis suggests that Lithuania belongs to the group of low performers.
- Hungary, Latvia, Poland and Estonia although scoring marginally better than the previous group are also underperformers. In about half of the models considered, they score in the lowest quartile, while obtaining scores in the IQR in the other half.
- Belgium, Cyprus, Spain, France, Luxembourg, Sweden and the Netherlands consistently score among the top seven performers in most of the models and are clustered in the group of countries with highest average efficiency scores.
- Scores for the remaining countries are generally around the IQR, depending on the model.

<sup>&</sup>lt;sup>34</sup> For instance, Hungary in most models scores consistently in the bottom 25<sup>th</sup> percentile, being one of the 7 worst performers overall, while France appears to be efficient across many models (above the 75<sup>th</sup> percentile), being among the top 7 performers overall. <sup>35</sup> Recall that there are three DEA models for each output measure.

<sup>&</sup>lt;sup>36</sup> Both K-means and an agglomeration hierarchical method are used.

<sup>&</sup>lt;sup>37</sup> They form a cluster in the k=2 analysis, while in the dendrogram they appear in a common branch (**Graph 7**).

		AVERAGI	EEFFICIEN	CYSCORE	S by health o	nutcome (1)			Efficiency DEA m (healt) expecta	across all odels on ny) life nncy (2)		Counting lo ac	ow and high ross models	performers (3)	Cluster a efficiency ac models o outcon	nalysis on ross all DEA n (health tes) (4)	
	Life expectancy at birth, years	Healthy life expectancy at birth, years	Life expectancy at age 65, years	Healthy life expectancy at age 65, years	Amenable mortality	Life expectancy at birth according to Heijink R. et al (2015), years	Healthy life expectancy at birth according to Heijink R. et al (2015), years		Average score	Ranking		Lowest to 25th	From 25 to 75th	75th to highest	On 3 clusters	On 2 clusters	
AT	0.98	0.92	0.94	0.83	0.51	0.98	0.94	AT	0.87	16	AT	1	20	0	1	1	AT
BE	0.98	0.92	0.93	0.85	0.68	0.98	0.94	BE	0.90	7	BE	1	15	5	1	1	BE
BG	0.97	0.91	0.95	0.80	0.48	0.96	0.95	BG	0.86	18	BG	4	15	2	1	1	BG
CY	0.99	0.96	0.97	0.83	0.64	0.99	0.97	CY	0.91	4	CY	0	8	13	1	1	CY
CZ	0.94	0.81	0.88	0.64	0.33	0.94	0.89	CZ	0.78	26	CZ	21	0	0	2	2	CZ
DE	0.98	0.90	0.95	0.87	0.51	0.98	0.95	DE	0.88	13	DE	4	15	2	1	1	DE
DK	0.97	0.86	0.93	0.86	0.58	0.97	0.94	DK	0.87	15	DK	8	11	2	1	1	DK
EE	0.96	0.93	0.91	0.76	0.53	0.95	0.91	EE	0.85	20	EE	8	10	3	2	1	EE
EL	0.98	0.93	0.95	0.80	0.55	0.98	0.95	EL	0.88	14	EL	1	16	4	1	1	EL
ES	0.99	0.96	0.96	0.86	0.66	0.99	0.97	ES	0.91	3	ES	0	7	14	1	1	ES
FI	0.98	0.94	0.95	0.87	0.50	0.97	0.96	FI	0.88	12	FI	1	14	6	1	1	FI
FR	0.99	0.97	0.95	0.92	0.81	0.99	0.95	FR	0.94	1	FR	0	5	16	1	1	FR
нк	0.98	0.92	0.91	0.71	0.52	0.98	0.94	нк	0.85	21	нк	6	11	4	1	1	нк
HU	0.95	0.88	0.91	0.71	0.38	0.96	0.92	HU	0.81	25	HU	8	11	2	2	1	HU
IE	0.98	0.92	0.97	0.89	0.53	0.97	0.97	IE IT	0.89	11	IE 17	1	12	8	1	1	IE 17
11	0.99	0.95	0.93	0.74	0.74	0.98	0.92		0.89	9		5	5	11	1	1	11
	0.92	0.87	0.85	0.54	0.33	0.92	0.87		0.76	2/		19	2	0	3	2	
	0.98	0.92	0.94	0.87	0.70	0.97	0.94		0.90	24		11	14	0	2	1	
MT	0.95	0.91	0.91	0.74	0.42	0.95	0.93		0.85	10		0	0	12	1	1	
NI	0.98	0.91	0.98	0.88	0.33	0.98	0.98	NI	0.83	2	NI	2	1	18	1	1	NI
PI	0.95	0.92	0.91	0.70	0.48	0.96	0.92	PI	0.84	23	PI	11	10	0	2	1	PI
PT	0.98	0.95	0.93	0.74	0.53	0.98	0.93	PT	0.86	17	РТ	4	15	2	1	1	PT
RO	0.98	0.93	0.95	0.79	0.38	0.98	0.95	RO	0.85	22	RO	3	18	0	1	1	RO
SE	0.99	0.93	0.98	0.93	0.54	0.98	0.97	SE	0.90	6	SE	0	11	10	1	1	SE
SI	0.97	0.92	0.91	0.74	0.53	0.97	0.92	SI	0.85	19	SI	6	15	0	1	1	SI
SK	0.92	0.80	0.84	0.50	0.24	0.92	0.85	SK	0.72	28	sк	21	0	0	3	2	SK
UK	0.98	0.93	0.96	0.97	UK	0.89	8	UK	0	14	7	1	1	UK			
-	Above the	75th perc	entile							-					Top cluster		
	Below the	25th perc	entile												Intermedia	e cluster	
															Bottom clus	ter	
															bottom cius		

#### Table 2 – Analysis of DEA efficiency scores and country clusters

Source: Own calculations.

*Notes*: 1. (1) = Mean score by health outcome measure;

(2) = Overall mean score across the 21 models;

(3) = Count the number of times a country scores in the lowest 25th percentile, in the interquartile range, and in the highest 75th percentile across all models;

(4) = A cluster analysis performed to classify countries.

We carried out clustering analysis of the 21 DEA model scores, using both k-means and hierarchical methods.<sup>38</sup> The last two columns of Table 2 present the clustering results using either 2 or 3 clusters. In Graph 6, we plot average efficiency scores for the 21 DEA models by cluster.<sup>39</sup> Results suggest that clusters have significantly different average efficiency scores. Graph 7 presents the dendrogram corresponding to a hierarchical agglomerative analysis of 21 DEA model scores. Based on the efficiency indicators, the dendrogram suggests proximity among a number of countries. As an example, for underperforming countries, the following groups of

<sup>&</sup>lt;sup>38</sup> K-means clustering aims to partition observations into k clusters such that the sum of squares ("distance") from points to the assigned cluster centres is minimised. Agglomerative (hierarchical) clustering is a "bottom up" approach, where each observation starts in its own cluster, and clusters are successively merged as one moves up in the hierarchy. A measure of dissimilarity or "distance" between clusters is needed to decide which clusters should be combined first. An "euclidean" metric is used. <sup>39</sup> K-means uses the function *kmeans* of the package *sets* in R. Agglomerative hierarchical clustering uses the function *agnes* in the package

cluster in R.

countries are found to be close: i) CZ, SK and LT; ii) EE and HU; while among better performing countries the following clusters are found to be close: i) AT and DE; ii) IE and the UK; iii) BE and LU; iv) CY, ES and FR.; and v) EL and RO.



Graph 6 – Average efficiency scores using 2 or 3 clusters





Having a closer look at the underlying characteristics of countries, it appears that on average countries with low life expectancy achieve lower efficiency scores (Graph 8, left panel). As regards per capita expenditure levels in PPP, it is interesting to observe that high spenders (75th percentile and above) tend also to have higher average

efficiency scores, although the range of efficiency levels is very wide for countries in the IQR for income (Graph 8, right panel).





Sources: Own calculations.

*Notes*: Countries grouped by life expectancy and health expenditure per capita measures in PPP; Average efficiency scores across all models based on measures of health outcomes.

Table 3 presents 5 outcome variables for the 7 worst performing countries reflected in their relatively low levels of (healthy) life expectancy and high amenable mortality rates (see Annex 10 for a graphical representation). These seven countries make up clusters 2 and 3 in the k-means clustering exercise that uses three clusters reported in Table 2. Median outcomes of low efficiency countries are considerably and systematically worse than those for other countries.

It is useful to recall that the type of macro-efficiency, as analysed in this paper, does not exclude inefficiencies in health service production and delivery at the micro- and meso-levels of health systems. There are numerous valid recommendations for health system reform for those countries, which are found to be efficient in our analysis (European Commission 2010). For example, Cyprus has substantial inefficiencies related to fragmented health system financing and provision of care (Theodorou, M. et al 2012), which culminated in the government's efforts to implement a key health system reform (Cylus, J. et al, 2014).

Countries with										
low efficiency	Life expecta	ncy at birth,	Healthy life	expectancy	Life expec	tancy at 65,	Healthy life	expectancy	Amenable	Mortality
scores	ye	ars	at birth	n, years	ye	ars	at age 6	55, years		
	Value	Ranking	Value	Ranking	Value	Ranking	Value	Ranking	Value	Ranking
CZ	78.0	19	69.9	18	17.6	21	12.1	18	2.7	9
EE	76.6 22		66.8	22	18.0	19	11.1	20	3.0	8
HU	75.1	24	65.3	25	16.6	26	9.6	24	3.5	5
LT	73.7	28	62.8	28	17.0	23	7.9 28		4.6	1
LV	73.9	27	63.9	26	16.6 26		9.1	25	4.2	2
PL	76.9	21	66.7	23	17.9	20	10.0	23	1.9	11
SK	76.1	23	66.1	24	16.8	24	9.0	26	3.9	4
Median Values										
High efficiency	tiency 80.9		74.3		19.8		16.1		1.5	
Low efficiency	76.1		66.1		17.0		9.6		3.5	

#### Table 3 – Health outcomes for countries with low DEA efficiency scores

Source: Own calculations.

*Notes*: A worse health status corresponds to a lower ranking in (healthy) life expectancy or to a higher ranking in amenable mortality rates.

## 6. Estimation of potential economic savings from improving efficiency

Another application of DEA is to estimate potential economic savings in health expenditure.<sup>40</sup> Using total expenditure, two methods are used to estimate potential economic savings in health expenditure: one based on the input-orientation of DEA models, another following some econometric work of De Cos and Moral-Benito (2011) (see Annex 8).

It must be noted that this is a highly mechanistic approach, and that results should not be taken at face value. For similar levels of life expectancy, spending levels differ substantially between countries. While naturally part of the differences in spending levels will be related to inefficiencies in the production process of health goods and services, other part will be related to factors, which cannot be (well) captured by macro-level estimation techniques, such as related to lifestyle differences, and institutional and policy settings.



**Graph 9 – Input- versus output-orientation DEA scores** 

Sources: Own calculations.

*Notes*: Horizontal/vertical distances to the production function measure the degree of inefficiency (using input or output orientation, respectively). The various production functions correspond to no bias correction (DEA), bias correction (Boot), a confidence interval of bias correction values (CI97.5), and a 80 years horizontal line (80).

Life expectancy at birth and health expenditure per capita, both adjusted for life-style differences according Heijink R. et al (2015).

In Sections 4 and 5, we discussed DEA scores in terms of output-orientation i.e. given a set of inputs by how much output(s) could be increased if resources would be used efficiently. Using a DEA model with per capita health care expenditure in PPP as the only input<sup>41</sup>, in this section we ask the related question: by how much could input(s) be reduced, while continuing to attain the same output(s) level(s), if resources would be used

<sup>&</sup>lt;sup>40</sup> For an application related to welfare spending see e.g. Matevz Hribernik and Rafał Kierzenkowski (2013), "Assessing the efficiency of welfare spending in Slovenia with Data Envelopment Analysis", OECD, ECO/WKP(2013)50,

http://www.oecd.org/official documents/public display document pdf/?cote=ECO/WKP(2013)50 & docLanguage=Ender (Context) and the second second

<sup>&</sup>lt;sup>41</sup> The same DEA model of Graph 2.

efficiently? The latter approach is called the input-orientation version of DEA. DEA results from input versus output orientation models are only equal in the case of constant returns to scale (CRS) technology. Recall that our DEA models do not assume a CRS technology, but instead a variable returns to scales technology (VRS).

Graph 9 exemplifies quite well the notion that in general the relevance of the two DEA model orientations depends on the context. For countries with life expectancy at or above 80 years, input-orientation appears to be the more relevant criterion, because of the low returns - in terms of added years of life expectancy - that can be gained by increasing per capita expenditure. Conversely, for a group of countries with relatively low life expectancy (PL, HU, CZ and SK), output orientation seems to be the more relevant criterion, particularly because other countries, spending equivalent per capita amounts, have considerable better outcomes (ES, IT, CY, PT, IE, EL and SI).<sup>42</sup>

Following De Cos and Moral-Benito (2011), we use a panel regression to provide a ballpark estimate of potential economic savings from improving efficiency. Saving estimates are calculated by exploring the following counterfactual: by how much could health care expenditure decrease if a country adopted the most efficient system, while keeping the same health outcome. Results suggest economic savings of about <sup>1</sup>/<sub>4</sub> of total health expenditure for the EU as a whole. 43

Graph 10 – Projected increase in public health expenditure in the EU until 2060 (Ageing Report, 2012) taking into account potential efficiency gains



Sources: Own calculations.

Notes: "No efficiency gains" - Projected increase in health expenditure according to Ageing Report (2012). "Improving efficiency by 0.5%" - Assumed projected increase in health expenditure according to Ageing Report, minus annual savings of 0.5% of 2010 expenditure levels due to efficiency gains. The 2012 Ageing Report uses only public health expenditure for expenditure projections. Therefore, in this calculation example we assume that annual savings of 0.5% apply to public health expenditure only.

It can be argued that both DEA (input oriented models) and econometric estimates are both overly optimistic regarding the scale of potential savings. However, achieving substantial savings may be more feasible if implemented over a very long-term period. Using the 2012 Ageing Report (European Commission, 2012) for long-term estimates of health care expenditure, cumulative savings of 22% over 50 years amount roughly to annual savings of 0.5% (22%  $\approx 1 - \frac{1}{1.005^{50}}$ ). Graph 10 suggests that an annual saving in health expenditure due to improved efficiency could halt the increase in healthcare expenditure-to-GDP ratio over the long-term.

<sup>&</sup>lt;sup>42</sup> Using DEA models, the estimation of potential savings can be done using input orientation in models with per capita expenditure as input. <sup>43</sup> For the EU, equivalent to 1.5% of GDP on average.

## 7. Conclusions

There is ample evidence of widespread inefficiency in health care systems. Although the relative ability of a particular health system in transforming resources into outcomes differs across countries, the consensus is that overall there is considerable waste, which contributes to excessive expenditure. This paper aims to estimate relative (technical) efficiency of health care systems across all EU countries.

The application of efficiency concepts to health care systems is challenging, raising both theoretical and practical problems. The relation between inputs and outcomes is complex, driven by factors outside the control of health system managers. This paper takes stock of the relevant empirical literature, particularly previous studies undertaken at the OECD (Joumard et al., 2010) and the WHO (2000), applying well-established methods to recent datasets in order to calculate efficiency scores for all EU MSs.

In order to account for data and model limitations as well as uncertainty, the paper uses a comprehensive battery of models with different combinations of input and output variables. Outputs are the commonly reported health outcome indicators, such as life expectancy, healthy life expectancy and amenable mortality rates. Inputs include (per capita) expenditure on health care, physical inputs and environmental variables. We emphasise the need to check the robustness of the efficiency scores obtained.

Results obtained in this paper are in line with previous empirical research. Large potential efficiency gains can be made in European health systems. On average in the EU, life expectancy at birth could be increased by 2.3% or 1.8 years (1.2 years at age 65) when moving from current positions to the efficiency frontier. Average healthy life expectancy in the EU could increase on average by 6.1 years at birth and 2.9 at age 65. As regards to average amenable mortality rates, they could be nearly halved in the EU, by moving to the efficiency frontier.

Countries are clustered using efficiency scores from more than 20 DEA models. Although efficiency scores can vary considerably across models, the resulting classification appears to identify consistent patterns, thereby supporting the use of DEA models on health systems for benchmarking.

Specifically, it appears that the Czech Republic, Lithuania and Slovakia have the lowest efficiency scores in most of the models used. Hungary, Latvia, Poland and Estonia, although scoring marginally better than the previous group are also underperformers. In about half of the models considered the latter group scores in the lowest quartile, while obtaining scores in the inter-quartile range in the other half. Belgium, Cyprus, Spain, France, Luxembourg, Sweden and the Netherlands consistently score among the top seven performers in most of the models and are clustered in the group of countries with highest efficiency scores. Scores for the remaining EU countries are generally around the inter-quartile range, depending on the model.

It is useful to recall that the type of macro-efficiency, as analysed in this paper, does not exclude inefficiencies in health service production and delivery at the micro- and meso-levels of health systems. There are numerous valid areas for micro-efficiency gains, which can be and are currently addressed in those countries.

Efficiency gains can be measured in two ways: either by increasing health outcomes, while keeping inputs at current levels (output-orientation), or by decreasing inputs, while keeping health outcomes at current levels (input-orientation). For countries with life expectancy at or above 80 years of age, input orientation appears to be the more relevant criterion, because of the low returns - in terms of added years of life expectancy - that can be gained by increasing resource use. Among countries with high life expectancy there are wide variations in per capita health care expenditure, which end up having only marginal effects on health outcomes.

Although varying significantly across countries, results from an econometric regression suggest that total efficiency costs could be potentially huge and may be a source of substantial savings both for the public and private payers. Using the 2012 Ageing Report (European Commission, 2012) for long-term estimates of public health care expenditure, efficiency gains could translate into a 0.5% reduction in the annual growth rate of public health expenditure, eventually halting the increase in the public healthcare expenditure-to-GDP ratio on the EU over the long-term.

#### References

AcademyHealth (2012), "The Economics of Prevention", Policy Brief, http://www.academyhealth.org/files/FileDownloads/RI Econ Prevention.pdf

Adler N. and Golany B. (2001), "Evaluation of deregulation airline networks using data envelopment analysis combined with principal component analysis with an application to Western Europe", European journal of operational research, 132 260-273.

Afonso A. and Aubyn M. (2006), "Relative efficiency of health provision: a DEA approach with non-discretionary inputs", manuscript.

Battese E. and Coelli J. (1992), "Frontier production functions, technical efficiency and panel data: with application to paddy farmers in India", Journal of productivity analysis, 3, 153-169.

Battese E. and Coelli J. (1995), "A model for technical inefficiency effects in a stochastic frontier production function for panel data", Empirical Economics, 20, 325-332.

Banker R., Charnes A., Cooper W. (1984), "Some models for estimating technical and scale inefficiencies in Data Envelopment Analysis", Management science, 30, 1078-1092.

Berwick D. and Hackbarth A. (2012), "Eliminating waste in US health care", JAMA 307(14), 1513-1516.

Busse Reinhard, Geissler Alexander, Aaviksoo Ain, Cots Francesc, Häkkinen Unto, Kobel Conrad, Mateus C., Or Z., O'Reilly J., Serden L., Street A., Tan SS, Quentin W. (2013), "Diagnosis related groups in Europe: moving towards transparency, efficiency, and quality in hospitals?", BMJ (British Medical Journal).

Charnes A, Cooper W. and Rhodes E. (1978), "Measuring the efficiency of decision making units", European journal of operational research 2 429-444.

Cohen D. et al. (2003), "Neighborhood physical conditions and health", American journal of public health, 99(3) 467-471.

Coelli J., Rao P., O'Donnell J. and Battese E. (2005), "An introduction to efficiency and productivity analysis", second edition Springer.

Codagnone C. and Lupiañez-Villanueva F. (2011), "A composite index for the benchmarking of eHealth development in European acute hospitals", JRC Scientific and Technical Reports.

Cylus, J. et al (2014), Moving forward: Lessons for Cyprus as it implements its health insurance scheme, Volume 110, Issue 1, April 2013, Pages 1–5.

De Cos P. and Moral-Benito E. (2011), "Eficiencia y regulacion en el gasto sanitaria en los países de la OECD", Banco de España, Documentos ocasionales No. 1107.

Espen, Erlandsen (2008), "Improving the Efficiency of Health Care Spending: What Can be Learnt from Partial and Selected Analyses of Hospital Performance?" ISSN 0255-0822 OECD Economic Studies No. 44, 2008/1

European Commission (2013), "Study on Corruption in the Healthcare Sector", Directorate-General Home Affairs, HOME/2011/ISEC/PR/047-A2, prepared by Ecorys.

Häkkinen U, Iversen T, Peltola M, Seppälä T, Malmivaara A, Belicza E, Fattore G, Numerato D, Heijink R, Medin E, Rehnberg C. (2013), "<u>Health care performance comparison using a disease-based approach: The EuroHOPE project</u>", Health Policy, Volume 112, Issues 1-2, September, Pages 100-109.

Medin E, Häkkinen U, Linna M, Anthun K, Kittelsen S, Rehnberg C. (2013), "<u>International hospital productivity</u> <u>comparison: experiences from the Nordic countries</u>", Health Policy, Volume 112, Issues 1-2, September, Pages 80-87.

Moger TA, Peltola M. (2014), "<u>Risk adjustment of health-care performance measures in a multinational register-based study: A pragmatic approach to a complicated topic"</u>, SAGE Open Medicine January, December, vol. 2.

Malmivaara A, Meretoja A, Peltola M, Numerato D, Heijink R, Engelfriet P, Wild SH, Belicza E, Bereczki D, Medin E, Goude F, Boncoraglio G, Tatlisumak T, Seppälä T, Häkkinen U. <u>Comparing ischaemic stroke in six</u> <u>European countries. The EuroHOPE register study.</u> Eur J Neurol. 2014 Sep 8. doi: 10.1111/ene.12560

European Commission (2010), "Joint Report on Health Systems", prepared by the European Commission and the Economic Policy Committee (AWG).

European Commission (2012), "The 2012 Ageing Report – economic and budgetary projections for 27 EU Member States (2010-2060)", European Economy N.2.

European Commission (2014), "Identifying fiscal sustainability challenges in the areas of pension, heath care and long-term care policies", Occasional Papers 201.

Färe R. and Lovell C. (1978), "Measuring the technical efficiency of production", Journal of Economic Theory, 19, 150-162.

Färe R. and Lovell C. (1983), "The relative efficiency of Illinois Electric Utilities", Resources and Energy, 5, 349-367.

Färe R., Grosskoff S., Lingren B, and Roos P. (1992), "Productivity changes in Swedish pharmacies 1980-1989: a non-parametric Malmquist approach", Journal of productivity analysis 3 85-101.

Färe R., Grosskoff S., Lingren B and Poullier J. (1997), "Productivity growth in healthcare delivery", 35(4) 355-366.

Farrell J. (1957), "The measurement of productive efficiency", Journal of the royal statistic society, series A, CXX, Part 3, 253-290.

Gay, J. G., et al. (2011), "Mortality Amenable to Health Care in 31 OECD Countries: Estimates and Methodological Issues", *OECD Health Working Papers*, No. 55, OECD Publishing.

Greene W. (2004), "Distinguishing between heterogeneity and inefficiency: stochastic frontier analysis of the World Health Organization's panel data on national health care systems", Health Economics 13 959-980.

Hart C., Smith G. and Blane D. (1998), "Inequalities in mortality by social class measured at 3 stages of the lifecourse", American journal of public health, 88(3) 471-474.

Herr A. (2008), "Cost and technical efficiency of German hospitals: does ownership matter?", Health Economics, 17(9), 1057-1071.

Heijink R. (2004), "Measuring health system performance", Tilburg University.

Heijink R. et al. (2015), "Comparative efficiency of health systems, corrected for selected lifestyle factors", funded by the European Commission and carried out by the Dutch National Institute for Public Health and the Environment (RIVM).

http://ec.europa.eu/health/systems performance assessment/docs/2015 maceli report en.pdf

Hjalmarsson L., Kumbhakar S. and Heshmati A. (1996), "DEA, DFA and SFA: a comparison", Journal of productivity analysis, 7 303-327.

Hollingsworth B. (2003), "Non-parametric and parametric applications measuring efficiency in health care", Health care management science, 6 203-218.

Hollingsworth B. and Smith P. (2003). "Use of Ratios in Data Envelopment Analysis", Applied Economic Letters 10(11) 733-735.

Hollingsworth B. and Wildman J. (2003), "The efficiency of health production: re-estimating the WHO panel data using parametric and non-parametric approaches to provide additional information", Health Economics, 12 493-504.

Hollingsworth B. and Spinks J. (2005), "Health production and the socioeconomic determinants of health in OECD countries: the use of efficiency models", Monash University, Working Paper No. 151.

Hollingsworth, B. (2007), "State of the Art: Efficiency Measurement in Health and Health Care: Non Parametric Efficiency Measurement", *iHEA 2007 6th World Congress: Explorations in Health* 

Economics Paper http://papers.ssrn.com/sol3/papers.cfm?abstract\_id=992983#PaperDownload

IBM Institute for Business Value (2010), "*The world's 4 trillion dollar challenge, Using a system-of-systems approach to build a smarter planet*", Somers, NY: IBM corporation. Available at: http://www-935.ibm.com/services/us/gbs/bus/html/ibv-smarter-planetsystem-of-systems.html

Berwick DM, Hackbarth AD. (2012), "Eliminating waste in US health care", JAMA, 307(14):1513 - 6.

Isaacs S. and Schroeder S. (2004), "The ignored determined of the nation's health", New England Journal of Medicine, 351(11) 1137-1142.

Jacobs, R., P.C. Smith and A. Street (2006), "Measuring efficiency in health care – Analytical Techniques and Health Policy", Cambridge University Press.

Jerrett M., Eyles J. and Cole D. (1998), "Socioeconomic and environmental covariates of premature mortality in Ontario", Social science and medicine, 47(1) 33-49.

Johansson S. and Sundquist J. (1999), "Change in lifestyle factors and their influence on health status and allcause mortality", International journal of epidemiology, 28(6) 1073-1080.

Joumard I., André C., and Nicq C. (2010), "Health care systems: Efficiency and institutions", OECD Economics Department Working Papers No. 769.

Laaksonen M. (2001), "Do health behaviour and psychosocial risk factors explain the European east-west gap in health status?", European journal of public health", 11(1) 65-73.

Maisonneuve C. and Martins J. (2013), "A projection method of public health and long-term care expenditures", OECD Economic Department Working Papers No. 1048.

Matevz H. and Kierzenkowski R. (2013), "Assessing the efficiency of welfare spending in Slovenia with data envelopment analysis", ECO/WKP(2013)50.

Medeiros J. and Schwierz (2013), "Estimating the drivers and projecting long-term public health expenditure in the European Union: Baumol's 'cost disease' revisited", European Economy, Economic Papers No. 507.

OECD (2014), Geographic Variations in Health Care: What Do We Know and What Can Be Done to Improve Health System Performance?, OECD Health Policy Studies, <u>http://dx.doi.org/10.1787/9789264216594-en</u>

OECD (2014), Health at a Glance: Europe 2014, http://www.oecd.org/health/health-at-a-glance-europe-23056088.htm.

Pastor J. (1996), "Translation invariance in data envelopment analysis: a generalization", Annals of operations research 66 93-102.

Ruggiero J. (1999), "Efficiency estimation and error decomposition in the stochastic frontier model: a Monte Carlo analysis", European journal of operational research 115(3), 555-563.

Simar L. and Wilson P. (1998), "Sensitivity analysis of efficiency scores: how to bootstrap in nonparametric frontier models", Management science 44, 49-61.

Simar L. and Wilson P. (2000), "A general methodology for bootstrapping in nonparametric frontier models", Journal of applied statistics, 27, 779-802.

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

Theodorou M, Charalambous C, Petrou C, Cylus J. (2012), "Cyprus: Health system review", Health Systems in Transition. 2012; 14(6):1–128.

Varabyova Y. and Schreyögg J. (2013), "International comparisons of the technical efficiency of the hospital sector: panel data analysis of OECD countries using parametric and non-parametric approaches", Health policy 112, 70-79.

Worthington A. (2004), "Frontier efficiency measurement in healthcare: a review of empirical techniques and selected applications", Medical care research and review 61(2) 1-36.

World Health Organization (2000), "Health systems: improving performance", The World Health Report.

## ANNEX 1: Definition of variables

Indicator group	Individual indicator	Year of data	Definition
Health outcomes	Life expectancy at birth	2012 or latest	Life expectancy at birth, 2012 or latest
	Life expectancy at age 65	2012 or latest	Life expectancy at age 65, 2012 or latest
	Healthy life expectancy at birth	2012 or latest	The indicator of healthy life years (HLY) measures the number of remaining years that a person
	Healthy life expectancy at age 65	2012 or latest	of specific age is expected to live without any severe or moderate health problems. The notion of health problem for Eurostat's HLY is reflecting a disability dimension and is based on a self- perceived question which aims to measure the extent of any limitations, for at least six months, because of a health problem that may have affected respondents as regards activities they usually do. The indicator is therefore also called disability-free life expectancy (DFLE); 2012, or latest.
	Amenable mortality	2011	Standardized death rates for causes of death with amenable mortality per 100 000 inhabitants. Causes of death selected are based on AMIEHS (2011) and availability in Eurostat. In AMIEHS, causes of death were identified that can be considered 'amenable'. International classification of diseases (ICD) 10 codes: Human immunodeficiency virus [HIV] disease (B20- B24); Malignant neoplasm of colon, rectosigmoid junction, rectum, anus and anal canal (C18- C21); Malignant neoplasm of breast (C50); Malignant neoplasm of cervix uteri (C53); Ischaemic heart diseases (I20-I25); Cerebrovascular diseases (I60-I69).
	Life expectancy at birth according to Heijink R. et al (2015), years	2011	Heijink R. et al. have simulated the impact of life-style differences on health outcomes across EU countries based on a micro-simulation model. In their model, health outcomes and health spending are based on individual level data, which allow for a better adjustment of health
	Healthy life expectancy at birth according to Heijink R. et al (2015), years	2011	outcomes and health spending due to life-style differences of individuals over their life-course than is possible using macro-level data.
Intermediate outputs	Outpatient contacts		Number of outpatient contacts with a physician (in a physician's office or at patient's home) excluding dentists consultations per capita
	Hospital discharges		Hospital discharges for all diagnoses (ICD 10: All causes of diseases (A00-Z99) excluding V00-Y98) per capita
	Hospital discharges - weighted		Using the shortlist for the International Classification of Diseases (ICD), containing 20 categories, discharges by diagnostic are aggregated into a weighted measure of total discharges using as weights the length of stay (LOS) by diagnostic category. This aggregation procedure follows the methodology developed by Herr (2008). First, a Global Length of Stay (GLOS) measure is calculated by adding the duration of all stays in number of days across all discharge categories in all countries and dividing by the number of all discharges. Second, an index of discharge weights (W) is constructed as the ratio of length of stay in a particular disease category (LOS) divided by GLOS. The final measure of aggregated discharges adjusted (Agg_disch_adj) by case severity is obtained by multiplying the discharges by categories by the discharge weights and summing across categories.
Inputs	Total health expenditure per capita, PPP	2012, or latest	Total (public and private) current expenditure including capital investment, in real values and expressed as purchasing power parities per capita using healthcare specific PPPs.
	Number of physians per 100 000 inhabitants	2012, or latest	Practicing physians per 100 000 inhabitants.
	inhabitants	2012, or latest	Practicing nursing and caring professionals including midwives per 100 000 inhabitants
Further and a state of the stat	Hospital beds per 1 000 pop	2012, or latest	Curative, psychiatric, long-term care and other hospital beds per 1 000 inhabitants.
Environmental variables	Alconol consumption	Average (1990 - 2003)	Alconol consumption in litres per capita
	Smoking	Average (1990 - 2003)	Regular smokers, % or population aged 15+ that are daily smokers
	Obesity	Average (1990 - 2003)	Obese population i.e. % of population with Body mass index BMI>=30Kg/m2
	Income	2011	
	Euucauon	2011	

ANNEX 2: DATASET	USED IN THE	e dea an	ALYSIS
------------------	-------------	----------	--------

	Total health expenditure per capita, in PPP	Life expectancy at birth, years	Life expectancy at age 65, years	Healthy life expectancy at birth, years	Healthy life expectancy at age 65, years	Amenable mortality rates per 100 000 population	Physician s per 100 000 population	Nurses per 100 000 population	Beds per 1 000 population	Life expectancy at birth according to Heijink R. et al (2015), years	Healthy life expectancy at birth according to Heijink R. et al (2015), years	Total health expenditure per capita adjusted base on Heijink R. et al (2014), in PPP	Alcohol consumption in litres per capita	Regular smokers, % of population aged 15+ that are daily smokers	Obese population i.e. % of population with Body mass index BMI>=30 Kg/m2	GDP per capita, in PPP	Education	
AT	2,935	81.1	20.1	74.0	15.3	1.6	4.8	7.9	7.6	82.0	76.2	2,965	13.8	27.7	8.8	31,588	77.1	AT
BE	2,780	80.7	19.9	73.8	15.8	1.2	2.9	15.4	6.3	81.8	76.0	2,805	11.0	28.1	11.5	29,520	68.1	BE
BG	1,319	74.2	15.8	67.6	10.4	3.2	3.7	4.7	6.4	75.5	69.7	1,401	10.0	33.7	12.4	11,259	76.0	BG
СҮ	1,619	81.2	19.3	73.6	13.3	1.2	3.0	4.9	3.5	82.5	76.1	1,643	9.2	30.7	12.3	23,133	71.7	СҮ
CZ	2,025	78.0	17.6	69.9	12.1	2.7	3.6	8.5	6.8	79.2	72.8	2,064	11.7	25.4	12.9	20,094	86.1	CZ
DE	3,625	80.8	19.8	74.9	16.2	1.5	3.8	11.6	8.2	82.0	77.3	3,661	11.2	24.5	12.2	30,172	81.6	DE
DK	3,015	79.9	18.8	74.1	16.1	1.4	3.5	15.7	3.5	81.3	76.1	3,053	12.2	35.0	8.6	31,405	69.3	DK
EE	1,315	76.6	18.0	66.8	11.1	3.0	3.3	6.5	5.3	78.3	69.7	1,350	7.3	31.6	13.3	17,218	82.3	EE
EL	2,032	80.8	19.8	74.3	14.2	1.5	6.1	3.3	4.9	82.6	77.0	2,059	10.0	38.2	10.7	20,159	62.9	EL
ES	1,937	82.6	21.0	76.4	16.2	1.1	4.1	5.5	3.1	83.8	79.6	1,964	11.9	31.8	11.3	23,793	53.0	ES
FI	2,461	80.6	19.9	75.4	16.1	1.9	2.7	10.7	5.5	81.5	78.2	2,473	8.8	24.1	10.1	28,560	77.1	FI
FR	3,127	82.3	21.7	75.5	17.1	0.9	3.1	9.0	6.4	83.4	77.2	3,160	14.8	28.1	7.3	27,812	68.9	FR
HR	1,266	77.2	17.0	63.9	8.6	3.2	2.8	5.7	5.8	78.8	67.8	1,302	12.9	27.4	11.0	14,703	74.7	HR
HU	1,978	75.1	16.6	65.3	9.6	3.5	3.0	6.4	7.2	76.8	68.5	2,027	12.8	34.9	18.5	16,433	76.2	HU
IE	1,962	80.9	19.5	78.5	18.1	1.6	2.7	12.6	2.9	81.5	80.6	1,976	12.6	29.3	13.0	31,933	70.3	IE
IT	2,117	82.4	20.5	73.3	13.6	1.3	4.1	6.6	3.4	82.8	74.9	2,137	9.5	25.2	8.4	25,380	54.6	IT
LT	1,335	73.7	17.0	62.8	7.9	4.6	3.7	7.2	7.4	75.6	66.0	1,380	7.1	28.6	16.0	16,413	84.1	LT
LU	3,259	81.1	19.8	74.8	16.1	1.2	2.8	11.6	5.4	81.2	76.5	3,272	14.9	28.3	16.5	63,892	70.9	LU
LV	1,352	73.9	16.6	63.9	9.1	4.2	2.9	4.9	5.9	76.2	67.9	1,406	8.7	32.3	15.5	14,439	80.5	LV
MT	2,218	80.9	19.4	77.1	16.3	1.8	3.1	6.8	4.5	82.2	80.1	2,252	6.4	25.2	23.0	21,524	41.1	MT
NL	3,172	81.3	19.8	77.4	17.9	1.1	3.0	8.6	4.7	82.7	79.5	3,211	9.9	34.1	7.8	31,853	68.3	NL
PL	1,598	76.9	17.9	66.7	10.0	1.9	2.2	5.8	6.5	78.8	70.0	1,652	8.4	34.9	11.4	16,092	82.5	PL
PT	1,748	80.7	19.9	67.8	10.4	1.5	3.0	6.3	3.4	81.3	70.5	1,761	13.7	22.0	12.2	19,500	35.8	РТ
RO	1,169	74.6	16.3	68.4	11.6	3.9	2.4	5.4	6.1	76.0	70.9	1,205	10.8	21.7	8.6	12,742	70.6	RO
SE	2,578	81.9	20.0	78.6	18.1	1.4	3.9	11.1	2.7	82.6	79.5	2,597	6.3	21.3	8.9	30,807	75.6	SE
SI	1,859	80.1	19.3	70.3	12.7	1.6	2.5	8.4	4.6	81.3	73.1	1,888	11.8	25.8	12.3	20,695	80.3	SI
SK	2,232	76.1	16.8	66.1	9.0	3.9	3.0	6.3	6.1	77.4	68.4	2,276	10.5	23.3	16.8	18,777	84.3	SK
UK	2,430	81.0	19.9	77.0	17.9	1.5	2.8	10.3	2.9	82.0	79.4	2,452	10.0	27.4	18.3	26,206	76.2	UK
EU -																		EU - 
median	<b>2,028</b>	80.7	19.5	73.7	13.9	1.6	3.0	7.0	5.5	81.4	76.1	2,062	10.7	28.1	12.2	22,328	75.2	median
Sources: (	Jwn calculat	lons base	a on Euro	ostat, OEC	$\mathcal{D}, \pi e \mu$	nk et al. (	2013).											

## ANNEX 3 Data envelopment analysis (DEA)

#### The method

Production possibility frontiers have been estimated using many different methods. The two principal methods that have been used are data envelopment analysis (DEA) and stochastic frontier analysis (SFA), respectively involving mathematical programming, and econometric methods (Coelli et al., 2005).<sup>44</sup>

DEA involves the use of linear programming methods to construct a non-parametric piece-wise surface (or frontier) enclosing the data. Efficiency measures are then calculated as measures of distance relative to this surface. Farrell (1957) first proposed estimating the production possibility frontier using a piece-wise linear convex hull approach, however it was Charnes et al. (1978) that first proposed a DEA model. First DEA applications used input orientation and assumed constant returns to scale (CRS). The CRS assumption is appropriate when all units are operating at an optimal scale. Various authors (e.g. Färe et al., 1983; Banker et al., 1984) adjusted the (initial) CRS DEA model to account for variable returns to scale (VRS) situations, thereby removing the impact of scale effects.

Graph 12 presents the DEA production possibilities frontier in the simple one input-one output case. Countries A, B and C are efficient. Their output scores (the Shepard measure:  $\frac{d1}{d1+d2}$ ) are equal to 1. Country D is not efficient and its score is smaller than 1.





The analytical description of the DEA linear programming problem to be solved, assuming VRS, is sketched below (Afonso and Aubyn, 2006). Note that problem (1) has to be solved for each one of n decision making units (DMUs) in order to obtain n efficiency scores.

<sup>&</sup>lt;sup>44</sup> This paper uses the statistical programme R, and for the frontier methods (DEA and SFA) packages "Benchmarking", "Frontier" and "Fear".

Suppose there are *k* inputs and *m* outputs for *n* DMUs. For the i-th DMU,  $y_i$  is the column vector of the outputs, and  $x_i$  is the column vector of the inputs. We can also define X as the (k×n) input matrix and Y as the (m×n) output matrix. The output-oriented DEA model is then specified as the following mathematical programming problem, for each i-th DMU (Coelli et al, 2005).<sup>45</sup>

 $\max_{\theta \lambda} \theta$ 

 $st - \theta y_i + Y\lambda \ge 0$  $x_i - X\lambda \ge 0$  $n1'\lambda = 1$ 

 $\lambda \ge 0$ 

where  $1 \le \theta < \infty$ , and  $\theta$ -1 is the proportional increase in outputs that could be achieved by the i-th DMU, with inputs held fixed. Note that  $1/\theta$  defines a technical efficiency score that varies between zero and one, which in this paper will be reported as the output-oriented technical efficiency score (Shephard measure).  $1/\theta$  measures the distance between a DMU and the efficiency frontier, defined as a linear combination of best practice observations. With  $\theta$ >1, the DMU is inside the frontier (i.e. it is inefficient), while  $\theta$ =1 implies that the DMU is on the frontier (i.e. it is efficient).

The vector  $\lambda$  is a (n×1) vector of constants, giving the weights used to compute the location of an inefficient DMU if it were to become efficient. The inefficient DMU would be projected on the production frontier as a linear combination, using those weights of the peers of the inefficient DMU. The peers are other DMUs that are more efficient and therefore are used as references for the inefficient DMU.

n1 is a n-dimensional vector of ones. The restriction  $n1'\lambda = 1$  imposes convexity of the frontier, accounting for variable returns to scale. Dropping this restriction would amount to assuming constant returns to scale.

DEA has a number of advantages, namely it can simultaneously deal with multiple inputs and outputs, and does not require any assumption on the functional form of the production possibilities frontier.

One important decision to take when performing DEA is whether to use input- or output-oriented efficiency measures. An input-oriented measure holds the current level of output constant and minimises inputs, whereas an output-oriented one maximises output keeping the amount of inputs constant. The output- and input-oriented efficiency measures are equivalent measures of technical efficiency only under CRS (Färe and Lovell, 1978). In this paper, we use Shephard's distant measure of output-oriented technical efficiency, which is bounded between zero and one (with one representing relative technical efficiency).

In this paper we always use the VRS model. Hollingsworth and Smith (2003) show that when using ratio data in DEA, which is the case in this paper (e.g. physical or financial variables are presented as per capital ratios), the VRS model is necessary for technical reasons.

#### Variables

The definition of variables and the dataset used in the DEA analysis are presented in Annexes 1 and 2, respectively.

DEA requires using output variables measured in such a way that "more is better". This is not the case with amenable mortality rates. That is why we use instead the inverse of the amenable mortality rate (inv\_amen), defined as:

 $Inv\_amen = \frac{100000}{amen}$ 

where *amen* is the amenable mortality rate.<sup>46</sup>

A well-known problem with DEA analysis is that when there are a large number of inputs and/or outputs relatively to the number of decision units, there tends to be an "inflation" of efficient units. Therefore, in some

(Eq. 1)

(Eq. 2)

<sup>&</sup>lt;sup>45</sup> Page 180.

<sup>&</sup>lt;sup>46</sup> Deaths considered avoidable due to medical intervention. Can be seen as an (inverse) indicator of healthcare quality.

DEA models (results not reported) we used principal component analysis (PCA) to aggregate indicators. Specifically, we applied PCA to the set of input variables, in order to reduce the dimensionality of the input dataset to the first 2 or 3 principal components, which typically represent about 80 percent of the total variation of input variables (Adler and Golany, 2001; Afonso and Aubyn, 2006).<sup>47</sup>

Hospital discharges have been weighted in the following way:

Using the shortlist for the International Classification of Diseases (ICD), containing 20 categories, discharges by diagnostic are aggregated into a weighted measure of total discharges using as weights the length of stay (LOS) by diagnostic category. This aggregation procedure follows the methodology developed by Herr (2008).

First, a Global Length of Stay (GLOS) measure is calculated by adding the duration of all stays in number of days across all discharge categories in all countries and dividing by the number of all discharges.

$$GLOS = \frac{\sum_{i} \sum_{c} Inpat_{i,c}}{\sum_{i} \sum_{c} Disch_{i,c}} = \frac{\sum_{i} \sum_{c} Los_{i,c} * Disch_{i,c}}{\sum_{i} \sum_{c} Disch_{i,c}}$$
Eq. 3

where *Inpat* is inpatients; *Disch* is discharges; and *i* and *c* are the country and ICD indices, respectively.

Second, an index of discharge weights (W) is constructed as the ratio of length of stay in a particular disease category (LOS) divided by GLOS:

$$\mathbf{W}_{\mathbf{c}} = \frac{\mathbf{Los}_{\mathbf{c}}}{\mathbf{GLOS}} = \frac{\frac{\sum_{i} \mathbf{Los}_{i,c} \cdot \mathbf{Disch}_{i,c}}{\sum_{i} \mathbf{Disch}_{i,c}}}{\mathbf{GLOS}} \mathbf{Eq. 4}$$

The final measure of aggregated discharges adjusted (Agg\_disch\_adj) by case severity is obtained by multiplying the discharges by categories by the discharge weights and summing across categories.

#### $Agg_disch_adj_i = \sum_c Disch_{i,c} * W_c$

The variable Agg\_disch\_adj<sub>i</sub> is then used as a measure of hospital output in some DEA models.

Strategy to estimate DEA scores

It is common practice to estimate DEA indexes of health care efficiency using a two-stage procedure. In the first stage, we determine output efficiency scores for each country, using Simar and Wilson's (1998) bootstrap method for the estimation of confidence intervals corrected for small sample bias.<sup>48</sup>

Bootstrapping consists in carrying out repeated simulations of the data generation process (DGP), creating multiple datasets, from which the sampling distribution of the original estimator can be calculated. The bootstrap procedure allows deriving the distribution of efficiency scores, the calculation of confidence intervals, and the correction of estimation bias.

In a second stage, bias-corrected DEA estimates are explained using regression analysis. Non-discretionary factors or "environmental variables" (i.e. outside the control of healthcare systems' decision makers) are used as explanatory factors. Simar and Wilson (2007) propose two alternative bootstrap methods. The efficiency scores obtained depend linearly on environmental variables, but the error term is a truncated, and not a censored, normal random variable. In empirical work, Afonso and Aubyn (2006) find that the estimated coefficients obtained applying bootstrap procedures (in the second stage) are close to the estimates derived from the more simple Tobit procedure.<sup>49</sup>

For simplicity, in this paper we adopted a one-stage approach. Specifically, we use the adjusted (healthy) life expectancy at birth for the impact of lifecycle variables based on Heijink R. et al. (2015), which used a micro-simulation model based on individual data.<sup>50</sup>

Eq. 5

 <sup>&</sup>lt;sup>47</sup> Use of PCA scores in DEA is relatively common in the empirical literature, being justified by a number of "translation invariance" results (Pastor, 1996). In VRS DEA models, output-oriented scores are invariant to input reduction.
 <sup>48</sup> The R package "Fear" is used.

<sup>&</sup>lt;sup>49</sup> Even if Tobit results are possibly biased, it is not clear that bootstrap estimates are more reliable. In fact, the latter are based on a set of assumptions concerning the data generation process and the perturbation term that may be questioned.
<sup>50</sup> http://ec.europa.eu/eahc/documents/health/tenders/2013/EN/EAHC\_2013\_05\_Specifications.pdf

## ANNEX 4 Some auxiliary diagnosis elements for DEA

#### Isoquant graphs

An alternative way of presenting efficiency scores is two draw isoquants. Production functions can be drawn for the one input-one output cases. Isoquants can be drawn for two inputs-one output cases. Graph 12 and Graph 13 present examples of isoquants, respectively, for a model with *life expectancy* as output and *beds* and *physicians* as inputs, and *life expectancy* as output and *the composite indicator* and *per capita health expenditure in PPP* as inputs.

In Graph 12 and Graph 13, the outer curve (black solid line) corresponds to the production possibilities frontier. The intermediate curve (green dashed line) corresponds to the bias corrected production possibilities frontier. The inner curve (red dotted line) corresponds to the lower confidence interval of the production possibilities frontier. For a selected number of countries, <sup>51</sup> not to clutter too much the graph, data points are connected to the origin, giving a "quick impression" of the respective efficiency scores.



Graph 12 – Isoquant of a DEA

<sup>&</sup>lt;sup>51</sup> DE, FR, IT, UK, ES, PL.





Notes: Own calculations

Table 6 reports results for the Kolmogorov-Smirnov (KS) test of the null hypothesis that two series have been drawn from the same (continuous) distributions. As an example, it applies this test to a few DEA models. The KS test cannot reject the null hypothesis that models A and B are drawn from the same distribution, but models A, B and C have not been drawn from the same distribution as model D. Apparently, changing the output variable from life expectancy to healthy life expectancy results in DEA scores that do not appear having come from the same distribution.

#### Table 4 – Kolmogorov-Smirnov tests (p-values)

	А	В	С	D	E	F								
А														
В	0.938													
С	C 0.005 0.056													
D	D 0.005 0.001 0.000													
Е	E 0.203 0.056 0.012 0.541													
F	0.027	0.203	0.203	0.012	0.346									
Signif. Code	s: 0.01 'blue'	' 0.05 'yellov	v'											
H0: Have the	ose sets of e	fficiency sco	res been dra	wn from the	e same conti	nuous								
distribution	?													

Source: Own calculations.

*Notes*: DEA models:

A) outputs: life expectancy; inputs: pc health exp. In PPP

B) outputs: life expectancy; inputs: pc health exp. In PPP & total composite indicator (TCI)

C) outputs: life expectancy; inputs: beds & physicians & nurses & TCI

D) outputs: healthy life expectancy; inputs: pc health exp. In PPP

E) outputs: healthy life expectancy; inputs: pc health exp. In PPP & TCI

F) outputs: healthy life expectancy; inputs: beds & physicians & nurses & TCI

	Life exp	years	at birth,	Life ex	pectancy 65, years	v at age	Healthy at	y life exp birth, ye:	ectancy ars	Healthy life expectancy at age 65, years			Amen	able mor	tality	Life exp accordi et al	pectancy ing to Hei (2015), y	at birth ijink R. ⁄ears	at birth jink R. ears Healthy life expectan at birth according t Heijink R. et al (201 years		ectancy ing to (2015),		Intern	ermediate outputs			
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	
Country																											
AT	0.98	0.98	0.98	0.94	0.94	0.95	0.92	0.92	0.92	0.83	0.83	0.84	0.50	0.51	0.53	0.98	0.98	0.98	0.94	0.94	0.94	0.84	0.85	0.54	0.78	0.82	AT
BE	0.97	0.97	0.98	0.93	0.93	0.93	0.91	0.91	0.92	0.85	0.85	0.84	0.68	0.69	0.66	0.97	0.97	0.98	0.94	0.94	0.93	0.72	0.42	0.58	0.59	0.58	BE
BG	0.95	0.98	0.98	0.94	0.95	0.95	0.85	0.94	0.94	0.77	0.81	0.82	0.59	0.45	0.38	0.93	0.98	0.98	0.94	0.95	0.95	0.49	0.65	NA 0.10	NA 0.21	0.73	BG
C7	0.99	0.99	0.98	0.98	0.98	0.93	0.97	0.97	0.94	0.64	0.63	0.85	0.77	0.75	0.39	0.99	0.99	0.98	0.98	0.98	0.96	0.21	0.19	0.19	0.21	0.74	C7
DF	0.94	0.94	0.94	0.88	0.88	0.95	0.82	0.90	0.80	0.87	0.88	0.87	0.50	0.52	0.28	0.94	0.94	0.94	0.95	0.95	0.95	0.40	0.38	0.75	0.78	0.89	DF
DK	0.97	0.96	0.97	0.94	0.94	0.93	0.86	0.85	0.89	0.87	0.86	0.86	0.55	0.52	0.69	0.97	0.97	0.97	0.94	0.94	0.93	0.54	0.30	0.36	0.38	0.67	DK
EE	0.98	0.98	0.93	0.93	0.93	0.85	0.97	0.96	0.86	0.82	0.82	0.65	0.65	0.64	0.30	0.96	0.96	0.94	0.94	0.94	0.86	0.42	0.54	0.65	0.73	0.75	EE
EL	0.97	0.98	0.98	0.93	0.96	0.95	0.92	0.93	0.94	0.75	0.84	0.82	0.64	0.62	0.38	0.98	0.98	0.98	0.94	0.95	0.96	0.39	0.54	0.29	0.50	0.73	EL
ES	0.99	0.99	0.98	0.96	0.97	0.95	0.97	0.96	0.94	0.86	0.90	0.82	0.90	0.72	0.37	0.99	0.98	0.98	0.97	0.97	0.96	0.40	0.30	0.54	0.52	0.74	ES
FI	0.97	0.97	0.99	0.95	0.95	0.96	0.92	0.93	0.96	0.87	0.87	0.88	0.48	0.48	0.55	0.97	0.97	0.99	0.96	0.96	0.97	0.59	0.82	0.32	0.77	0.85	FI
FR	0.99	0.99	0.99	0.95	0.95	0.96	0.98	0.98	0.96	0.92	0.91	0.92	0.87	0.83	0.73	0.99	0.99	0.99	0.95	0.95	0.95	0.65	0.34	0.53	0.54	0.50	FR
HR	0.98	0.98	0.97	0.89	0.95	0.89	0.94	0.93	0.88	0.63	0.81	0.69	0.66	0.49	0.40	0.98	0.98	0.99	0.94	0.95	0.93	0.58	0.56	NA	NA	0.82	HR
HU	0.90	0.98	0.98	0.82	0.95	0.95	0.77	0.93	0.94	0.50	0.81	0.82	0.28	0.47	0.38	0.91	0.98	0.98	0.84	0.95	0.95	0.50	0.55	0.83	0.78	0.74	HU
IE	0.97	0.97	0.98	0.98	0.98	0.95	0.91	0.91	0.94	0.94	0.92	0.82	0.61	0.59	0.38	0.97	0.96	0.98	0.98	0.98	0.96	0.79	0.34	0.27	0.39	0.73	IE
IT	0.99	0.99	0.99	0.92	0.92	0.93	0.96	0.96	0.95	0.72	0.72	0.77	0.76	0.76	0.69	0.98	0.98	0.98	0.92	0.92	0.92	0.44	0.43	NA	NA	0.83	IT
	0.94	0.94	0.89	0.88	0.88	0.80	0.92	0.91	0.78	0.58	0.58	0.44	0.40	0.39	0.19	0.93	0.93	0.90	0.89	0.89	0.82	0.48	0.78	0.71	0.69	0.81	
	0.98	0.98	1.00	0.93	0.95	0.94	0.90	0.90	0.93	0.87	0.88	0.85	0.62	0.65	0.85	0.97	0.97	0.98	0.94	0.94	0.95	0.51	0.58	0.52	0.54	0.49	
MT	0.94	0.94	0.98	0.89	0.90	0.95	0.89	0.91	0.94	0.00	0.73	0.85	0.41	0.40	0.55	0.93	0.93	0.98	0.91	0.93	0.95	0.50	0.37	0.38	0.01	0.74	MT
NI	0.98	0.98	0.99	0.98	0.98	0.98	0.90	0.90	0.94	0.97	0.97	0.93	0.68	0.69	0.84	0.98	0.98	1.00	0.98	0.98	0.97	0.61	0.23	0.52	0.48	0.39	NI
PL	0.94	0.94	0.98	0.89	0.90	0.95	0.91	0.90	0.94	0.64	0.66	0.82	0.50	0.55	0.38	0.95	0.95	0.98	0.90	0.91	0.95	0.45	0.53	0.64	0.63	0.72	PL
PT	0.98	0.98	0.98	0.88	0.95	0.95	0.97	0.94	0.94	0.61	0.78	0.82	0.66	0.59	0.35	0.97	0.98	0.98	0.89	0.96	0.95	0.56	0.23	0.35	0.35	0.74	РТ
RO	0.97	0.98	0.98	0.93	0.95	0.95	0.92	0.93	0.94	0.72	0.81	0.83	0.32	0.47	0.36	0.97	0.98	0.98	0.94	0.95	0.95	0.23	-0.08	-0.05	0.33	0.73	RO
SE	0.99	0.99	0.98	0.99	0.99	0.95	0.92	0.93	0.94	0.97	0.98	0.83	0.61	0.62	0.39	0.98	0.98	0.98	0.98	0.98	0.95	0.41	0.37	0.23	0.35	0.73	SE
SI	0.97	0.97	0.98	0.90	0.89	0.95	0.91	0.92	0.94	0.70	0.69	0.83	0.61	0.61	0.36	0.97	0.97	0.98	0.91	0.90	0.95	0.45	0.54	0.50	0.63	0.73	SI
SK	0.92	0.92	0.93	0.83	0.83	0.86	0.78	0.78	0.82	0.48	0.48	0.56	0.24	0.24	0.23	0.92	0.92	0.93	0.84	0.84	0.86	0.44	0.57	0.82	0.81	0.71	SK
UK	0.98	0.98	0.98	0.97	0.97	0.95	0.92	0.92	0.94	0.96	0.95	0.83	0.60	0.59	0.39	0.97	0.97	0.98	0.98	0.97	0.95	0.69	0.39	0.38	0.47	0.74	υκ
EU	0.98	0.98	0.98	0.94	0.95	0.95	0.92	0.92	0.94	0.82	0.83	0.83	0.61	0.57	0.39	0.97	0.98	0.98	0.94	0.95	0.95	0.50	0.48	0.52	0.54	0.74	EU

## ANNEX 5 DEA efficiency scores for all models

#### Source: Own calculations.

Notes: Green: above the 75<sup>th</sup> percentile; Red: below the 25<sup>th</sup> percentile; White: Inter Quartile Range.

Output-oriented and bias corrected Shephard DEA efficiency scores: Models 1 to 3 as defined in **Table 1**;Model 4: Output: Hospital discharges; Inputs: total health expenditure per capita in PPP; Model 5: Output: Outpatient consultations; Inputs: total health expenditure per capita in PPP; Model 6: Output: Hospital discharges weighted by length of stay as described in Annex 2; Model 7: Output: Hospital discharges weighted by length of stay as described in Annex 2; Model 7: Output: Hospital discharges weighted by length of stay as described in Annex 2; Inputs: Indicators of physical inputs (hospital beds, nurses, physicians per capita) and a composite indicator of the socio-economic environment (GDP per capita, educational attainment) and lifestyle factors (lagged consumption of alcohol and tobacco, obesity) as in model 3.

		Life ex	pectancy a years	t birth,	Life exp	ectancy a years	t age 65,	Healthy b	life expec irth, year	ctancy at 's	Healthy aş	life expec ge 65, year	tancy at rs	Ame	nable mor	tality	Life ex according (2	pectancy a g to Heijin 2014), year	at birth 1k R. et al rs	Healthy birth acco et al	v life expe ording to l (2014), y	ctancy at Heijink R. zears		Inter	mediate o	utputs	
		Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Life expectance at birth	Model 1	1.00																									
vears	Model 2	0.86	1.00																								
	Model 3	0.37	0.42	1.00	)																					<u> </u>	
Life expectancy at age 65.	Model 1	0.60	0.47	0.34	1.00																				<u> </u>	└──	
years	Model 2	0.57	0.57	0.40	0.83	1.00																			<u> </u>	<u> </u>	_
	Model 3	0.16	0.29	0.62	0.43	0.60	1.00	1.00																	<b> </b> '	──	
Healthy life expectancy at	Model 1	0.74	0.58	0.11	0.29	0.34	0.09	1.00	1.00																	<u> </u>	
birth, years	Model 3	0.31	0.05	0.20	0.13	0.31	0.27	0.77	1.00	1.00															<u> </u>	<u> </u>	-
	Model 1	0.51	0.30	0.74	0.31	0.30	0.37	0.20	-0.04	0.28	1.00														<u> </u>	<u> </u>	
Healthy life expectancy at	Model 2	0.56	0.45	0.42	0.89	0.84	0.35	0.21	0.05	0.33	0.94	1.00															-
Healthy life expectancy at age 65, years	Model 3	0.34	0.28	0.51	0.63	0.45	0.51	-0.03	-0.13	0.29	0.71	0.67	1.00														
	Model 1	0.80	0.64	0.40	0.43	0.42	0.17	0.66	0.52	0.28	0.39	0.39	0.15	1.00													
Amenable mortality	Model 2	0.78	0.65	0.51	0.51	0.43	0.22	0.62	0.43	0.43	0.52	0.51	0.28	0.89	1.00												
	Model 3	0.55	0.45	0.61	0.56	0.43	0.26	0.06	-0.04	0.35	0.66	0.61	0.74	0.42	0.48	1.00											
Life expectancy at birth	Model 1	0.89	0.73	0.34	0.60	0.64	0.30	0.69	0.42	0.22	0.55	0.58	0.44	0.72	0.72	0.55	1.00										
according to Heijink R. et	Model 2	0.75	0.84	0.47	0.49	0.71	0.56	0.54	0.60	0.32	0.38	0.49	0.36	0.64	0.61	0.45	0.84	1.00									
al (2014), years	Model 3	0.39	0.34	0.67	0.43	0.58	0.59	0.16	0.17	0.49	0.35	0.46	0.48	0.48	0.40	0.57	0.42	0.48	1.00						<u> </u>	└──	
Healthy life expectancy at	Model 1	0.59	0.43	0.34	0.97	0.85	0.43	0.33	0.14	0.27	0.88	0.88	0.59	0.41	0.50	0.52	0.63	0.49	0.45	1.00					<u> </u>	<u> </u>	
birth according to Heijink <b>R</b> at al (2014) years	Model 2	0.49	0.52	0.39	0.78	0.95	0.62	0.31	0.32	0.37	0.63	0.78	0.41	0.37	0.37	0.31	0.52	0.64	0.51	0.81	1.00	0				──	
R. et al (2014), years	Model 3	0.14	0.15	0.47	0.54	0.68	0.83	0.13	0.11	0.48	0.39	0.49	0.44	0.16	0.24	0.19	0.33	0.45	0.57	0.60	0.74	1.00	1.00		<u> </u>	┝───	-
	Model 5	0.07	-0.06	0.27	0.18	0.12	0.12	-0.18	-0.28	0.10	0.30	0.30	0.44	0.01	-0.04	0.50	0.04	-0.04	0.28	0.14	0.11	0.02	0.15	1.00	<u> </u>		
Intermediate outputs	Model 6	-0.43	-0.45	-0.33	-0.45	-0.48	-0.23	-0.33	-0.24	-0.31	-0.30	-0.44	-0.22	-0.33	-0.00	-0.20	-0.43	-0.41	-0.30	-0.44	-0.47	-0.30	0.13	1.00	1.00	<u> </u>	
and include outputs	Model 7	-0.44	-0.43	-0.35	-0.58	-0.65	-0.32	-0.39	-0.33	-0.33	-0.48	-0.49	-0.22	-0.47	-0.34	-0.28	-0.50	-0.49	-0.38	-0.56	-0.64	-0.48	0.11	0.00	0.85	1.00	5
N N	Model 8	0.08	0.08	-0.12	-0.03	0.06	0.00	0.35	0.36	-0.05	-0.15	-0.08	-0.14	-0.13	-0.18	-0.09	0.08	0.13	-0.15	0.07	0.14	0.06	-0.03	0.42	0.00	0.18	3 1

## ANNEX 6 Spearman rank order correlations for DEA models

Source: Own calculations.

Notes: Based on bias corrected Shephard DEA output-oriented efficiency scores, and models described in Table 1 and Annex 5.

## ANNEX 7 Stochastic frontier analysis (SFA)

There is a wide variety of methods to assess technical efficiency.<sup>52</sup> Available methods are usually divided between non-parametric (e.g. DEA) and parametric (e.g. SFA) ones. Non-parametric methods have been criticised for lack of a statistical base. In addition, any deviation between observed data and the estimated production possibilities frontier is attributed to inefficiency. However, due to the work of Simar and Wilson (1998, 2000 and 2007), the underlying data generation process has been explored to analyse the sensitivity of estimated efficiency scores to sampling variation.

Stochastic frontier analysis (SFA) is the other major strand of the literature regarding the estimation of efficiency measures of production. SFA estimates the production possibilities frontier assuming a given functional form and decomposing the error term into two components. One part represents random events outside the control of the decision making unit, while the other is a non-negative term capturing inefficiency.

#### 11.1. Panel data models

As regards our basic SFA panel data model estimations, we assume time-varying efficiency (Table 7, models 1, 2 and 3), and use the error components frontier model (ECF, Balttese and Coelli, 1992). Model 4 in Table 7 assumes time invariant efficiency scores. All four models in Table 7 assume that inefficiency has a truncated normal distribution.

Specifically, we assume the following Cobb-Douglas production function:<sup>53</sup>

$$lny_{i,t} = \alpha + lnx_{i,t} * \beta + v_{i,t} - u_{i,t}$$

Eq. 6

where  $y_{i,t}$  is life-expectancy (as a proxy of output of the healthcare system) in country *i* and year *t*;  $\beta$  is a vector of unknown parameters;  $x_{i,t}$  are the input variables;  $v_{i,t}$  represent stochastic shocks; and  $u_{i,t} \ge 0$  is a non-negative stochastic variable associated with technical inefficiency.  $v_{i,t}$  are assumed to be normally iid variables i.e.  $v_{i,t} \sim iidN(0, \sigma_{v_i}^2)$ , while  $u_{i,t}$  are assumed to be truncated normally iid variables i.e.  $u_{i,t} \sim iidN^+(\mu, \sigma_u^2)$ .

Of the four ECF SFA models presented in Table 7, only model 4 is statistical significant, which assumes time invariant efficiency scores. For the other three models presented in Table 7, a constant is the best estimate of inefficiency. In addition, likelihood ratio tests reject the null hypothesis of equality of the ECF SFA models presented in Table 7 with the corresponding OLS models (with no inefficiency).

<sup>&</sup>lt;sup>52</sup> Recall that technical efficiency does not necessary imply economic efficiency, because the latter also requires allocative efficiency (Coelli et al., 2005).

<sup>&</sup>lt;sup>53</sup> Using Monte Carlo simulations, Ruggiero (1999) shows that a misspecified translog function performs rather poorly despite its flexibility if the sample size is small. In their empirical work, Varabyova and Schreyögg (2013) report that a translog function does not fit their data, whereas in contrast a Cobb-Douglas model provides an excellent fit.

	Peg 1			Í				Beg 2			
	E atime at a		5. <u> </u>	D-(-)			Fatter at a		5. <u> </u>	D-(-)	
	Estimate	Sta. Error	z value	Pr(> z )			Estimate	Std. Error	z value	Pr(>[z])	
(Intercept)	4.717	0.068	69.661	0.000	***	(Intercept)	4.608	0.053	86.642	0.000	***
log(pc HC in PPP)	-0.033	0.008	-4.065	0.000	***	log(pc HC in PPP)	-0.010	0.006	-1.540	0.124	
sigmaSq	0.001	0.000	5.888	0.000	***	log(CI_Total)	0.113	0.017	6.470	0.000	***
gamma	0.979	0.004	230.321	0.000	***	sigmaSq	0.002	0.000	8.483	0.000	***
mu	0.076	0.015	5.247	0.000	***	gamma	0.987	0.002	440.354	0.000	***
time	0.041	0.002	21.382	0.000	***	mu	0.084	0.013	6.652	0.000	***
						time	0.024	0.002	10.525	0.000	***
	Reg. 3						Reg. 4				
	Estimate	Std. Error	z value	Pr(> z )			Estimate	Std. Error	z value	Pr(> z )	
(Intercept)	4.562	0.994	4.591	0.000	***	(Intercept)	3.894	0.037	104.453	0.000	***
log(Beds)	-0.060	0.681	-0.089	0.929		log(pc HC in PPP)	0.067	0.005	13.197	0.000	***
log(Physicians)	0.050	0.762	0.066	0.948		sigmaSq	0.001	0.001	2.340	0.019	*
log(CI_Total)	0.024	0.998	0.025	0.980		gamma	0.936	0.028	33.739	0.000	***
sigmaSq	0.004	0.986	0.004	0.997		mu	0.042	0.012	3.535	0.000	***
gamma	0.956	1.000	0.956	0.339							
mu	0.000	1.000	0.000	1.000							
time	0.000	1.000	0.000	1.000							
Signif. Codes: 0 `***' 0.01 `**' 0.05 `*'											
Dependent variable: life expectancy											

Table 5: SFA, Error components frontier (ECF) model

pc HC in PPP: per capita health care expenditure in PPP; CI\_Total: a composite indicator, including GDP per capita, education attainment, obsesity, tobacco and alcohol consumption.

Notes: Commission services, own calculations

#### **11.2.** Accounting for the production environment

We also estimate a SFA panel data model assuming time-varying efficiency, and using the efficiency effects frontier model (EFF, Balttese and Coelli, 1995).

 $lny_{i,t} = \alpha + lnx_{i,t} * \beta + lnz_{i,t} * \gamma + \nu_{i,t} - u_{i,t}$ 

Eq. 7

where  $\mathbf{z}_{i,t}$  is a vector of environmental variables and  $\boldsymbol{\gamma}$  is a vector of unknown parameters. This model has exactly the same error structure as the model described in equation 6.

		Reg. 5							
		Estimate	Std. Error	z value	Pr(> z )				
	(Intercept)	4.274	0.036	119.433	0.000	***			
log(pc HC in PPP)		0.015	0.005	3.162	0.002	***			
	Z_log(GDP_pc)	-0.136	0.015	-9.229	0.000	***			
	Z_log(EDU)	0.331	0.036	9.209	0.000	***			
	Z_log(old_ratio)	0.037	0.021	1.792	0.073				
	sigmaSq	0.001	0.000	6.838	0.000	***			
	gamma	0.679	0.071	9.557	0.000	***			
	Signif. Codes: 0 `***' 0.01 `**' 0.05 `*'								
	Dependent variable: life expectancy.								
	pc HC in PPP: per capita health care expenditure in PPP; EDU:								
	education attainment (Isced 3&6); GDP_pc: per capita GDP in								
	PPP; old_ratio: ratio of old age population (65+) over total								
	population (15_74).								
N	Notes: Commission semijase, own calculations								

Table 6: Efficient effects frontier (EEF) model

Notes: Commission services, own calculations

The single EEF SFA model presented in Table 8 is overall significant. In addition, likelihood ratio tests reject the null hypothesis of equality of the EEF SFA model presented in Table 8 with the corresponding OLS model (with no inefficiency).

#### 11.1. Spearman rank correlations between DEA and SFA efficiency scores

Table 9 presents Spearman rank correlations between DEA and SFA scores.<sup>54</sup> Rank correlations are relatively high. Of particular interest are the rank correlations between model SFA5 (SFA-Efficiency effects frontier) which is the only significant SFA model with time-varying efficiency and the three DEA models reported in Table 9 (in bold).<sup>55</sup>

	DEA1	DEA2	DEA3	SFA1	SFA2	SFA3	SFA5	
DEA1	1.00							
DEA2	0.94	1.00						
DEA3	0.59	0.55	1.00					
SFA1	0.68	0.62	0.77	1.00				
SFA2	0.61	0.55	0.70	0.77				
SFA3	0.54	0.50	0.90	0.78	0.70	1.00		
SFA5	0.65	0.59	0.74	0.76	0.94	0.69	1.00	
DEA1:	DEA1: output=life expectancy; input= per capita health care expenditure in PPP.							
DEA2: output=life expectancy; inputs=per capita health care expenditure in PPP & total composite indicator.								
DEA3: output=life expectancy; inputs= beds & physicians & composite indicator.								

Table 7: Spearman correlations between DEA vs. SFA scores

SFA1, SFA2, SFA3 and SFA5 correspond to the SFA regressions 1, 2, 3 and 5 in Tables 8

and 9.

Notes: Commission services, own calculations

<sup>&</sup>lt;sup>54</sup> DEA scores were obtained for the period 2003 to 2010 by solving for each year a DEA programme.

<sup>&</sup>lt;sup>55</sup> Empirical results in the literature (e.g. Hjalmarsson et al., 1996; and De Cos and Moral-Benito, 2011) suggest that correlations are stronger within each method (i.e. DEA vs. SFA) than across methods.

## ANNEX 8 Potential economic savings of improving efficiency

This section follows De Cos and Moral-Benito (2011) in providing country-specific estimates of potential economic gains of improving efficiency. Saving estimates explore the following counterfactual: by how much could health care expenditure decrease if a country adopted the most efficiency system, while keeping the same health outcome?

The following panel regression is estimated:

$$O_{i,t} = \alpha_1 * EFF_{i,t} + \alpha_2 * H_{pc_{i,t}} + \gamma_1 * \mathbf{Z}_{i,t} + \varepsilon_{i,t}$$

where *O* is life expectancy of country *i* in year *t*; *EFF* is the efficiency level in the production of health (either a DEA or a SFA score);  $H_{pc}$  is per capita health care expenditure; and *Z* is a matrix with other socioeconomic characteristics that can affect life expectancy. Equation 8 was estimated with variables either in linear or logarithmic form and using either DEA or SFA estimates.<sup>56</sup>

	Estimate	Std. Error	z value	Pr(> z )				
(Intercept)	4.112	0.038	107.847	0.000	***			
log(DEA11)	0.993	0.053	18.595	0.000	***			
log(pc HC in PPP)	0.070	0.006	11.857	0.000	***			
log(EDU)	-0.012	0.004	-3.344	0.001	***			
log(pc GDP in PPP)	-0.023	0.008	-2.810	0.005	**			
log(old_ratio)	-0.009	0.007	-1.404	0.162				
Signif. Codes: 0 `***' 0.01 `**' 0.05 `*'								
Dependent variable: (log of) life expectancy.								
DEA11: output=life expectancy; input= per capita health care								
expenditure in PPP; pc HC in PPP: per capita health care expenditure in								
PPP; EDU: education attainment (Isced 3&6); pc GDP in PPP: per capita								
GDP in PPP; old_ratio: ratio of old age population (65+) over total								
population (15_74).								
PPP for GDP are used (instead of PPP for health).								

Table 8: Panel regression estimates (2013-2010) of life expectancy

Notes: Commission services, own calculations

**Table 10** presents results for one regression using life expectancy as the outcome variable of health, and DEA for the efficiency level in production. The estimates for efficiency and health care expenditure are significant and correctly signed. As expected, an increase in efficiency, or in (per capita) expenditure, raises life expectancy.

Potential savings in the health care expenditure-to-GDP ratio  $(sav_{gdp})$  are calculated as follows. Using the estimated parameters of equation 8, we calculate the change in the expenditure-to-GDP ratio compatible with maintaining the health outcome  $\hat{O}$ , measured by life expectancy, while adopting the maximum efficiency level over the panel (EFF<sup>max</sup>).

In the linear and logarithmic models, GDP savings are calculated, respectively as:

$$sav_{gdp} \equiv \frac{HE_{pc} - HE_{pc}^{min}}{GDP_{pc}} = \frac{\frac{\alpha_1}{\alpha_2} * (EFF^{max} - EFF)}{GDP_{pc}}$$
 Eq. 9

or

Eq. 8

<sup>&</sup>lt;sup>56</sup> However only results using DEA scores and the logarithmic specification are reported.

$$sav_{gdp} \equiv \frac{HE_{pc} - HE_{pc}^{min}}{GDP_{pc}} = \left[1 - \frac{1}{\left(\frac{EFF^{max}}{EFF}\right)^{\frac{\alpha_1}{\alpha_2}}}\right] \frac{HE_{pc}}{GDP_{pc}}$$
Eq. 10

where  $\alpha_1$  and  $\alpha_2$  are the estimated coefficients in equation 8 for the efficiency and expenditure variables, respectively.

Although varying significantly across country, on average around 1/4 of total health expenditure could be saved: equivalent to (non-weighted) average savings of 1.5% of GDP.<sup>57</sup>

<sup>&</sup>lt;sup>57</sup> Using this methodology, Cos and Moral-Benito (2011) estimate efficiency savings of 2.6% of GDP for the OECD.

## ANNEX 9 Estimating Malmquist indexes

In this Annex, we calculate Malmquist total factor productivity (TFP) indexes to estimate changes in productivity growth in the health sector, breaking it down between technical efficiency (catching-up) and technological progress (innovation). We use DEA scores to produce Malmquist indexes, and report cumulated changes in productivity over a period (and not annual changes).

In the context of estimating total factor productivity growth in the health sector, Malmquist indexes have a number of advantages, because they can handle multiple inputs and outputs, do not require information on prices of outputs or inputs,<sup>58</sup> and do not presume that producers behave either as cost minimisers or profit maximisers.

Given the close links between DEA and Malmquist productivity indexes, the latter will be calculated using DEA indicators. For illustrative purposes, we calculate two sets of Malmquist productivity indexes using the DEA models: i) (output: life expectancy; inputs: per capita health care in PPP and a total composite index); and ii) (output: life expectancy; inputs: physicians, beds and a total composite index).<sup>59</sup>

Tables 9 and 10 present Malmquist TFP indexes and their breakdown between changes in technical efficiency and technological progress in the cumulated period 2003-2010. Results for the TFP index and its breakdown in efficiency and technological changes depend significantly on the model chosen to calculate the production frontier. In addition, data limitations allow us to make calculations only for a few years, basically the first decade of the 21<sup>st</sup> century.

Results tentatively suggest stagnation (or even deterioration) in productivity in the first decade of the 21<sup>st</sup> century, mainly driven by worsening technology conditions, which were not sufficiently compensated by efficiency improvements.

<sup>&</sup>lt;sup>58</sup> This is particularly important in the health sector, because prices do not necessarily reflect opportunity costs, but rather views concerning access to care and the impact of numerous institutional settings. <sup>59</sup> GDP PPP are used.

	Prod	Eff	Tech				
	(1)=(2)*(3)	(2)	(3)				
AT	0.982	1.147	0.856				
BE	0.955	1.112	0.858				
СҮ	0.924	1.019	0.906				
CZ	0.911	1.106	0.823				
DE	0.974	1.154	0.844				
DK	1.070	1.245	0.860				
EE	0.789	1.168	0.675				
EL	0.919	1.090	0.843				
ES	0.937	1.107	0.846				
FI	0.939	1.042	0.901				
FR	0.981	1.134	0.865				
HR	0.871	1.102	0.790				
HU	0.967	1.061	0.912				
IE	0.944	1.109	0.852				
IT	0.962	1.113	0.864				
LT	0.788	1.133	0.696				
MT	0.831	0.995	0.835				
NL	0.984	1.149	0.856				
PL	0.779	1.095	0.711				
РТ	0.920	1.067	0.862				
SE	0.984	1.102	0.893				
SI	0.926	1.031	0.899				
SK	0.666	0.884	0.753				
UK	0.916	1.070	0.856				
Geo. Avg. a)	0.909	1.091	0.833				
Sources: Owi	n calculations						
Notes: Derive	ed using a DE	A model with	life				
expectancy as output; and per capita health care in							
PPP and a total composite indicator as inputs.							
Prod: Malmquist total factor productivity index; Eff:							
efficiency changes; Tech: technological changes.							
Values greater (less) than one signal improvement							
(deterioration).							
a) Geometric	a) Geometric average						

 Table 9: Cumulated productivity, efficiency, and technological changes: 2003-2010

	Prod		Tech		
	(1)=(2)*(3)	(2)	(3)		
AT	0.934	1.008	0.926		
BE	0.988	1.062	0.931		
CY	0.993	1.005	0.989		
CZ	1.004	1.080	0.930		
DE	0.948	1.020	0.930		
EE	1.041	1.127	0.923		
ES	1.075	1.000	1.075		
FR	1.022	1.100	0.930		
HR	0.924	0.993	0.931		
HU	1.020	1.102	0.926		
LT	0.985	1.059	0.930		
LV	0.988	1.065	0.928		
PL	1.091	1.142	0.956		
PT	0.990	1.000	0.990		
RO	0.880	0.909	0.968		
SE	1.107	1.000	1.107		
SI	0.976	1.062	0.920		
UK	1.018	1.000	1.018		
Geo. Avg. a)	0.998	1.039	0.960		
Sources: Own calculations.					

Table 10: Cumulated productivity, efficiency, and technological change: 2003-2010

Derived using a DEA model with life expectancy

as output; and beds, physicians and a total

Prod: Malmquist total factor productivity index;

Eff: efficiency changes; Tech: technological

Values greater (less) than one signal

improvement (deterioration).

a) Geometric average

#### Some technical aspects



Graph 14 - The output distance function and the production possibility set

Notes: Reproduced from Coelli et al. (2005)

As in DEA, Malmquist total factor productivity (TFP) indexes use the concept of distance function.<sup>60</sup> The output distant function  $(d_0)$  defined on the output set P(x) is:

$$\mathbf{d}_{\mathbf{o}}(\mathbf{X},\mathbf{Y}) = \min\left\{\delta: \left(\frac{\mathbf{Y}}{\delta}\right) \in \mathbf{P}(\mathbf{X})\right\}$$
(Eq. 11)

where X and Y are respectively the input and output vectors.

As an example, Graph 14 draws the production possibilities frontier (PPF) for a firm producing two outputs  $(y_1$ and  $y_2$ ) and using an input vector X. The production possibility set, P(X), is the area bounded by the PPF and the axes. The distance value for firm A is given by the ratio:  $\delta = \frac{OA}{OB}$ . It is the reciprocal of the factor by which the production of all output quantities could be increased while still remaining within the feasible production possible set.

Using DEA terminology,  $\delta$  corresponds to the output-oriented Shephard measure of technical efficiency (Coelli et al, 2005).<sup>61</sup>

The output-oriented Malmquist TFP index  $(m_o)$  is the geometric average of distance ratios, using the technologies of two periods (s and t):

$$m_o(y_s, y_t, x_s, x_t) = \left[\frac{d_0^s(x_t, y_t)}{d_0^s(x_s, y_s)} * \frac{d_0^t(x_t, y_t)}{d_0^t(x_s, y_s)}\right]^{0.5}$$
Eq. 12

where  $d_0^{\nu}$  represents the output-oriented distance function calculated using technology of period v.

<sup>&</sup>lt;sup>60</sup> As in DEA, distance functions used in Malmquist TFP indexes can either be output-oriented or input-oriented. These two alternative approaches result in the same numeric measure if the technology in periods s and t exhibits constant returns to scale. <sup>61</sup> Points B and C on the PPF have distance function values of 1 i.e. are efficient.

In the presence of technical inefficiency, equation 12 can be rewritten as:

$$m_{o}(y_{s}, y_{t}, x_{s}, x_{t}) = \frac{d_{0}^{t}(x_{t}, y_{t})}{d_{0}^{s}(x_{s}, y_{s})} * \left[\frac{d_{0}^{s}(x_{t}, y_{t})}{d_{0}^{t}(x_{t}, y_{t})} * \frac{d_{0}^{s}(x_{s}, y_{s})}{d_{0}^{t}(x_{s}, y_{s})}\right]^{0.5}$$
Eq. 13

where the ratio outside the square brackets measures the change in the output-oriented measure of technical efficiency between periods s and t, and the geometric mean of the two ratios inside the square brackets captures the shift in technology between the two periods.

Specifically, change in technical efficiency is given by:<sup>62</sup>

Eq. 14 
$$\frac{d_0^t(x_t,y_t)}{d_0^s(x_s,y_s)}$$

and technological change/progress by:

 $\left[\frac{d_0^s(x_t, y_t)}{d_0^t(x_t, y_t)} * \frac{d_0^s(x_s, y_s)}{d_0^t(x_s, y_s)}\right]^{0.5}$ 

Eq. 15

 $<sup>^{62}</sup>$  Change in technical efficiency is measured by a ratio of Shephard technical efficiency indexes calculated for periods *t* and *s*. More exactly, the reciprocal of the output distance function is equivalent to Farrell output-oriented technical efficiency, which is equivalent to the Shephard measure in the CRS model.

## ANNEX 10 Country profiles -low efficiency scores



## Czech Republic: health care indicators

Source : Own calculations. Eurostat, OECD Data

Note: In all panels data points outside the average circle indicate that the level of the variable for the group or the country under scrutiny is higher than the average OECD country (e.g. Latvia has higher amenable mortality rates than the EU average).

In Panel A, data points outside the average circle indicate that the country performs better than the EU average. In Panel B, data points outside the average circle indicate that the country performs better than the EU average.

consultations than the EU average.



## Hungary: health care indicators

Source : Own calculations. Eurostat, OECD Data

Note: In all panels data points outside the average circle indicate that the level of the variable for the group or the country under scrutiny is higher than the average OECD country (e.g. Latvia has higher amenable mortality rates than the EU average). In Panel A, data points outside the average circle indicate that the country performs better than the EU average. In Panel B, data points outside the average circle indicate that the country has higher health outcomes and, more hospital discharges and outpatient consultations than the EU average.

### Lithuania: health care indicators



#### A. Efficiency scores

B. Outputs

EU average 🗕 🗕 Lithuania

Source : Own calculations. Eurostat, OECD Data

Note: In all panels data points outside the average circle indicate that the level of the variable for the group or the country under scrutiny is higher than the average OECD country (e.g. Latvia has higher amenable mortality rates than the EU average). In Panel A, data points outside the average circle indicate that the country performs better than the EU average. In Panel B, data points outside the average circle indicate that the country has higher health outcomes and, more hospital discharges and outpatient

## Latvia: health care indicators



#### A. Efficiency scores





EU average 🛛 🗕 🗕 Latvia

Source : Own calculations. Eurostat, OECD Data

Note: In all panels data points outside the average circle indicate that the level of the variable for the group or the country under scrutiny is higher than the average OECD country (e.g. Latvia has higher amenable mortality rates than the EU average). In Panel A, data points outside the average circle indicate that the country performs better than the EU average. In Panel B, data points outside the average circle indicate that the country has higher health outcomes and, more hospital discharges and outpatient

## Poland: health care indicators



#### A. Efficiency scores

#### B. Outputs

EU average - - Poland

Source : Own calculations. Eurostat, OECD Data

Note: In all panels data points outside the average circle indicate that the level of the variable for the group or the country under scrutiny is higher than the average OECD country (e.g. Latvia has higher amenable mortality rates than the EU average). In Panel A, data points outside the average circle indicate that the country performs better than the EU average. In Panel B, data points outside the average circle indicate that the country has higher health outcomes and, more hospital discharges and outpatient

### Slovakia: health care indicators



#### A. Efficiency scores

#### B. Outputs

EU average Slovakia

Source : Own calculations. Eurostat, OECD Data

Note: In all panels data points outside the average circle indicate that the level of the variable for the group or the country under scrutiny is higher than the average OECD country (e.g. Latvia has higher amenable mortality rates than the EU average). In Panel A, data points outside the average circle indicate that the country performs better than the EU average. In Panel B, data points outside the average circle indicate that the country has higher health outcomes and, more hospital discharges and outpatient

#### **ECONOMIC PAPERS**

As of n° 120, Economic Papers can be accessed and downloaded free of charge at the following address: <u>http://ec.europa.eu/economy\_finance/publications/economic\_paper/index\_en.htm</u>

Alternatively, hard copies may be ordered via the "Print-on-demand" service offered by the EU Bookshop: <u>http://bookshop.europa.eu</u>.

## HOW TO OBTAIN EU PUBLICATIONS

#### Free publications:

- one copy: via EU Bookshop (<u>http://bookshop.europa.eu</u>);
- more than one copy or posters/maps:
  - from the European Union's representations (<u>http://ec.europa.eu/represent\_en.htm</u>);
  - from the delegations in non-EU countries (<u>http://eeas.europa.eu/delegations/index\_en.htm</u>);
  - by contacting the Europe Direct service (<u>http://europa.eu/europedirect/index\_en.htm</u>) or calling 00 800 6 7 8 9 10 11 (freephone number from anywhere in the EU) (\*).
  - (\*) The information given is free, as are most calls (though some operators, phone boxes or hotels may charge you).

#### Priced publications:

• via EU Bookshop (<u>http://bookshop.europa.eu</u>).

