Chapter 6
Evaluating Healthcare System Efficiency of OECD Countries: A DEA-Based Study

Zehra Önen and Serpil Sayın

6.1 Introduction

Healthcare is one of the important public policy issues of our time. Quality of healthcare services, its distribution and accessibility are often perceived as being related to quality of life of individuals. As advances in technology and sciences are reflected into healthcare domain and as demand increases, costs and expenditures have been increasing, reaching levels between 10–20% of gross domestic product (GDP) in several countries (OECD 2013). This has contributed to a higher emphasis being placed on performance evaluation for healthcare institutions (Kazandjian and Lied 1999; Özcan 2008). While healthcare organizations can be viewed and evaluated as service systems to some extent, it is difficult to assess the quality of the healthcare system of a country. In WHO (2000), the World Health Organization published a ranking of healthcare systems of 191 countries and found France as the country that has the best healthcare system. Rankings of healthcare systems of countries appear in the media from time to time as in Capell (2008), The Guardian (2003) and The New York Times (2007). Most of the time, these articles are based on WHO reports or Organisation for Economic Cooperation and Development (OECD) annual reports as presented in WHO (2015), De Looper and Lafortune (2009), and OECD (2013).

Z. Önen

Faculty of Engineering and Natural Sciences, Istanbul Bilgi University, Santralistanbul, Kazim Karabekir, Cad. No: 2/13 34060 Eyüp, İstanbul
e-mail: zonen@ku.edu.tr

S. Sayın (✉)
College of Administrative Sciences and Economics, Koç University, 34450 Sarıyer, Istanbul, Turkey
e-mail: ssayin@ku.edu.tr

The focus placed on the efficiency of the healthcare system can vary not only across countries but also across the years within a country. Different countries can have different priorities depending on the current state of their economy and other relevant parameters. In addition, when the political system changes and technology progresses, these priorities may also change.

Past research has often identified life expectancy (LE) at birth and infant mortality rate (IMR) as key outcomes of the healthcare system. Mohan and Mirmirani (2007) investigated the significance of different factors that can influence LE and IMR using panel data of 12 years, from 1990 to 2002. Two regression models were built, one where LE was chosen as the dependent variable and another where IMR was chosen as the dependent variable. Several independent variables were used in the regression models. Empirical results indicated that the level of healthcare expenditure among OECD countries has been an important factor in extending LE but did not have much impact on lowering IMR. Education level used as an indicator of health awareness was significant in both of the regressions.

Data Envelopment Analysis (DEA) has been among the methods used for performance evaluation of healthcare systems (Greene 2004; Tandon et al. 2001; Jacobs 2001). DEA is a methodology based on linear programming that can be used to analyze the relative efficiency of similar Decision Making Units (DMUs). DEA has been applied in very different areas with success (Afonso and Aubyn 2005; Galterio et al. 2009; Johnes 2006; Zhou et al. 2008). In healthcare, the methodology has been used for performance evaluation of units such as hospitals (Jacobs 2001; Özcan 2008), or different units within hospitals (Wang and Yu 2006) as well as countries (Afonso and Aubyn 2005; Reinhardt et al. 2002; Varabyova and Schreyögg 2013; Borisov et al. 2012).

Afonso and Aubyn (2005) compared performances of 24 OECD countries in education and healthcare using two different methodologies, Free Disposal Hull (FDH) and DEA. They used number of physicians, nurses and beds per 1000 population (for brevity, the term *per 1000 population* will be omitted when referring to these widely used inputs in this manuscript) as input measures and infant survival rate (ISR) and LE at birth as output measures. They used data from the year 2000. According to their results, 11 out of 24 countries were efficient in the FDH analysis whereas 8 of them remained efficient in the DEA analysis.

Greene (2004) analyzed 191 countries based on WHO data, using stochastic frontier analysis. He used disability adjusted LE and a composite measure of health care delivery as output measures. As input measures, health expenditure per capita in 1997 and average years of schooling were considered. He also used different variables that were considered as indicators of cross country heterogeneity such as the Gini coefficient that measures income inequality, and OECD membership. As a result, Greene pointed out that expenditure is a major component of healthcare system performance and should be taken into account. The author also found that OECD membership explained much of the variation in the outcome measures and the distribution of income is a significant factor.

In another study, Retzlaff-Roberts et al. (2004) developed different types of DEA models using the OECD 2000 database with data of 1998. As in Afonso

and Aubyn (2005), ISR and LE at birth were considered as output measures but these measures were dealt with separately. For input measures, they considered four healthcare-related inputs and three social environment inputs. The healthcare inputs included the number of physicians, beds, healthcare expenditure as a percentage of GDP and MRI units per million population. The social environment input variables included the expected number of years of education, the Gini coefficient and the maximum value of the percentage of male and female smokers. From their different models, the authors concluded that countries with relatively modest outcomes like Turkey and Mexico turned out to be efficient while other countries with good health outcomes were not necessarily using their resources efficiently.

Varabyova and Schreyögg (2013) analyzed the hospital care efficiency in OECD countries using panel data between 2000 and 2009. They used DEA and Stochastic Frontier Analysis (SFA) methods and compared them. For both methodologies, they considered hospital discharges and mortality as output measures whereas hospital resources such as number of beds, physicians, nurses and hospital employments were the inputs in the models. Total hospital employments refer to the number of persons employed (including self-employed and full-time equivalent employed) in general and special hospitals. Varabyova et al. argued that their analyses are good indicators of efficient use of resources. Hence, countries with good health outcomes in terms of longevity like Japan can be inefficient while developing countries like Turkey can be efficient.

Samut and Cafrı (2015) analyzed the efficiency of hospitals for 29 OECD countries between 2000 and 2010 to understand the parameters that have an impact on efficiency. They used a two stage model in which DEA and Panel Tobit were used for the first and second stage respectively. As a result they found that countries that were fully efficient during this 10-year period were Mexico, Turkey, and the United Kingdom. On the other hand, Japan, Iceland, France and Belgium had under-average efficiency scores for the same period. From the Panel Tobit analysis the authors pointed out that wealthier countries had better hospital efficiency. Indeed, they found a positive relation between GDP and education along with a positive relation between GDP and efficiency.

Frogner et al. (2015) analyzed health data of 25 OECD member countries using panel data analysis. They applied stochastic frontier analysis and fixed effect analyses over 11 input variables, including health care resources, health-related behavior, and economic and environmental factors. They built 36 different models and by comparing their results from different models, they revealed the fragility of the results of ranking models. However, they were not able to demonstrate that the U.S., in particular, performed significantly better than its WHO ranking in these alternative ranking models.

In this study, we experiment with Assurance Region Global (ARG) models based on the expectation that they may produce more conservative and consistent evaluations than those produced by standard DEA models. In addition to the widely utilized outputs of ISR and LE at birth, we also explore a different approach considering survival rates from major causes of death as output variables. Our goal is to see how efficiency and inefficiency with respect to these models differ from

those obtained by traditional models. We also provide a comprehensive evaluation of OECD member countries as some past studies exclude some member countries due to a lack of data observations (Mohan and Mirmirani 2007; Jacobs 2001; Afonso and Aubyn 2005; Retzlaff-Roberts et al. 2004; Frogner et al. 2015). We use OECD data for the years 2008 and 2012 therefore making it possible to observe any shifts in healthcare system performance at different points in time in the studied countries. In the next section, we explain the DEA methodology and present basic models. In Sect. 6.3, we describe how the data is processed, give our model results and discuss them. In Sect. 6.4, we build a model with specific causes of death as outputs. Finally, we conclude and provide further research directions in Sect. 6.5.

6.2 DEA

DEA methodology was originally introduced by Charnes, Cooper and Rhodes in 1978 as a method for evaluating the relative efficiency of Decision Making Units (DMUs) performing essentially the same task. The methodology got its name because of the idea of *enveloping* the observations to identify an efficient frontier. This frontier is computed via a ratio where multiple inputs produce multiple outputs (Joro et al. 1998; Cooper et al. 2006). The main idea of this linear programming (LP) model is to have a score between 0 and 1 representing the degree of efficiency of a DMU where 1 represents an efficient DMU. The model provides also an identification of sources and amounts of possible inefficiency and a direction of improvement based on orthogonal projection of the observation to the frontier.

Let n denote the number of DMUs to be evaluated. Suppose that there are m input and s output variables and let the input and output data be denoted by matrices X and Y of size $m \times n$ and $s \times n$ respectively. DEA methodology seeks to attach weights to each input and output variable. In the formulation below, these are decision variables denoted as $v_i, \ i = 1 \ldots, m$ and $u_j, \ j = 1, \ldots s$ respectively.

The first model introduced by Charnes, Cooper and Rhodes, denoted as CCR, sets up a fractional problem (FP_k) for an arbitrary DMU_k for $k \in \{1, \ldots, n\}$ which can be expressed as:

$$\max \theta = \frac{u_1 y_{1k} + u_2 y_{2k} + \ldots + u_s y_{sk}}{v_1 x_{1k} + v_2 x_{2k} + \ldots + v_m x_{mk}}$$

$$subject \ to \frac{u_1 y_{1j} + u_2 y_{2j} + \ldots + u_s y_{sj}}{v_1 x_{1j} + v_2 x_{2j} + \ldots + v_m x_{mj}} \leq 1, \forall j = 1, \ldots, n \qquad (6.1)$$

$$v_1, \ldots, v_m \geq 0, \qquad (6.2)$$

$$u_1, \ldots, u_s \geq 0, \qquad (6.3)$$

The objective is to obtain the optimal weights such that the ratio of weighted output over weighted input is maximized for DMU_k. In other words, the objective is to find the most favorable weights for DMU_k that maximizes weighted output obtained per

weighted input used. Constraint (6.1) ensures that for these weights, output-to-input ratios for all DMUs are between zero and one. Constraints (6.2) and (6.3) are non-negativity constraints for the weights. Also the values for inputs and outputs are assumed to be positive. At optimality, (v_k^*, u_k^*) represents the set of most favorable weights for DMU_k that maximizes the ratio scale. Each weight shows how highly the associated input or output is evaluated relatively. Note that the above formulation is a fractional program that can be linearized relatively easily. In addition, an *output-oriented* version of this model can be written as opposed to this *input-oriented* version. The orientation is named based on the objective function of the dual problem. Discussion of equivalence between these two versions, derivations of equivalent linear programming transformations as well as a discussion of alternative DEA models can be found in Cooper et al. (2006). In this study, we employ the output oriented BCC model which differs from the CCR model by an additional convexity constraint in the dual formulation. This constraint translates into a new variable free in sign in the LP model in which the observations for the n DMUs may be combined, thus allowing a variable returns to scale in the production frontier. Formulation (BCC-O-FP) represents the dual problem with the convexity constraint and the corresponding fractional problem of an output-oriented BCC model.

$$(BCC - O - FP) \quad \min \bar{\theta} = \frac{v_1 x_{1k} + v_2 x_{2k} + \ldots + v_m x_{mk} - v_0}{u_1 y_{1k} + u_2 y_{2k} + \ldots + u_s y_{sk}}$$

$$subject\ to\ \frac{v_1 x_{1j} + v_2 x_{2j} + \ldots + v_m x_{mj} - v_0}{u_1 y_{1j} + u_2 y_{2j} + \ldots + u_s y_{sj}} \geq 1, \forall j = 1, \ldots, n \quad (6.4)$$

$$v_1, \ldots, v_m \geq 0, \quad (6.5)$$

$$u_1, \ldots, u_s \geq 0, \quad (6.6)$$

$$v_0 \quad free\ in\ sign \quad (6.7)$$

Note that since in BCC there is an additional constraint compared to CCR, the feasible region of the latter problem contains the feasible region of BCC. Hence, any BCC efficient DMU is CCR efficient.

DEA methodology has been used in a variety of application areas such as education, healthcare, energy efficiency, among others (Galterio et al. 2009; Johnes 2006; Zhou et al. 2008; Liu et al. 2013). One drawback of the methodology is that it may present a very favorable outlook of a DMU since it is designed to choose input and output weights in a way to benefit that particular DMU the most. In other words, the methodology is capable of overemphasizing strengths of a DMU while ignoring its weaknesses by setting respective weights to zero. As a way to alleviate this shortcoming, imposing restrictions on weight vectors u and v has been proposed. Thompson et al. (1986) developed the assurance region approach which is based on imposing constraints on the magnitude of the weights for specific inputs or outputs of DMUs relative to each other. Another proposal presented in Charnes et al. (1990), known as the cone ratio approach has been to restrict input and output weights to predetermined cones via additional constraints. Wong and Beasley (1990) propose limiting the proportion of total output of DMU k devoted to output measure i by imposing limits as follows:

$$L_i \leq \frac{u_i y_{ik}}{\sum_{j=1}^{s} u_j y_{jk}} \leq U_i. \tag{6.8}$$

Similar constraints can be defined for inputs and their corresponding weights. This approach has the advantage of restricting relative weights instead of virtual weights and therefore may be more intuitive. While Wong and Beasley (1990) motivated this approach as a means of incorporating value judgments in a DEA model, Cooper et al. (2006) mentioned that it can also be used to establish some consistency in weight choices of different DMUs by a careful choice of the bounds L_i and U_i. We employ this approach as a variation of our base model and label it as the ARG model following the terminology in Cooper et al. (2006). Since there are additional constraints in ARG models, it may be expected that some DMUs that were formerly classified as efficient may become inefficient once the constraints are imposed.

6.3 LE, Infant Mortality and Efficiency

Our data comes from the OECD online library. OECD is an organization that aims to promote policies that will improve the economic and social well-being of people around the world (OECD 2016). It has been established in 1961 and it consisted of 34 member countries at the time analyses were conducted for this study. With Latvia becoming a member in July 2016, the number of member countries has reached 35. OECD collects data from its members in order to develop policies with respect to its mission. It keeps track of significant amounts of data about not only economics, taxes, trade finance but also education, health, environment and social issues. Within healthcare domain, OECD collects a variety of data from various expenditure figures to amounts of tobacco consumption. The reader can find the list of variables related to healthcare in Health at a Glance report of OECD (2013). Table 6.1 provides a list of variables that we use in this part of our study. In these models, we pick our two output measures as ISR and LE at birth. ISR is computed using IMR as given in Afonso and Aubyn (2005) by:

$$ISR = \frac{1000 - IMR}{IMR} \tag{6.9}$$

ISR represents the ratio of children that survived their first year to the number of children that died. As inputs, we pick number of physicians, number of nurses and number of hospital beds. A vast majority of the countries account for number of physicians and nurses as practicing professionals. A few countries report the number of physicians and nurses by including practicing physicians or nurses plus others working in the healthcare sector as managers, educators and researchers, adding another 5–10% to each group. We use the figures as reported in the database. We take two cross sections of data for the 34 countries from 2008 and 2012 to be able to take two snapshots in time and to see if any differences can be observed.

Table 6.1 Description of input and output variables

Variable	Description
Number of physicians	Number of practising, professionally active or licensed to practice physicians per 1000 population.
Number of nurses	Number of practising, professionally active or licensed to practice nurses per 1000 population.
Number of hospital beds	All hospital beds regularly maintained, staffed and immediately available per 1000 population.
	(including curative care beds, rehabilitative care beds, long-term care beds and other hospital beds).
ISR	Infant Survival Rate. Computed via Eq. (6.9) using IMR.
	IMR is the number of deaths in children under 1 year of age per 1000 live births that occurred in a given year.
LE at birth	The average number of years that a person at birth is expected to live,
	assuming that age-specific mortality levels remain constant.

The missing data in our data set were estimated via the previous and future available data and by means of linear interpolation.[1] In the literature, it is suggested that the number of DMUs should exceed 3 times the total number of inputs and outputs. With 34 DMUs, 3 inputs and 2 outputs, we obey this guideline. All of our models are based on output-oriented BCC approach. In addition to the base models, we build ARG models where weights on outputs are imposed with the values of $L_i = 0.4$ and $U_i = 0.6$ for $i = 1, 2$ in constraint (6.8). These bounds are chosen so that nearly equal importance is given to both output variables. We use DEA-Solver-Learning Version developed by Kaoru Tone where the platform is Microsoft Excel 2003 (Cooper et al. 2006).

6.3.1 2008 Models with Respect to LE and Infant Mortality

In terms of inputs and outputs, our base BCC model is parallel to the work of Afonso and Aubyn (2005). The descriptive statistics of the variables are presented in Table 6.2. In addition to the implementation of the base model on 2008 data, we run our ARG model as well and observe the differences between the results of the two models. As we use the same inputs and outputs, in the ARG models we expect to see a subset of the efficient countries of the base model. The efficiency scores of all countries for both BCC and ARG models can be found in Table 6.3. In the following tables with efficiency scores, a score of 1 is presented in bold in order to highlight countries that are efficient. In 2008 BCC model, Canada, Chile, Greece, Italy, Japan, South Korea, Luxembourg, Mexico, New Zealand, Spain, Sweden and Switzerland

---

[1]A total of 14 data items, 8 for 2008, 6 for 2012 of the total 340 have been estimated.

Table 6.2 Descriptive statistics of variables for 2008

	Number of physicians	Number of nurses	Number of beds	LE	ISR
Mean (σ)	3.04 (0.91)	8.35 (3.91)	5.08 (2.39)	79.24 (2.55)	269.8 (104.2)
Max (DMU)	6 (Greece)	14.9 (Switzerland)	13.8 (Japan)	82.7 (Japan)	554.6 (Luxembourg)
Min (DMU)	1 (Chile)	0.5 (Chile)	1.7 (Mexico)	73.6 (Turkey)	64.8 (Mexico)

Table 6.3 Scores of all OECD countries for 2008 models

DMU	BCC	ARG	DMU	BCC	ARG
Australia	0.996	0.776	Japan	**1**	0.992
Austria	0.980	0.769	South Korea	**1**	0.994
Belgium	0.972	0.699	Luxembourg	**1**	**1**
Canada	**1**	0.801	Mexico	**1**	**1**
Chile	**1**	**1**	Netherlands	0.981	0.772
Czech Republic	0.945	0.872	New Zealand	**1**	0.973
Denmark	0.965	0.756	Norway	0.992	0.929
Estonia	0.903	0.682	Poland	0.936	0.727
Finland	0.980	0.886	Portugal	0.974	0.955
France	0.987	0.771	Slovak Republic	0.916	0.627
Germany	0.975	0.727	Slovenia	0.983	0.975
Greece	**1**	**1**	Spain	**1**	**1**
Hungary	0.903	0.652	Sweden	**1**	**1**
Iceland	0.997	0.869	Switzerland	**1**	0.693
Ireland	0.978	0.728	Turkey	0.940	0.632
Israel	0.995	0.910	United Kingdom	0.982	0.786
Italy	**1**	0.934	United States	0.964	0.698

are efficient. We can see that some of the developed economies (e.g. Germany) are inefficient whereas some of the developing economies are efficient (e.g. Chile).[2] A similar counter intuitive result was pointed out in the work of Retzlaff-Roberts et al. (2004) with data from OECD 2000 database. Note that due to the nature of DEA methodology, inefficiency does not necessarily imply a deficiency in outputs. The developed countries may be using more inputs compared to developing ones for getting certain level of outputs whereas developing countries do relatively well with their limited available resources. A closer look at the output weights of the efficient

---

[2]The distinction between developing and developed economies is made based on the classification of the UN (2015) World Economic Situation and Prospect report.

Table 6.4 Descriptive statistics of variables for 2012

	Number of Physicians	Number of Nurses	Number of Beds	LE	ISR
Mean (σ)	3.22 (0.93)	8.97 (4.07)	4.80 (2.50)	80.18 (2.44)	307.2 (151.1)
Max (DMU)	6.27 (Greece)	16.97 (Switzerland)	13.36 (Japan)	83.2 (Japan)	908.1 (Iceland)
Min (DMU)	1.73 (Turkey)	1.79 (Turkey)	1.57 (Mexico)	74.4 (Mexico)	74.2 (Mexico)

countries reveals that in 2008, Greece, Luxembourg, Spain and Sweden have zero weight on their LE at birth output. 16 out of the 22 inefficient countries achieve their best by nullifying one of their output weights.

In the 2008 ARG model, six of the BCC-efficient countries (Chile, Greece, Luxembourg, Mexico, Spain and Sweden) are still efficient. On the other hand, Canada, Italy, Japan, South Korea, New Zealand and Switzerland lost their efficiency once the additional constraints are added. We note that countries that lose their efficiency when a more balanced combination of outputs is enforced are developed countries. This strengthens the interpretation that the inefficiencies may stem from abundant inputs rather than poor outputs. We also observe that scores now come from a wider range between 0.627 and 1, as expected.

6.3.2 2012 Models with Respect to LE and Infant Mortality

The descriptive statistics of the variables for 2012 are presented in Table 6.4. In the 2012 BCC model, some countries that were efficient in 2008 are not efficient whereas some countries that were not efficient in 2008 are observed as efficient in 2012. The countries that were efficient in 2008 but not in 2012 are Italy, Luxembourg, New Zealand and Switzerland. In 2012, Iceland, Israel, Slovenia and Turkey join the list of efficient countries. Even more inefficient countries (21 out of 22) have zero weight on one of their output weights. In 2012, Canada, Israel and Sweden are the countries that are BCC efficient but inefficient with respect to the ARG model. Again, it is mainly developed countries that are forced to inefficiency by a balanced use of outputs. Furthermore in the ARG model scores vary between 0.523 and 1. All the results for BCC and ARG models for 2012 can be found in Table 6.5.

Table 6.5 Scores of all OECD countries for 2012 models

DMU	BCC	ARG	DMU	BCC	ARG
Australia	0.995	0.673	Japan	**1**	**1**
Austria	0.978	0.730	South Korea	**1**	**1**
Belgium	0.976	0.611	Luxembourg	0.990	0.777
Canada	**1**	0.676	Mexico	**1**	**1**
Chile	**1**	**1**	Netherlands	0.981	0.565
Czech Republic	0.945	0.794	New Zealand	0.993	0.575
Denmark	0.969	0.572	Norway	0.982	0.657
Estonia	0.928	0.756	Poland	0.950	0.730
Finland	0.976	0.751	Portugal	0.975	0.864
France	0.991	0.658	Slovak Republic	0.925	0.591
Germany	0.975	0.598	Slovenia	**1**	**1**
Greece	**1**	**1**	Spain	**1**	**1**
Hungary	0.913	0.623	Sweden	**1**	0.832
Iceland	**1**	**1**	Switzerland	0.997	0.523
Ireland	0.991	0.732	Turkey	**1**	**1**
Israel	**1**	0.970	United Kingdom	0.992	0.722
Italy	0.997	0.915	United States	0.966	0.565

6.3.3   Discussion of Results

Looking at the overall results, it is possible to observe that Chile, Greece, Mexico and Spain are efficient in both models for both years. On the other hand, 18 countries (including Austria, France, the United Kingdom and the United States) out of 34 are always inefficient regardless of the model types and years.

We conducted reference set analyses on the 2008 and the 2012 ARG models. Tables 6.6 and 6.7 report the reference set members along with the associated weights[3] for the inefficient countries for 2008 and 2012, respectively. In Table 6.6 we can see that Luxembourg is dominant with an occurrence of 26 out of 28. Furthermore, Luxembourg is the country with the highest reference set weight value for more than half of the countries. Therefore it can be stated that Luxembourg, with the highest ISR (554.6) and its LE at birth being relatively good (80.7), can be seen as the "ideal" country in terms of healthcare system performance. This outcome is worth attention because Retzlaff-Roberts et al. did not include Luxembourg in their data set due to lack of information in OECD 2000 database. Chile, Sweden and Spain are the other countries that appear frequently in the reference sets. An interesting observation is that Luxembourg is not efficient in either of the 2012 models although it has a strong appearance in 2008 models. DEA model results include the values of the slacks that give a direction of improvement from the computation of the model. We remark that most of the excess is present in the number of nurses and number of hospital beds according to ARG models.

---

[3]Weight are rounded to two digit accuracy and the total may deviate from 1.00 for some countries.

Table 6.6 Reference set for inefficient countries for ARG model in 2008

DMU	Reference set
Australia	Sweden (0.49); Luxembourg (0.38); Chile (0.13)
Austria	Luxembourg (0.56); Greece (0.44)
Belgium	Luxembourg (1.00)
Canada	Chile (0.43); Sweden (0.31); Luxembourg (0.26)
Czech Republic	Luxembourg (0.64); Greece (0.27); Chile (0.09)
Denmark	Sweden (0.70); Luxembourg (0.29); Chile (0.01)
Estonia	Luxembourg (0.49); Greece (0.28); Chile (0.22)
Finland	Luxembourg (0.87); Chile (0.1); Greece (0.03)
France	Luxembourg (0.65); Greece (0.22); Chile (0.12)
Germany	Luxembourg (0.97); Greece (0.03)
Hungary	Luxembourg (0.48); Chile (0.27); Greece (0.25)
Iceland	Luxembourg (1.00)
Ireland	Luxembourg (0.76); Sweden (0.2); Chile (0.04)
Israel	Spain (0.68); Luxembourg (0.14); Chile (0.09); Greece (0.09)
Italy	Spain (0.53); Luxembourg (0.3); Chile (0.12); Sweden (0.04)
Japan	Luxembourg (0.67); Chile (0.33)
South Korea	Chile (0.59); Luxembourg (0.36); Greece (0.05)
Netherlands	Luxembourg (0.65); Spain (0.29); Chile (0.06)
New Zealand	Sweden (0.42); Mexico (0.35); Chile (0.22)
Norway	Sweden (0.75); Luxembourg (0.25)
Poland	Chile (0.48); Luxembourg (0.43); Greece (0.09)
Portugal	Spain (0.89); Luxembourg (0.09); Sweden (0.02)
Slovak Republic	Luxembourg (0.47); Greece (0.31); Chile (0.22)
Slovenia	Luxembourg (0.7); Chile (0.27); Greece (0.03)
Switzerland	Luxembourg (0.86); Sweden (0.14)
Turkey	Chile (0.86); Spain (0.08); Sweden (0.05)
United Kingdom	Sweden (0.41); Chile (0.32); Luxembourg (0.27)
United States	Chile (0.42); Sweden (0.4); Luxembourg (0.18)

A reference set analysis for 2012 is given in Table 6.7. It can be observed that Slovenia is dominant with an occurrence of 21 out of 25 times. Also Iceland has an occurrence of 17 out of 25. Chile, Greece, Spain and Turkey are the other countries that appear frequently in the reference sets.

6.4 Survival from Major Causes of Death and Efficiency

Our results above show that when LE at birth and ISR are taken as the outputs that indicate system efficiency, developing countries may have an advantage as they achieve relatively good results in relation to the inputs they provide to the system.

Table 6.7 Reference set for inefficient countries for ARG model in 2012

DMU	Reference set
Australia	Slovenia (0.42); Iceland (0.38); Spain (0.2)
Austria	Slovenia (0.93); Greece (0.07)
Belgium	Slovenia (0.81); Iceland (0.19)
Canada	Chile (0.57); Iceland (0.4); Slovenia (0.03)
Czech Republic	Slovenia (0.98); Greece (0.02)
Denmark	Iceland (0.91); Mexico (0.09)
Estonia	Slovenia (0.54); Greece (0.23); South Korea (0.18); Turkey (0.05)
Finland	Slovenia (0.54); Iceland (0.46)
France	Slovenia (0.87); Iceland (0.13)
Germany	Iceland (0.64); Slovenia (0.36)
Hungary	Slovenia (0.51); South Korea (0.3); Greece (0.18)
Ireland	Iceland (0.53); Chile (0.46); Slovenia (0.01)
Israel	Spain (0.67); Turkey (0.21); Slovenia (0.11); Greece (0.01)
Italy	Spain (0.69); Slovenia (0.27); Iceland (0.04)
Luxembourg	Slovenia (0.77); Iceland (0.23)
Netherlands	Iceland (0.57); Slovenia (0.43)
New Zealand	Iceland (0.51); Chile (0.46); Slovenia (0.03)
Norway	Iceland (1.00)
Poland	South Korea (0.42); Slovenia (0.39); Turkey (0.18); Greece (0.01)
Portugal	Spain (0.73); Slovenia (0.26); Greece (0.01)
Slovak Republic	Slovenia (0.44); Greece (0.26); South Korea (0.24); Turkey (0.05)
Sweden	Iceland (0.62); Mexico (0.38)
Switzerland	Iceland (1.00)
United Kingdom	Mexico (0.42); Iceland (0.35); Slovenia (0.21); Spain (0.02)
United States	Chile (0.49); Iceland (0.35); Slovenia (0.16)

Obviously our models correspond to a very high level analysis of the healthcare system. In order to focus on the system's ability to deal with major health issues, we analyze some major causes of death and their corresponding survival rates as outputs. According to WHO fact sheet, the first two major causes of death in the world are ischemic heart disease and stroke. Death from all types of cancer is also listed as a major cause of death in WHO (2014). We build new models by using survival rates of ischemic heart disease, cerebrovascular disease (since stroke is the most common type of cerebrovascular diseases) and malignant neoplasms as outputs of the system while the inputs remain the same. We label our new models that consider survival rates from three major causes of death (ischemic heart disease, cerebrovascular disease, malignant neoplasms) as outputs along with the same inputs as in previous models BCCs and ARGs. Data from OECD library is presented in the form of age standardized mortality rates per 100,000 population. These rates are calculated by the OECD Secretariat, using the total OECD population for 2010 of each corresponding country as the reference population. They use the method

Table 6.8 Descriptive statistics of mortality rates from common causes of death for 2008

	Ischemic heart MR	Cerebrovascular MR	Malignant neoplasms MR
Mean (σ)	135.72 (86.31)	78.75 (29.35)	215.06 (32.66)
Max (Country)	428.4 (Slovak Republic)	148.4 (Slovak Republic)	294.7 (Hungary)
Min (Country)	41.9 (Japan)	41.8 (Israel)	124.5 (Mexico)

of standardization for age-standardized calculations to be able to compare the level of mortality across countries and over time. We compute survival rates based on mortality rates using Eq. (6.9) replacing 1000 by 100,000.

6.4.1 2008 Models with Respect to Survival from Major Causes of Death

For 2008, Turkey is the only country with missing data. We use data from 2009 for Turkey. The descriptive statistics of the output variables for 2008 are presented in Table 6.8.

Table 6.9 shows the results of the BCC and ARG models. In the BCC model nearly half of the countries (16 out of 34) are efficient. The efficient countries are Canada, Chile France, Israel, Italy, Japan, South Korea, Luxembourg, Mexico, Netherlands, Portugal, Slovenia, Switzerland, Turkey, the United Kingdom and the United States. Compared to the BCC model in Sect. 6.3.1, France, Israel, Netherlands, Portugal, Slovenia, Turkey, the United Kingdom and the United States became efficient with the new output variables. On the other hand, Greece, New Zealand and Sweden are no longer efficient. Looking at the ARG model results, we see that Portugal, Slovenia, Turkey and the United Kingdom are not listed as efficient anymore. Also looking at the scores, we see that the discrepancy is much more apparent when the output variables changed. This can be explained by high standard deviation values of the output variables.

6.4.2 2012 Models with Respect to Survival From Major Causes of Death

The most recent and complete data about survivals from the three conditions belong to 2012. Data on survivals from these conditions are not available for Canada, Iceland and Slovenia. We use the most recent data for these countries: 2011 for Canada, 2010 for Iceland and 2009 for Slovenia. The descriptive statistics of the variables are presented in Table 6.10. Table 6.11 shows the results of both models for BCC and ARG models using 2012 data.

Table 6.9 Scores of all OECD countries for 2008 mortality output models

	BCCs	ARGs		BCCs	ARGs
Australia	0.794	0.777	Japan	1	1
Austria	0.812	0.673	South Korea	1	1
Belgium	0.794	0.791	Luxembourg	1	1
Canada	1	1	Mexico	1	1
Chile	1	1	Netherlands	1	1
Czech Republic	0.505	0.417	New Zealand	0.783	0.745
Denmark	0.773	0.773	Norway	0.781	0.781
Estonia	0.547	0.412	Poland	0.558	0.554
Finland	0.784	0.631	Portugal	1	0.845
France	1	1	Slovak Republic	0.494	0.373
Germany	0.778	0.700	Slovenia	1	0.642
Greece	0.737	0.701	Spain	1	1
Hungary	0.450	0.334	Sweden	0.766	0.766
Iceland	0.804	0.691	Switzerland	1	1
Ireland	0.694	0.619	Turkey	1	0.903
Israel	1	1	United Kingdom	1	0.688
Italy	1	1	United States	1	1

Table 6.10 Descriptive statistics of mortality rates from common causes of death for 2012

	Ischemic heart MR	Cerebrovascular MR	Malignant neoplasms MR
Mean (σ)	114.32 (71.79)	66.27 (23.92)	207.19 (32.03)
Max (Country)	318.6 (Slovak Republic)	122.9 (Slovak Republic)	293.3 (Hungary)
Min (Country)	37.7 (Japan)	37.1 (Switzerland)	117.1 (Mexico)

We observe that Canada, Chile, France, Israel, Japan, South Korea, Mexico, Spain and Turkey are efficient in both models. Portugal and Switzerland are BCC efficient but ARG inefficient. France, Portugal and Switzerland are efficient when survival rates from major causes of death are considered as outputs instead of LE at birth and ISR. On the other hand, Greece, Iceland Slovenia and Sweden are not efficient when survival rates from major causes of death are considered as outputs although they are efficient with respect to LE at birth and ISR. Again, we observe that efficiency and inefficiency do not necessarily align with the status of countries as developed or developing. Looking at the reference sets presented in Table 6.12, even though France is not efficient when classical outputs are considered, this country can be followed as the model country for most of the inefficient countries of the ARG model when main causes of mortality are considered as output variables. Also Canada, Mexico and Spain appear frequently in the reference sets.

Table 6.11 Scores of all OECD countries for 2012 mortality output models

	BCCs	ARGs		BCCs	ARGs
Australia	0.868	0.856	Japan	**1**	**1**
Austria	0.833	0.653	South Korea	**1**	**1**
Belgium	0.829	0.826	Luxembourg	0.865	0.865
Canada	**1**	**1**	Mexico	**1**	**1**
Chile	**1**	**1**	Netherlands	0.921	0.902
Czech Republic	0.526	0.414	New Zealand	0.717	0.713
Denmark	0.865	0.849	Norway	0.777	0.773
Estonia	0.659	0.476	Poland	0.598	0.597
Finland	0.742	0.622	Portugal	**1**	0.853
France	**1**	**1**	Slovak Republic	0.465	0.381
Germany	0.790	0.699	Slovenia	0.628	0.611
Greece	0.802	0.768	Spain	**1**	**1**
Hungary	0.431	0.359	Sweden	0.804	0.804
Iceland	0.724	0.702	Switzerland	**1**	0.944
Ireland	0.761	0.752	Turkey	**1**	**1**
Israel	**1**	**1**	United Kingdom	0.794	0.794
Italy	0.767	0.760	United States	0.961	0.910

6.5 Conclusion

This research aims to provide an evaluation of healthcare system efficiency of 34 OECD countries. DEA methodology is used with different modeling techniques and with different output measures. OECD data for 2008 and 2012 are used. The two output measures we use in the first models, LE at birth and ISR, are generally accepted measures as system output in the literature. We suggest the use of ARG models in a way to balance the two outputs as a remedy against the overly optimistic nature of the DEA methodology. To our knowledge, this is the first study that suggests using an ARG model for evaluating healthcare system efficiency at the country level. We observe a more consistent result when constraints on relative weights are imposed as weight restrictions require a more balanced relative output generation. A reference set analysis on the ARG model of 2008 displays Luxembourg as a good role model for inefficient countries whereas for 2012, Slovenia assumes that role.

In addition to using LE and ISR as outputs as is traditionally done, we experiment with survival rates from major causes of death as possible measures of outputs and implement this model with 2008 and 2012 data. When compared to LE at birth as a measure, this model rules out deaths from less common causes as well as deaths whose occurrence have a more indirect relationship to the performance of the healthcare system, such as deaths from accidents, homicides and suicides. In other words, under the assumption that most countries spend more effort on increasing their healthcare system capacity geared towards conditions surrounding

Table 6.12 Reference set for inefficient countries for ARGs model in 2012

DMU	Reference set
Australia	France (0.38); Canada (0.35); Mexico (0.28)
Austria	France (0.67); Mexico (0.22); Japan (0.11)
Belgium	France (0.65); Mexico (0.19); Japan (0.14); Canada (0.02)
Czech Republic	France (0.79); South Korea (0.15); Mexico (0.06)
Denmark	Spain (0.82); Canada (0.1); France (0.05); Mexico (0.02)
Estonia	France (0.47); South Korea (0.19); Mexico (0.29); Israel (0.05)
Finland	France (0.74); Mexico (0.24); Japan (0.02)
Germany	France (0.72); Japan (0.28)
Greece	Turkey (0.3); South Korea (0.3); Spain (0.22); Mexico (0.18)
Hungary	France (0.44); South Korea (0.38); Mexico (0.18)
Iceland	Mexico (0.42); Canada (0.29); France (0.29)
Ireland	Canada (0.65); Mexico (0.25); France (0.1)
Italy	Spain (0.37); Mexico (0.35); France (0.28)
Luxembourg	France (0.58); Mexico (0.18); Chile (0.16); Japan (0.06); Canada (0.03)
Netherlands	France (0.58); Spain (0.26); Mexico (0.16)
New Zealand	Mexico (0.47); Canada (0.35); France (0.18)
Norway	France (0.46); Mexico (0.37); Canada (0.17)
Poland	Mexico (0.61); Japan (0.31); Canada (0.05); France (0.03)
Portugal	Spain (0.87); France (0.13)
Slovak Republic	France (0.4); Mexico (0.32); South Korea (0.28)
Slovenia	France (0.4); Mexico (0.33); Chile (0.19); Japan (0.08)
Sweden	Mexico (0.44); Canada (0.26); Israel (0.2); France (0.09)
Switzerland	France (0.61); Canada (0.28); Mexico (0.11)
United Kingdom	Canada (0.48); Spain (0.22); Mexico (0.18); France (0.08); Chile (0.05)
United States	Canada (0.71); Mexico (0.12); France (0.11); Chile (0.06)

most likely causes of death, efficiency with respect to this model can be taken as an indication of their success. We observe that while there is significant overlap among efficient countries of different models, there are also differences. Countries that are not efficient with respect to LE and ISR but are efficient with respect to survival rates from major causes of death may focus on other causes of death to improve their standing. On the other hand, countries that are efficient with respect to LE and ISR but are not efficient with respect to survival rates from major causes of death need to focus more on these major causes of death. A reference set analysis on the ARG model of 2008 and 2012 displays France as a good role model for inefficient countries with these two outputs whereas for the previous models.

Lack of output measures that capture the quality of health services remains as a limitation of the current study. This is mainly due to the difficulty of quantifying system quality and obtaining associated complete data. Bringing commonly accepted appropriate pseudo-measures that quantify system quality into our models remains as future work. Another future research direction is to study how efficiency

of healthcare systems change over time by using, for instance, the Malmquist Index, named after Malmquist (1953). This may be more revealing in terms of the nature of the changes that take place in the studied time interval. Although we studied two snapshots in time to observe any possible shifts with respect to different DEA models and different output measures and also to better relate to past literature, a multi-year study designed to track changes in system performance might provide more insights to policy makers.

References

Afonso A, Aubyn MS (2005) Non-parametric approaches to education and health efficiency in OECD countries. J Appl Econ 8(2):227–246

Borisov D, Cicea C, Turlea C (2012) DEA model for assessing efficiency in providing health care and management decisions. Manage Res Pract 4(1):5–18

Capell K (2008) The french lesson in health care. The Guardian, 26 June 2008. http://www.guardian.co.uk/environment/2008/jun/26/climatechange.scienceofclimatechange

Charnes A, Cooper WW, Huang ZM (1990) Polyhedral cone-ratio DEA models with an illustrative application to large commercial banks. J Econometrics 46(1–2):73–91

Cooper WW, Seiford LM, Tone K (2006) Introduction to data envelopment analysis and its uses with DEA-solver software and references, Springer, Berlin

De Looper M, Lafortune G (2009) Measuring disparities in health status and in access and use of health care in OECD countries. Organisation for Economics Co-operation and Development

Frogner BK, Frech H, Parente ST (2015) Comparing efficiency of health systems across industrialized countries: a panel analysis. BMC Health Serv Res 15(1):1

Galterio L, Helton J, Langabeer J, 2nd, DelliFraine J (2009) Data envelopment analysis: performance normalization and benchmarking in healthcare. J Health Inf Manag 23(3):38–43

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 Econ 13(10):959–980

Jacobs R (2001) Alternative methods to examine hospital efficiency: data envelopment analysis and stochastic frontier analysis. Health Care Manage Sci 4:103–115

Johnes J (2006) Data envelopment analysis and its application to the measurement of efficiency in higher education. Econ Educ Rev 25(3):273–288

Joro T, Korhonen P, Wallenius J (1998) Structural comparison of data envelopment analysis and multiple objective linear programming. Manage Sci 44(7):962–970

Kazandjian VA, Lied T (1999) Healthcare performance measurement: systems design and evaluation. ASQ Quality Press, Wisconsin

Liu JS, Lu LY, Lu WM, Lin BJ (2013) A survey of dea applications. Omega 41(5):893–902

Malmquist S (1953) Index numbers and indifference surfaces. Trabajos de Estadistica y de Investigacion Operativa 4(2):209–242

Mohan R, Mirmirani S (2007) An assessment of OECD health care system using panel data analysis. Southwest Bus Econ J 16:21–35

OECD (2013) Health at a glance 2013: OECD indicators. OECD

OECD (2016) Secretary-general's report to ministers 2016. http://dx.doi.org/10.1787/sg_report-2016-en

Özcan YA (2008) Health care benchmarking and performance evaluation: an assessment using data envelopment analysis. International Series in Operations Research and Management Science, vol 120. Springer, Berlin

Reinhardt UE, Hussey PS, Anderson GF (2002) Cross-national comparisons of health systems using OECD data, 1999. Health Affair 21(3):169–181

Retzlaff-Roberts D, Chang FC, Rubin RM (2004) Technical efficiency in the use of health care resources: a comparison of OECD countries. Health Policy 69:55–72

Samut PK, Cafrı R (2015) Analysis of the efficiency determinants of health systems in OECD countries by DEA and panel Tobit. In: Social indicators research. Springer, Berlin, pp 1–20

Tandon A, Murray CJ, Lauer JA, Evans DB (2001) Measuring overall health system performance for 191 countries. World Health Organization

The Guardian (2003) NHS slips down world healthcare league. The Guardian-Society, 14 November 2003. http://www.theguardian.com/society/2003/nov/14/politics.medicineandhealth

The New York Times (2007) World's best medical care? The New York Times, 12 August 2007. http://www.nytimes.com/2007/08/12/opinion/12sun1.html?pagewanted=all&_r=0

Thompson RG, Singleton FDJ, Thrall RM, Smith BA (1986) Comparative site evaluations for locating a high-energy physics lab in Texas. Interfaces 16:35–49

UN (2015) World economic situation and prospects statistical annex 2015. http://www.un.org/en/development/desa/policy/wesp/wesp_archive/2015wespstats-en.pdf

Varabyova Y, 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(1):70–79

Wang R, Yu M (2006) Evaluating the efficiency of hospital's departments using DEA. In: Proceedings of IEEE international conference on service operations and logistics, and informatics (SOLI 2006), pp 1167–1170

WHO (2000) World health organization (2000) the world health report 2000: health systems: improving performance. http://www.who.int/whr/2000/en/

WHO (2014) The top 10 causes of death. Fact sheet no 310, May 2014. http://www.who.int/mediacentre/factsheets/fs310/en/

WHO (2015) World Health Organization (2015) world health statistics 2015. http://www.who.int/gho/publications/world_health_statistics/2015/en/

Wong YHB, Beasley JE (1990) Restricting weight flexibility in data envelopment analysis. J Oper Res Soc 41(9):829–835

Zhou P, Ang BW, Poh KL (2008) A survey of data envelopment analysis in energy and environmental studies. Eur J Oper Res 189(1):1–18