Value efficiency analysis of health systems: does public financing play a role?

González, Eduardo; Cárcaba, Ana; Ventura, Juan

ORIGINAL ARTICLE

Value efficiency analysis of health systems: does public financing play a role?

Eduardo González · Ana Cárcaba · Juan Ventura

Abstract

Aims/Background This study measures the performance of the health system in 165 countries and its relationship with public financing.

Methods We use value efficiency analysis (VEA), a refinement of data envelopment analysis (DEA), to measure the efficiency of the health systems using data on healthy life expectancy and disability adjusted life years as health outcomes. Expenditure on health and education are used as inputs to the health system.

Results The group of high income OECD countries shows the largest indexes of efficiency and also the lowest dispersion. In contrast, low income countries also have the most inefficient health systems, which implies that there are more opportunities for improvement. The average efficiency score is 0.96 for high income countries, 0.83 for upper-middle income countries, 0.86 for lower-middle income countries and only 0.76 for low income countries. Only 17 countries have a score equal to 1 and therefore are completely efficient and can be taken as referents. The index of efficiency is found to be positively associated with government expenditure on health as a percentage of total expenditure on health.

Conclusions The analysis of the results shows that the public share in health expenditure and the weight of health expenditure in public budgets are two factors positively associated with the performance of the health systems. The study also highlights the advantages of using VEA over DEA in the measurement of performance.

Keywords Data envelopment analysis · Value efficiency analysis · Health system performance · Health outcomes

E. González (✉) · J. Ventura
Department of Business Administration, Universidad de Oviedo,
Av. Cristo s/n 33006,
Oviedo, Spain
e-mail: efidalgo@uniovi.es

J. Ventura
e-mail: ventura@uniovi.es

A. Cárcaba
Department of Accounting, Universidad de Oviedo,
Av. Cristo s/n 33006,
Oviedo, Spain
e-mail: acarcaba@uniovi.es

Introduction

The increasing efforts of rich and developing countries to improve the quality and quantity of health services call for an objective and accurate assessment of the performance of their health systems. Both policy makers and citizens demand the best possible outcomes from the health system, given the considerable amount of resources devoted to it. Cumulative evidence shows that rich countries are installed in the diminishing returns zone of the production function in what has been named flat-of-the-curve medicine (Enthoven 1980; Hertzman 1999; Fusch 2004). In these countries, increasingly costly innovations and services can barely achieve modest improvements in the general health level of the population. The debate suggests that redirecting resources from flat-of-the-curve medicine to other programs that would promote healthy habits or education could perhaps make a better contribution to health outcomes in rich countries.

On the other hand, the returns that modest investments in health services can produce in underdeveloped countries can be dramatically important in terms of lives saved, increases in life expectancy and improvements in living conditions. However, the magnitude of these effects also

critically depends on the way in which resources are employed. Basic vaccine programs, for instance, can save thousands of lives in these countries. With extremely limited resources these countries cannot aspire today to many sophisticated and costly health technologies that would have a much lesser impact on the population. Again, the way in which resources are allocated to their most productive uses in terms of improving the general health status of the population may vary greatly from one country to another. Measuring the efficiency with which health resources are transformed into health outcomes in different countries can indicate the countries that can serve as referents and indicate the opportunities for improvement in the other, less efficient countries.

Unfortunately, the measurement of productive efficiency in health care is a far from easy task. Traditionally, the main difficulty has been the correct measurement of the outcome of the health system (Kooreman 1994). The usual approach is to use measurable intermediate indicators of the services provided (outputs), which are assumed to have a fundamental impact on the health status of the population (outcome). Health system outcomes may be defined as those changes in the health status of the population that can be attributed to spending on health care (Häkkinen and Joumard 2007). The World Health Organization database reports information on many variables that conform to this definition (life expectancy, infant mortality, inequality in access, prevalence of certain diseases, etc.) for a broad sample of countries. Although there may be some controversy regarding the appropriateness of some of these variables as relevant outcomes of the health system (Häkkinen and Joumard 2007), most analyses at the system level have relied on the use of life expectancy and infant mortality to approximate the outcomes of the health system (Or 2000; WHO 2001; Retzlaff-Roberts et al. 2004; Afonso and St. Aubyn 2005). However, it can be argued that while infant mortality is a very dramatic problem in underdeveloped countries, it is no longer a relevant issue in most developed countries.

The measurement of health efficiency at the system level of analysis is also subject to a second problem. The health status of a country's population is not only affected by expenditure on health services. Many other factors related to the social, economic and natural environment also play an important role (Naylor et al. 2002). Therefore, influences external to the health system must be accounted for in order to provide accurate estimations of performance. Among these external factors, education has been widely recognized as the main driver of health status (Ross and Wu 1995; Grossman and Kaestner 2004; Grossman 2005; Cutler and Lleras-Muney 2006). Educated people are in a better position to interpret and evaluate information and are therefore able to make better choices that improve and preserve health conditions.

In recent years a variety of research studies have been published that attempt to measure the efficiency of health systems across the world. Most of these restrict the sample to the OECD member countries and apply non-parametric frontier measurement techniques such as data envelopment analysis (DEA) or free disposal hull (FDH) (Retzlaff-Roberts et al. 2004; Räty and Luoma 2005; Afonso and St. Aubyn 2005; Lauer et al. 2004). On the other hand, Evans et al. (2001) relied on parametric frontier estimation techniques and covered a wide sample of 191 countries. The most influential study was carried out by the World Health Organization in the 2000 World Health Report (WHO 2001). In this case the analysis included 191 member countries, and a synthetic index was constructed weighting five dimensions representing the goals of the health systems. In this paper, we apply analytical frontier techniques from the DEA family to measure the performance of the countries included in the WHO sample. In doing so, we use information on health and non-health inputs and health outcomes.

The extreme flexibility of DEA and its ability to handle multiple outputs and inputs in the specification of the production process explains its extensive use in the measurement of health efficiency (Hollingsworth et al. 1999; Puig 2000; Worthington 2004). However, DEA also has some drawbacks that severely limit its application in practice. One of the most important limitations of DEA is its low discriminatory power, especially when many dimensions are taken into account and the sample size is limited (Ali 1994). In these cases, DEA results show a considerable number of efficient DMUs (decision-making units), even though some of them would be considered as low performers under closer inspection of the data. These DMUs obtain a score of 100% simply because they are not comparable to the rest of the sample in one or more dimensions. In fact, the DEA score is a weighted index of inputs and outputs, and there is total flexibility with regard to the weights that can be assigned to each country. In the WHO World Health Report 2000, the five health goal dimensions received fixed weights for all the countries in the sample. DEA does just the opposite: there is complete freedom with regard to the weights assigned to each country, and the country's performance is compared to that which would be achieved by other countries using the same weights. In fact both approaches may seem to be extreme cases. Some degree of flexibility in order to capture differences in specific country goals or values may be desirable (Richardson et al. 2003), but not to the extent of preventing meaningful comparisons.

A recent advance in DEA methodology, namely value efficiency analysis (VEA), provides a way of dealing with this problem, though at the cost of an increased level of analytical complexity. This technique is based on DEA, but

adds a constraint on how input and output weights can be chosen for the different countries within the sample. As a result, VEA significantly improves both the discriminatory power of DEA and the consistency of the weights upon which the evaluation is based. We will use VEA to obtain the final performance scores for the health systems.

Some policy implications can also be derived from the performance indicators obtained for the countries in the sample. In particular, we will focus on the relationship between the commitment of governments towards financing the health system and its performance. While systems dominated by private provision of health services may be more efficient in producing and delivering services, they also can incur higher administrative costs than statutory health insurance (Thomson and Mossialos 2004). In turn, public provision of health can better provide access to basic services for a larger fraction of the population. If this is so, we should find a positive relationship between the estimated indicator of performance and variables that reflect the commitment of public resources to the health system.

The rest of the paper is structured as follows. "Methods" briefly reviews the VEA model as an extension of conventional DEA. "Data" presents the data, and "Results" discusses the empirical results. Concluding remarks are provided in the final section.

Methods

To compute the VEA efficiency scores, we must first obtain the DEA frontier for the countries in the sample. The DEA frontier identifies the countries that would be considered completely efficient under certain (conservative) assumptions. Even though there are many variants of DEA programs, in this paper we follow the traditional specifications of Charnes et al. (1978) for the constant returns to scale frontier (CRS) and Banker et al. (1984) for the variable returns to scale frontier (VRS). For the orientation of the efficiency model, we chose an output orientation because we believe that the main concern of governments and citizens regarding health during recent decades has been to improve the quality and quantity of health services and not cost containment. The CRS DEA model with an output orientation requires solving the following mathematical program for each DMU i in the sample:

$$\min \frac{\sum\limits_{m=1}^{M} v_m x_{im}}{\sum\limits_{s=1}^{S} u_s y_{is}}$$
$$s.t:$$
$$\frac{\sum\limits_{m=1}^{M} v_m x_{jm}}{\sum\limits_{s=1}^{S} u_s y_{js}} \geq 1 \quad , \qquad \forall j \tag{1}$$
$$u_s, v_m \geq 0 \quad , \qquad \forall s, m$$

where x_{im} represents the consumption of input m by DMU i, y_{is} represents the production of output s by DMU i, v_m is the shadow price of input m, and u_s is the shadow price of output s. The program finds the set of shadow prices that minimizes the production cost of unit i with respect to the value of its product, conditioned on obtaining ratios larger than or equal to 1 for all the other DMUs in the sample. If DMU i is efficient, the optimal shadow prices will give the minimum possible value of the ratio, i.e., 1. Inefficiency would be reflected by a value greater than 1 for the objective function. The fractional program (1) entails some computational complexities. Thus, it may be preferable to solve the following equivalent linear program:

$$\min \sum\limits_{m=1}^{M} v_m x_{im}$$
$$s.t:$$
$$\sum\limits_{s=1}^{S} u_s y_{is} = 1 \tag{2}$$
$$\sum\limits_{s=1}^{S} u_s y_{js} - \sum\limits_{m=1}^{M} v_m x_{jm} \leq 0 \quad , \qquad \forall j$$
$$u_s, v_m \geq 0 \quad , \qquad \forall s, m$$

This program finds the shadow prices that minimize the cost of DMU i, but normalizing the output value to 1. If DMU i is efficient it will obtain a cost equal to 1, while if it is inefficient it will obtain a value greater than 1. If DMU i is inefficient then the solution to the linear program must identify another DMU in the sample that obtains the minimum cost of 1 with the shadow prices that are most favourable to DMU i. Program (2) is solved for every DMU in the sample, providing each unit with its most favorable set of shadow prices for inputs and outputs and the corresponding scores of relative efficiency. For ease of interpretation, it is common to use the inverse of the objective function in (2) as the efficiency score. Therefore, the score is bounded within the (0,1] interval, and values lower than 1 reflect the degree of productive inefficiency.

Banker et al. (1984) relax the CRS assumption modifying linear program (2) to allow for VRS:

$$\min \sum\limits_{m=1}^{M} v_m x_{im} + e_i$$
$$s.t:$$
$$\sum\limits_{s=1}^{S} u_s y_{is} = 1 \tag{3}$$
$$\sum\limits_{s=1}^{S} u_s y_{js} - \sum\limits_{m=1}^{M} v_m x_{jm} - e_i \leq 0 \quad , \qquad \forall j$$
$$u_s, v_m \geq 0 \quad , \qquad \forall s, m$$

where the intercept e_i is added to relax the CRS condition that forced the objective function to pass through the origin in (2). In program (3), that condition will only be satisfied if $e_i^* = 0$. For values greater than or smaller than 0, the reference on the frontier for the DMU will be located in a

local zone with decreasing or increasing returns to scale respectively. Most productive activities are subject to variable returns to scale, and health is just one of them. For example, it is relatively easy to achieve large improvements in the health level of the population of poor countries with a very limited expenditure on vaccination campaigns, information campaigns regarding common diseases, etc., so that increasing returns are observed within these countries. The opposite occurs in rich countries. Obtaining additional increases in the health status of the population in rich countries is much more expensive because of decreasing returns (flat-of-the-curve hypothesis). Thus, we consider that the VRS frontier is the most appropriate for evaluating health efficiency in order to avoid scale problems in the efficiency scores due to effects of the large differences in the sizes of the countries. To further avoid scale problems, we will use data that are measured as ratios or per-capita values, which eliminate the size component in the data.

A distinctive feature of DEA is the absolute flexibility in the way the linear program can assign weights (shadow prices) to each particular DMU in the sample. Recall that the program is solved independently for each DMU and that shadow prices for inputs and outputs may then be completely different from one DMU to another. The main argument in defense of the extreme weight flexibility in DEA is that this has the advantage of providing an evaluation of the inefficiency of each DMU under its most favorable scenario. However, extreme flexibility is also the object of criticism because it often produces an extreme inconsistency in the values of the shadow prices across DMUs. To avoid this inconsistency the DEA literature has suggested some solutions to restrict the range of acceptable values for those weights (Thompson et al. 1996; Dyson and Thanassoulis 1988; Allen et al. 1997; Roll et al. 1991; Wong and Beasley 1990; Pedraja et al. 1997; Sarrico and Dyson 2004).

In turn, the problem of weight restriction methods is that they require making value judgments about the range of shadow prices that is considered appropriate. In order to facilitate the implementation of weight restrictions in practice, Halme et al. (1999) proposed an alternative methodology under the name value efficiency analysis (VEA). The objective of VEA is to restrict weights using a simple piece of additional information that must be supplied by an outside expert. The most notable difference between VEA and conventional methods of weight restriction is that instead of establishing appropriate ranges for shadow prices, the expert is simply asked to select one of the DEA-efficient DMUs as his most preferred solution (MPS). Once the MPS is selected, the standard DEA program is supplemented with an additional constraint that forces the weights of the DMU under evaluation (*i*) to make the MPS (*o*) efficient. In other words, the new linear

program requires that the optimal shadow prices selected for DMU *i* must also be good for the MPS in the sense that they ensure that the MPS is on the frontier. As this requirement is made for all the DMUs in the sample, the optimal sets of shadow prices for all the linear programs must make the MPS efficient. The use of the MPS therefore ensures a high degree of consistency in the sets of shadow prices across DMUs. An immediate effect of the VEA constraint is that DMUs that obtained a DEA score of 1 simply because they had an extreme value in one input or output will only obtain a VEA score equal to 1 if they can withstand the additional comparison with the MPS.

The VRS VEA program with an output orientation can be expressed as follows:

$$
\begin{aligned}
&\min \sum_{m=1}^{M} v_m x_{im} + e_i \\
&s.t: \\
&\sum_{s=1}^{S} u_s y_{is} = 1 \\
&\sum_{s=1}^{S} u_s y_{is} - \sum_{m=1}^{M} v_m x_{jm} - e_i \leq 0 \quad , \quad \forall j \\
&\sum_{m=1}^{M} v_m x_{om} + e_i - \sum_{s=1}^{S} u_s y_{os} = 0 \\
&u_s, v_m \geq 0 \quad , \quad \forall s, m
\end{aligned}
\tag{4}
$$

Program (4) is identical to program (3) with the MPS constraint added. Thus, the MPS (*o*) must obtain a value of 1 with the shadow prices of DMU (*i*). Indirectly, this requirement restricts the permissible range of shadow prices to the range that makes the MPS (*o*) efficient. We used the software LINGO to solve the DEA and VEA programs in our study. While many packages are pre-programmed to solve DEA, we are not aware of any that can solve VEA. However, any mathematical programming software can be used.

A controversial issue in VEA is how to select the MPS (Korhonen et al. 1998). Our empirical setting is designed to evaluate the efficiency of countries regarding their health achievements. In this context, it would be difficult to find an expert that could provide the MPS. We propose a new alternative method that avoids supplying any external information. We will run various VEA analyses considering each of the countries that are DEA efficient as the MPS. Then, we will compute the average reduction in the variation of the shadow prices of the variables included in the VEA model. The country that achieves the greatest reduction in the variation of shadow prices will be our selected MPS. This approach has two advantages. First, it is objective and does not require the implication of an outside expert. Second, it obtains the greatest possible congruence in the shadow prices from the set of linear programs that are computed to calculate the value efficiency scores. This seems highly consistent with the objective we were pursuing with the implementation of VEA instead of DEA.

Data

We are interested in measuring value efficiency scores for all the countries in the WHO database for the year 2004. However, the required data were not available for some of the countries in the database, and therefore they were excluded from the analysis. The final sample includes a total of 165 countries, which is sufficiently large for the techniques we will use and representative of the vast majority of the world's population. While we followed the existing literature in order to choose the variables that could reasonably approximate the relevant dimensions of health production at the system level of analysis, we are innovative in the precise specification of the outputs.

To approximate the output of the health system, the specification that has been common in previous empirical research has relied on life expectancy and infant mortality rates (Retzlaff-Roberts et al. 2004; Afonso and St. Aubyn 2005; Räty and Luoma 2005). Life expectancy is a variable that is widely available and can be considered as a long run global result of the health system of a country (Evans et al. 2001). Countries with poor access to health services, poor quality of health care centers and physicians and low expenditure on medicines will, in general, have a low average life expectancy. In contrast, it is reasonable to assume that many (though not all) health care services should translate into a higher life expectancy for the population. However, it could be argued that the living conditions of these additional years of life should also be taken into account when specifying the goals of a health system. Thus, adding life to years may emerge as a more reasonable health system goal. We believe that the variable Healthy Life Expectancy provided by the WHO for the year 2002 covers both aspects, namely adding years to life and life to years, in that it measures the expectancy of life in good living conditions.

The second output variable that has been widely used in previous literature is infant mortality. This variable is an outcome of a health system and also an indicator of inequality in access to resources. However, we believe that this variable is not relevant for many developed countries, since infant mortality is no longer an issue in those countries. Instead, we propose using a more general variable that is relevant for all the countries of the sample and that indirectly captures the effects of infant mortality in those countries in which it is an important problem. The variable we will use is provided by the WHO for the year 2004 with the name Disability Adjusted Life Years (DALY). It is a measure of the years lost due to premature death and also includes the equivalent years lost due to disability. In particular, we use the age-standardized DALY rates per 100 inhabitants. Obviously, this variable must be transformed in order to use it as an output in a DEA

specification (i.e., a larger value must indicate a better performance). We therefore take the inverse of the variable and multiply by 100 for ease of interpretation. This transformed variable can be interpreted as the number of people in the population that corresponds to the loss of a year of life to disability or premature death. For example, Ghana has a value of 33.3 for the original DALY variable, which means that 33.3 years are lost to death or disability per 100 members of the population. Our inverse variable takes the value 3, which means that there is 1 year of life lost to death or disability for every three people in the population. The higher this value is, the better the performance of the health system. While in the DEA and VEA models we will use this transformed variable, in the tables summarizing results we will use the original DALY variable in order to improve the readability of the paper.

The input variables should capture the magnitude of the resources committed to health production services and other environmental or social factors that influence the health status of the population. The resources devoted to the health system are approximated by per capita total expenditure in health in purchasing power parity of $US (PPP). Using PPP expenditure facilitates cross-country comparisons (Gupta and Verhoeven 2001). However, health resources are not the only input involved in the health production process. It has been argued that the social environment greatly influences the health status of the population (Naylor et al. 2002). An appropriate and available proxy for the beneficial influences of the social environment is the level of education. The basic argument is that a person that is better educated would make healthier choices (Cowell 2006). Additionally, education is an excellent proxy for other social dimensions that may influence health (nutrition, hygiene, use of health services, working conditions, etc.). As an indicator of the level of education we use the School Life Expectancy (Years). This variable was taken from the UNESCO online database.

Given the enormous disparities that exist among countries, we decided not to pool all the countries of the sample under the same production frontier. Instead we separated the sample into four sub-samples of countries based on the degree of development of the country. We followed the classification made by the World Bank that separates countries into four groups according to gross national income per capita. The groups are low income ($935 or less), lower-middle income ($936–$3,705), upper-middle income ($3,706–$11,455) and high income ($11,456 or more). In order to provide a more accurate comparison of countries, high income non-OECD countries were included in the group of upper-middle income countries. This subgroup of countries is composed of those such as Brunei, Bahrain, Barbados, Cyprus, Kuwait, etc., that are more similar to the group of upper-middle income

countries in terms of the variables used as an input in our DEA model (education and per capita expenditure in health). Indeed, they are not comparable to the group of OECD high income countries in terms of health expenditure or health results.

Table 1 succinctly describes the variables used as inputs and outputs in our empirical analysis for the four groups of countries. It is clear from the table that there is great variation across countries in every dimension of the health production model, as reflected by the standard deviations (SD). Perhaps the most striking differences concern the DALYs, which represent the years lost to premature death or disability per 100 members of the population. The minimum of eight corresponds to Japan and reflects the situation of most developed countries in the high income group. In contrast, Zimbabwe with 82.8 is the most extreme example of an underdeveloped country in the low income group that also suffers the effects of violence. If we look at an input dimension we obtain a similar picture. The Democratic Republic of the Congo is the country that spends less in health per capita ($15 PPP) and is representative of the situation in most underdeveloped countries, most of which are in Africa and Asia. On the opposite extreme, the USA spends $6,014 PPP followed by Luxembourg with $5,317 PPP, representing the situation in the developed and rich part of the world. However, these figures do not guarantee that this money is being spent efficiently on the production of relevant health outcomes in

any of these countries. The relationship between resources and outcomes is what gives us the indicator of efficiency.

Results

The DEA model was run separately for each of the four groups of countries to obtain an initial efficiency frontier for each group. This is a necessary step to identify which countries are located on the frontier and can thus be considered as candidates to be the MPS for the VEA analysis. Table 2 summarizes the DEA results for the 165 countries in the four income groups considered, further classified into broad geographical areas. High income OECD countries show the largest average efficiency and also the lowest dispersion. In contrast, low income countries have the lowest average and the largest dispersion. These results show that the health systems in high income countries have more similarities among them than the health systems in low income countries. Therefore, when comparisons are made within each group, the much larger dispersion for low income countries also reflects larger distances to the best practice frontier. This means that there is more room for improvement in low income countries than in high income countries, a finding consistent with the flat-of-the-curve hypothesis.

With respect to geographical areas, European countries are among the most efficient on average, closely followed by American, Oceanic and Asian countries. Since most low

Table 1 Descriptive statistics of inputs and outputs

	Average	Min	Max	SD
High income countries				
Total expenditure in health PPP ($)	2,830.2	1,058	6,014	1,137.8
School life expectancy	16.1	13.6	20.4	1.5
Healthy life expectancy (2002)	70.9	64.9	75	2.2
Age std. DALYs per 100	10.7	8	15	1.5
Upper-middle income countries				
Total expenditure in health PPP ($)	724.4	236	4,457	646
School life expectancy	13.4	9.6	16.5	1.5
Healthy life expectancy (2002)	62.1	35.7	71.4	6.39
Age std. DALYs per 100	18.6	9.7	53.4	9
Lower-middle income countries				
Total expenditure in health PPP ($)	251.6	31	588	149.5
School life expectancy	11.1	3.68	14.3	2.3
Healthy life expectancy (2002)	55.7	31.4	64.4	8.3
Age std. DALYs per 100	24.8	15.5	65	10.8
Low income countries				
Total expenditure in health PPP ($)	70	15	188	43.5
School life expectancy	7.8	3.5	12.4	2
Healthy life expectancy (2002)	44.5	28.6	61.4	8
Age std. DALYs per 100	41.5	17	82.8	14.4

Table 2 Summary of DEA results by income and geographical areas

	n	Average	Min	Max	SD	Efficient
High income	27	0.964	0.923	1.00	0.02	4
Asia	2	1.00	1.00	1.00	0.0	2
Europe	21	0.963	0.931	1.00	0.02	2
North America	2	0.942	0.924	0.961	0.03	0
Oceania	2	0.967	0.965	0.968	0.00	0
Upper-middle income	50	0.928	0.530	1.00	0.08	6
Africa	7	0.802	0.530	0.973	0.16	0
Asia	13	0.949	0.858	1.00	0.05	3
Europe	11	0.955	0.916	1.00	0.02	1
North America	1	0.968	–	–	–	0
Latin America and Caribbean	17	0.946	0.883	1.00	0.04	2
Oceania	1	0.900	–	–	–	0
Lower-middle income	45	0.924	0.536	1.00	0.10	8
Africa	13	0.872	0.536	1.00	0.17	4
Asia	14	0.946	0.845	1.00	0.04	3
Europe	4	0.968	0.920	0.992	0.03	0
Latin America and Caribbean	10	0.936	0.875	0.989	0.04	0
Oceania	4	0.941	0.838	1.00	0.07	1
Low income	43	0.843	0.537	1.00	0.13	8
Africa	31	0.811	0.537	1.00	0.13	5
Asia	11	0.922	0.798	1.00	0.08	2
Oceania	1	1.00	–	–	–	1
Total	165	0.911	0.530	1.00	0.10	26

income countries are located in Africa, it is not surprising that this continent is the least efficient in terms of health attainment. It is noticeable though that more than one third of the DEA-efficient countries are also located in Africa. On the other hand, we were not able to find any efficient country within the North American area. It can also be noticed that the standard deviation is very high in Africa, whereas it is moderate in Europe and America. It is also noticeable that all the Asian countries included in the high income group (Japan and Korea) are on the frontier, whereas the other two frontier countries are European (Luxembourg and Slovakia).

Overall, the minimum efficiency score is obtained by Botswana in Southern Africa. With 35.7 years of Healthy Life Expectancy and 53.4 DALYs per 100 members of the population, this country obtains an index of just 0.53, which means that an improvement of 88% in health outcomes could be achieved with a better use of resources. In this particular case, this improvement would add 31.6 more years of healthy life and avoid 25.1 DALYs per 100 members of the population.

A total of 26 countries in the sample obtain a DEA score equal to 1, which means they cannot make any (relative) improvement, given the data observed and the structure of the DEA programs that generate the best practice reference frontier. Some of these are countries with favorable health outcomes, given the resources deployed in health production

and their standards of comparison. However, other countries are on the frontier simply because DEA is very flexible in evaluating countries with extreme data. These countries may be assigned unreasonable weights to inputs and/or outputs in the DEA program to reach complete DEA efficiency. For example, the Democratic Republic of the Congo is the country with the lowest per capita expenditure on health, and this makes it DEA-efficient regardless of its Healthy Life Expectancy or DALY figures. The DEA program will assign an extremely high weight to per capita expenditure on health, and no other country could then be comparable to the Democratic Republic of the Congo.

In our view, the presence of these "extreme data" countries on the DEA frontiers only provides evidence for the important limitations of DEA. Many countries with poor results are considered efficient simply because there is no other country that does better in some dimension of the production setting. In other words, the flexibility of the weights allows some countries to be assigned a very low value in those dimensions in which they perform poorly and a high value in those dimensions in which they perform better. As we noticed before, the Democratic Republic of the Congo achieves full DEA efficiency because of the extremely high weight given to health expenditure. It would not matter if this country reduced its life expectancy by one half: it would still be DEA-efficient.

To increase the discriminatory power of DEA and achieve a higher degree of congruence in the shadow prices assigned to the different countries in the DEA programs, we solved 26 VEA programs using as MPS each of the 26 countries appearing on the DEA-efficient frontiers. As before, the analysis was carried out separately for each income group. For each VEA analysis we computed the coefficient of variation of the weights of each input and output and the average coefficient of variation. The country that achieved the lowest average was taken as the MPS. These countries were Japan for the high income group, Oman for the upper-middle income group, Algeria for the lower-middle income group and the Solomon Islands for the low income group.

The VEA constraints produce more coherent results in terms of the weights selected for each country in the DEA program to justify its index of efficiency. The reductions in the coefficients of variation of the weights within the different groups of countries are 15% (high income), 77% (upper-middle income), 70% (lower-middle income) and 69% (low income). The reductions are notable in all groups except for the group of high income countries. Also, the discriminatory power of DEA is enhanced with the VEA specification. Of the 26 countries that form the DEA frontiers, only 17 remain after the VEA constraints are added to the linear programs. Again, the gains in discriminatory power are higher for the three groups with lower income. The group of high income countries maintains its four DEA-frontier countries on the VEA frontier. This means that within the high income group there is less internal variation, and adding a constraint to add coherence to the weights is of little help since the internal coherence of the DEA results was already very high. In fact, in the group of high income countries there are no differences in the performance scores obtained with DEA and VEA. In contrast, VEA significantly improves the results of the DEA analysis in the other three groups of countries.

A summary of the VEA results is shown in Table 3. As we mentioned above, the number of efficient countries drops from 26 (DEA) to 17 (VEA). This means that only 17 countries in the sample are efficient when using weights that are reasonable for the MPS of their groups. To see how unreasonable some DEA results can be, the VEA score for the Democratic Republic of the Congo is just 0.74, whereas it was completely efficient under the DEA program for the simple reason that it had the lowest per capita expenditure on health. Other countries such as Ethiopia or Niger also fall from complete DEA efficiency to VEA scores under 0.8. The least VEA-efficient country is Swaziland with a score of 0.47, followed by Zimbabwe (0.49) and Sierra Leone (0.49). Europe is no longer the most VEA-efficient geographical area. When consistency in the weights is

required with the VEA program, the average efficiency of European countries in the upper and lower-middle income groups falls dramatically. The leading edge of health system efficiency in the upper-middle income group corresponds to American and Asian countries, while Oceania obtains the highest average in the lower-middle income group. Africa remains the most inefficient area in the three groups in which it has representative countries. The standard deviation is also high in Africa and Asia, while it remains moderate in the rest of the world. The complete results for the 165 countries are provided in Tables 4, 5, 6 and 7.

Now that we have estimated the value of the efficiency scores for the entire sample, we are interested in testing whether there is some relationship between the implication of the public sector in the provision and financing of health services and the global performance of the system. For this purpose, we have compiled data on two variables from the WHO database, namely the "general government expenditure on health as percentage of total expenditure on health" (GIMP) and the "general government expenditure on health as percentage of total government expenditure" (HREL). We refer to these variables as government implication in health financing (GIMP) and health relevance in public budgets (HREL).

Table 8 shows the averages of the variables GIMP and HREL for the six geographical zones considered and also the rates of growth of these variables as well as the growth of per capita expenditure in health (PCEXP) over the period 1995–2004. There is considerable variation among the countries in the sample with respect to these variables and their evolution on time. The increase in per capita expenditure on health (PCEXP) is highest in Europe (93%), followed by Asia (75%), Africa (68%) and North America (64%). The growth in Latin America and the Caribbean and Oceania was moderate at 52%. With respect to the variable GIMP, Oceania and Europe are the zones with the largest proportion of the health system publicly financed (78% and 71%, respectively). In contrast, Asia and Africa, with 50% and 51% respectively, are the zones with the lowest public financing. In America the proportion varies little from the 53% of North America to the 55% of Latin America and the Caribbean.

Somewhat surprisingly, however, the weight of health in public budgets (HREL) is greatest in North America (16%), followed by Europe and Oceania (both with 13%). Again, these figures contrast with the low figures of 8% and 9% in Asia and Africa. While this was the situation in 2004, the evolution since 1995 shows large increases in GIMP, especially in Africa and Oceania where these figures rose by 18% and 7%, respectively. The growth of GIMP was negative in Europe (−3%) and remained almost unchanged in the rest of the world. Finally, the weight of health in public budgets (HREL) rose significantly in all zones

Table 3 Summary of VEA results by income and geographic areas

	n	Average	Min	Max	SD	Efficient
High income	27	0.964	0.923	1.00	0.02	4
Asia	2	1.00	1.00	1.00	0.0	2
Europe	21	0.963	0.931	1.00	0.02	2
North America	2	0.942	0.924	0.961	0.03	0
Oceania	2	0.967	0.965	0.968	0.00	0
Upper-middle income	50	0.835	0.523	1.00	0.10	4
Africa	7	0.729	0.523	0.855	0.13	0
Asia	13	0.870	0.705	1.00	0.09	2
Europe	11	0.803	0.684	0.873	0.07	0
North America	1	0.883	–	–	–	0
Latin America and Caribbean	17	0.870	0.737	1.00	0.08	2
Oceania	1	0.823	–	–	–	0
Lower-middle income	45	0.862	0.469	1.00	0.13	4
Africa	13	0.769	0.470	1.00	0.18	1
Asia	14	0.901	0.729	1.00	0.09	2
Europe	4	0.890	0.816	0.992	0.08	0
Latin America and Caribbean	10	0.890	0.797	0.970	0.05	0
Oceania	4	0.930	0.802	1.00	0.09	1
Low income	43	0.763	0.495	1.00	0.14	5
Africa	31	0.725	0.495	1.00	0.12	2
Asia	11	0.849	0.591	1.00	0.12	2
Oceania	1	1.00	–	–	–	1
Total	165	0.845	0.470	1.00	0.13	17

except Latin America and the Caribbean, where this figure dropped by 5%. Again, the increases were especially notable in Africa (27%) and Oceania (37%).

From these figures we can conclude that there has been an evolution towards more public participation in health financing across the world in recent years, except perhaps in Europe and Latin America and the Caribbean. It would

therefore be interesting to show the relationship between these variables (GIMP and HREL) and efficiency in the use of resources as measured by VEA scores. Before proceeding, it should be warned that regression analysis is not an appropriate statistical tool to test the relationship between DEA or VEA scores and possible explanatory variables. The reason is that DEA or VEA scores are not normally distributed

Table 4 Complete VEA and DEA results: high income countries

Country	VEA score	VEA rank	DEA score	DEA rank	Country	VEA score	VEA rank	DEA score	DEA rank
Australia	0.968	11	0.968	11	Japan	1.000	1	1.000	1
Austria	0.952	19	0.952	19	Luxembourg	1.000	1	1.000	1
Belgium	0.948	22	0.948	22	The Netherlands	0.949	21	0.949	21
Canada	0.961	13	0.961	13	New Zealand	0.965	12	0.965	12
Czech Republic	0.996	5	0.996	5	Norway	0.960	15	0.960	15
Denmark	0.931	25	0.931	25	Portugal	0.956	18	0.956	18
Finland	0.958	16	0.958	16	Rep. Korea	1.000	1	1.000	1
France	0.960	14	0.960	14	Slovakia	1.000	1	1000	1
Germany	0.958	17	0.958	17	Spain	0.987	6	0.987	6
Greece	0.947	23	0.947	23	Sweden	0.978	7	0.978	7
Hungary	0.949	20	0.949	20	Switzerland	0.976	8	0.976	8
Iceland	0.971	9	0.971	9	UK	0.942	24	0.942	24
Ireland	0.931	26	0.931	26	USA	0.924	27	0.924	27
Italy	0.969	10	0.969	10					

Table 5 Complete VEA and DEA results: Upper-middle income countries

Country	VEA score	VEA rank	DEA score	DEA rank	Country	VEA score	VEA rank	DEA score	DEA rank
Argentina	0.737	45	0.932	31	Libya	0.764	39	0.973	14
Bahrain	0.797	34	0.939	28	Lithuania	0.741	42	0.932	32
Barbados	0.765	38	0.947	26	Malaysia	0.877	15	0.950	24
Belarus	0.778	37	0.916	35	Malta	0.848	26	1.000	1
Belize	0.822	32	0.914	37	Mauritius	0.833	29	0.933	30
Botswana	0.523	50	0.530	50	Mexico	0.883	14	0.968	17
Brazil	0.740	43	0.883	44	Oman	1.000	1	1.000	1
Brunei	0.894	13	0.984	10	Panama	0.894	12	0.985	9
Bulgaria	0.866	22	0.960	22	Poland	0.813	33	0.965	20
Chile	0.902	11	0.998	7	Qatar	0.907	10	0.967	19
Costa Rica	1.000	1	1.000	1	Romania	0.873	17	0.952	23
Croatia	0.870	19	0.967	18	Russia	0.764	40	0.875	45
Cyprus	0.871	18	0.976	12	St. Kitts and Nevis	0.868	20	0.916	36
Dominica	0.874	16	0.964	21	Saint Lucia	0.864	23	0.950	25
Eq. Guinea	0.838	28	0.838	47	St. Vinc. and Grenada	0.919	9	0.923	33
Estonia	0.751	41	0.944	27	Saudi Arabia	0.847	27	0.914	38
Fiji	0.823	31	0.900	40	Seychelles	0.855	25	0.905	39
Gabon	0.701	47	0.786	48	Slovenia	0.684	48	0.973	13
Grenada	0.867	21	0.889	43	South Africa	0.587	49	0.653	49
Israel	0.790	35	1.000	1	Suriname	0.856	24	0.896	42
Jamaica	1.000	1	1.000	1	Trinidad and Tobago	0.933	7	0.935	29
Kazakhstan	0.705	46	0.858	46	Turkey	0.986	5	0.986	8
Kuwait	1.000	1	1.000	1	United Arab Emirates	0.922	8	0.969	16
Latvia	0.739	44	0.922	34	Uruguay	0.782	36	0.971	15
Lebanon	0.827	30	0.898	41	Venezuela	0,961	6	0.981	11

as they are bounded by one. More importantly, the scores are not IID. Therefore, non-parametric rank-based tests are preferable. For further discussion about the application of non-parametric rank-based statistics to efficiency scores, see Brockett and Golany (1996) and Sueyoshi and Aoki (2001).

We ranked countries based on their VEA score and then assigned them to five efficiency groups of equal size (N=33). Then, we used the Kruskal-Wallis non-parametric test (H-KW) to test for the existence of significant differences across the five efficiency groups with respect to the variables that reflect the government role in the health system. Table 9 shows the average values for the VEA score and the variables of government implication (GIMP) and health relevance in public budgets (HREL). It is clear from the table that the most efficient groups of countries also have a higher government implication in financing health services. In the most efficient countries (group 5), an average of 63.8% of health expenditure is public expenditure, which means that almost two thirds of the system has a public basis. In contrast, low efficiency countries show an average of public expenditure on health of around 50%. The differences across groups are statistically significant at the 0.1 level. The

relevance of health expenditure within the government budget seems also to be positively associated with the performance of the health system. The most efficient countries dedicate averages of above 11% of the government budget to the health system, while this figure drops to 10% or 9% in countries with lower efficiency scores. However, these differences are not statistically significant at conventional levels. All in all, the results suggest that having a high percentage of the health system publicly financed also creates a more efficient system. Similar results are obtained if we use the DEA scores instead of the VEA scores to construct the five efficiency groups.

Concluding remarks

This paper provides additional evidence on the lack of discriminatory power of DEA when the weights of inputs and outputs in the linear programs are freely selected for DMUs. There are three ways to improve the discriminatory power of DEA. First, the simplest procedure is to reduce the number of input-output dimensions to be considered in

Table 6 Complete VEA and DEA results: Lower-middle income countries

Country	VEA score	VEA rank	DEA score	DEA rank	Country	VEA score	VEA rank	DEA score	DEA rank
Albania	0.917	19	0.976	14	Jordan	0.782	37	0.948	26
Algeria	1.000	1	1.000	1	Kiribati	0.802	35	0.838	41
Angola	0.585	43	0.929	31	Lesotho	0.518	44	0.536	45
Armenia	0.957	11	0.974	16	Maldives	0.729	40	0.898	36
Azerbaijan	0.947	14	0.955	24	Mongolia	0.934	16	0.934	28
Bhutan	0.914	20	0.958	23	Morocco	0.955	12	0.992	10
Bolivia	0.865	28	0.882	38	Namibia	0.586	42	0.684	43
Cameroon	0.706	41	0.756	42	Nicaragua	0.970	8	0.989	11
Cape Verde	0.978	7	0.986	12	Paraguay	0.928	17	0.962	20
China	0.995	5	1.000	1	Peru	0.954	13	0.963	19
Colombia	0.797	36	0.966	17	Philippines	0.946	15	0.959	22
Congo	0.820	33	1.000	1	Rep. Moldova	0.992	6	0.992	9
Djibouti	0.756	39	1.000	1	Samoa	0.958	10	0.961	21
Dominican Rep.	0.874	25	0.926	33	Sudan	0.848	29	1.000	1
Egypt	0.904	22	0.922	34	Swaziland	0.470	45	0.562	44
El Salvador	0.870	27	0.929	32	Thailand	0.897	23	0.933	29
Georgia	1.000	1	1.000	1	Macedonia	0.837	32	0.986	13
Guatemala	0.881	24	0.933	30	Timor-Leste	0.840	31	0.845	40
Guyana	0.841	30	0.875	39	Tonga	0.961	9	0.964	18
Honduras	0.919	18	0.937	27	Tunisia	0.873	26	0.974	15
India	0.913	21	0.950	25	Ukraine	0.816	34	0.920	35
Indonesia	1.000	1	1.000	1	Vanuatu	1.000	1	1.000	1
Iran	0.767	38	0.895	37					

the model specification. The cost of this approach is that information that may be relevant to discrimination is overlooked. Second, the sample size may be increased. Theoretically, this would be the best solution although, unfortunately, it may be not feasible (in practice) when the researcher is working with complete populations, as is often the case. A third approach is to improve the discriminatory power of the model by supplying some additional information on how the discrimination should be carried out. Value efficiency analysis (VEA) was developed in order to easily incorporate a piece of qualitative information into the DEA specification. This information corresponds to the identification of a most preferred solution (MPS) that acts as an ideal weighting reference in the eyes of an expert. Our results show that VEA significantly increases the discriminatory power of DEA and achieves congruence in the weights of inputs and outputs.

The paper applied both DEA and VEA methodologies to health data on a sample of 165 countries during the year 2004. The sample includes all the countries for which we were able to compile the required data on inputs and outputs. Our sample comprises nearly the whole population of countries in the world (around 86%). Thus, it is not possible to significantly improve the discriminatory power

of DEA by increasing the sample size. The DEA scores show moderately high levels of efficiency in health provision, with an average of 0.91. However, VEA analysis shows an average of only 0.84 when consistency in shadow prices is forced into the measurement model. By simply incorporating information on an efficient country that is considered as an appropriate general referent for the weights (MPS) within each group of countries, VEA notably increases the discriminatory power of DEA.

From 26 DEA-efficient countries, we obtain just 17 VEA-efficient referents. What is happening is that VEA allows a simple identification of the countries whose DEA score is based on unrealistic values for the shadow prices of inputs and outputs. These countries (Democratic Republic of the Congo or Ethiopia, for instance) benefit from the extreme flexibility of DEA, but do not withstand any further analysis of their activity data. For example, a DMU may obtain a DEA score of 1 simply because it is the unit that produces the largest quantity of an output, thus assigning a very large weight to that variable. VEA does not allow this extreme flexibility with regard to weights. Behavior must be globally acceptable, and the MPS indicates what is considered as globally acceptable behavior in terms of weighting inputs and outputs.

Table 7 Complete VEA and DEA results: low income countries

Country	VEA score	VEA rank	DEA score	DEA rank	Country	VEA score	VEA rank	DEA score	DEA rank
Afghanistan	0.591	38	0.819	25	Mali	0.679	31	0.712	36
Bangladesh	0.929	6	0.972	11	Mauritania	0.795	17	0.815	27
Benin	0.754	23	0.800	28	Mozambique	0.643	35	0.677	38
Burkina Faso	0.753	24	0.910	16	Nepal	0.851	11	0.919	13
Burundi	0.647	34	0.913	15	Níger	0.749	25	1.000	1
Cambodia	0.726	29	0.798	29	Nigeria	0.674	33	0.753	33
Chad	0.732	27	0.783	31	Pakistan	1.000	1	1.000	1
Comoros	1.000	1	1.000	1	Rwanda	0.622	36	0.677	37
Côte d'Ivoire	0.698	30	0.751	34	San Tome and Principe	0.819	15	0.918	14
D. R. Congo	0.743	26	1.000	1	Senegal	0.871	9	0.898	18
Eritrea	1.000	1	1.000	1	Sierra Leone	0.496	42	0.537	43
Ethiopia	0.780	19	1.000	1	Solomon I.	1.000	1	1.000	1
Gambia	0.856	10	0.902	17	Tajikistan	0.910	7	0.984	10
Ghana	0.846	12	0.893	19	Togo	0.730	28	0.796	30
Guinea	0.777	20	0.815	26	Uganda	0.617	37	0.725	35
Guinea-Bissau	0.761	22	0.844	23	Tanzania	0.764	21	0.868	22
Kenya	0.679	32	0.780	32	Uzbekistán	0.904	8	0.990	9
Kyrgyzstan	0.832	13	0.956	12	Viet Nam	1.000	1	1.000	1
Lao PDR	0.786	18	0.832	24	Yemen	0.808	16	0.870	21
Liberia	0.557	40	0.645	39	Zambia	0.590	39	0.638	40
Madagascar	0.824	14	0.890	20	Zimbabwe	0.495	43	0.569	42
Malawi	0.544	41	0.623	41					

Table 8 Trends in government financing of health throughout the world

	2004		Growth 1995–2004		
	GIMP	HREL	GIMP	HREL	PCEXP
Africa	51.5	9.2	18.6	27.3	68.9
Asia	50.0	7.9	−0.6	7.6	75.8
Europe	71.2	13.5	−3.3	19.9	93.0
North America	53.8	16.3	2.5	18.6	64.7
Latin America and Caribbean	55.2	1,136	−0.3	−5.4	52.7
Oceania	78.3	13.3	7.7	37.2	52.1
Total	57.4	10.6	5.2	15.9	72.3

Table 9 Government role in the health system and performance

	Group 1	Group 2	Group 3	Group 4	Group 5	H-KW
N	33	33	33	33	33	
VEA score	0.64	0.79	0.87	0.94	0.99	
GIMP	50.4	56.6	57.9	58.2	63.8	2.01[a]
HREL	10.5	9.3	10.3	11.0	11.7	1.49

[a] Significance level 0.1

We followed an innovative approach to objectively selecting the MPS. This involves the estimation of VEA scores using each of the 26 DEA-efficient countries as the MPS. Then, the reduction in the average dispersion of the weights of inputs and outputs within each group of countries is computed, and the country that achieves the highest reduction in dispersion is selected as the MPS for that group. Using this method, Japan (high income), Oman (upper-middle income), Algeria (lower-middle income) and Solomon Islands (low income) were selected as the MPSs of the four groups of countries. They achieve reductions of 15%, 77%, 70% and 69% in the coefficients of variation of the weights in their respective groups of countries. Thus, the improvement in the discriminatory power of VEA is obtained through a more rational selection of weights in the mathematical programs.

A look at the scores shows that high income countries lead the efficiency edge of health provision in terms of efficiency, while low income countries have the lowest efficiency scores on average. By geographical regions, North America and Oceania achieve the highest averages in the VEA scores, while Africa shows the poorest results. There are, of course, rich countries that also show important inefficiencies. The USA, for instance, has a score of just 0.92, which means that an 8% improvement in health outcomes could be achieved without increasing resources deployed to the health system. Denmark, the UK and the Netherlands are other examples of rich countries with VEA scores below 0.95. This means that considerable resources that are devoted to health do not have the desired impact on health outcomes within these countries. The results are consistent with the flat-of-the-curve medicine hypothesis that predicts the moderate marginal impact on average health outcomes from additional investments in health in rich countries. However, the most worrying fact is the confirmation that poor countries with poor health outcomes are also the countries that use the scant resources they devote to health in the most inefficient manner, especially in Africa.

The role of governments in financing the health system is a controversial issue. It is commonly stated that private health insurance tends to incur higher management and administrative costs than statutory health insurance (Thomson and Mossialos 2004). The need to generate a profit is another opportunity cost that public systems do not incur. Our results partially support public financing of health services. The most efficient countries in our sample have around 64% of the health system publicly financed, while the least efficient countries barely reach 50%. The weight of health in the government budget is also positively associated with efficiency, but in this case the association is not statistically significant at conventional levels. However, we can conclude that the countries in which governments show a deeper commitment to the development and financing of the health system also use the resources more efficiently in the achievement of relevant health outcomes.

References

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

Ali AI (1994) Computational aspects of data envelopment analysis. In: Charnes A, Cooper WW, Lewin AY, Seiford LM (eds) DEA: theory, methodology and applications. Kluwer Academic Publishers, Boston, pp 63–88

Allen R, Athanassopoulos R, Dyson G, Thanassoulis E (1997) Weights restrictions and value judgements in data envelopment analysis: evolution, development, and future directions. Ann Oper Res 73:13–34

Banker RD, Charnes A, Cooper WW (1984) Some models for estimating technical and scale inefficiencies. Manage Sci 39:1261–1264

Brockett PL, Golany B (1996) Using rank statistics for determining programmatic efficiency differences in data envelopment analysis. Manage Sci 42(3):466–472

Charnes A, Cooper WW, Rhodes E (1978) Measuring the efficiency on decision making units. Eur J Oper Res 2:429–444

Cowell AJ (2006) The relationship between education and health behavior: some empirical evidence. Health Econ 15:125–146

Cutler DM, Lleras-Muney A (2006) Education and health: evaluating theories and evidence. NBER Working Paper N. 12352

Dyson RG, Thanassoulis E (1988) reducing weight flexibility in data envelopment analysis. J Oper Res Soc 6:563–576

Enthoven AC (1980) Health plan: the only practical solution to the soaring cost of medical care. Addison-Wesley, Reading, Massachusets

Evans DB, Tandon A, Murray CJL, Lauer JA (2001) Comparative efficiency of national health systems: cross national econometric analysis. Br Med J 323:307–310

Fusch VR (2004) Perspective: more variation in the use of care, more flat of the curve medicine. Health Aff 23:104–107

Grossman M, Kaestner R (2004) The effects of education on health. In: Behrman JR, Stacey N (eds) The social benefits of education. University of Michigan Press, Michigan, pp 69–113

Grossman M (2005) Education and nonmarket outcomes. NBER Working Paper N. 11582

Gupta S, Verhoeven M (2001) The efficiency of government expenditure. Experiences from Africa. J Policy Model 23:433–467

Häkkinen U, Joumard I (2007) Cross-country analysis of efficiency in oecd health care sectors: options for research. OECD Economics Department Working Papers N. 554. OECD Publishing, Paris

Halme M, Joro T, Korhonen P, Salo S, Wallenius J (1999) A value efficiency approach to incorporating preference information in data envelopment analysis. Manage Sci 45(1):103–115

Hertzman C (1999) The social dimensions of public health. In: World Social Science Report, UNESCO Publishing, London, pp 341–351

Hollingsworth B, Dawson PJ, Maniadakis N (1999) Efficiency measurement of health care: a review of nonparametric methods and applications. Health Care Manage Sci 2:161–172

Kooreman P (1994) Nursing home care in the Netherlands: a nonparametric efficiency analysis. J Health Econ 13(3):301–316

Korhonen P, Siljamäki A, Soismaa M (1998) Practical aspects of value efficiency analysis. Interim Report IR-98-42, International Institute for Applied Systems Analysis

Lauer JA, Lovell CAK, Murray CJL, Evans DB (2004) World health system performance revisited: the impact of varying the relative importance of health system goals. BMC Health Serv Res 4:19

Naylor CD, Iron K, Handa K (2002) Measuring health performance: problems and opportunities in the era of assessment and accountability. In: Smith PC (ed) Measuring up: improving health system performance in OECD countries. OECD Publishing, Paris, pp 13–34

Or Z (2000) Determinants of health care on mortality across OECD countries: a pooled, cross-country, time-series analysis. OECD Economic Studies Nº 30. OECD Publishing, Paris

Pedraja F, Salinas J, Smith P (1997) On the role of weight restrictions in data envelopment analysis. J Prod Anal 8:215–230

Puig J (2000) Efficiency in primary care: a critical review of frontier indicators. Rev Esp Salud Pública 74:483–495

Räty T, Luoma K (2005) Nonparametric country rankings using indicators and OECD health data. Valtion Taloudellinen Tutkimuskeskus 74:1–19

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

Richardson J, Wildman J, Robertson IK (2003) A critique of the World Health Organization's evaluation of health system performance. Health Econ 12:355–366

Roll Y, Cook WD, Golany B (1991) Controlling factor weights in data envelopment analysis. IIE Trans 23:2–9

Ross CE, Wu C (1995) The links between education and health. Am Sociol Rev 60:719–745

Sarrico CS, Dyson RG (2004) Restricting virtual weights in data envelopment analysis. Eur J Oper Res 159:17–34

Sueyoshi T, Aoki S (2001) A use of a nonparametric statistic for DEA frontier shift: the Kruskal and Wallis rank test. Omega 29(1):1–18

Thompson RG, Singleton F, Thrall R, Smith B (1996) Comparative site evaluations for locating a high energy physics lab in Texas. Interfaces 16:35–49

Thomson S, Mossialos E (2004) What are the equity, efficiency, cost containment and choice implications of private health-care funding in Western Europe?. Health Evidence Network Report. WHO Regional Office for Europe. Available at http://www.euro.who.int/Document/E83334.pdf. Accessed 10 Mar 2009

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

World Health Organization (2001) The World Health Report 2000. Health systems: improving performance. Geneva

Worthington AC (2004) Frontier efficiency measurement in health care: a review of empirical techniques and selected applications. Med Care Res Rev 61(2):135–170