Ann Oper Res (2006) 145:281–299
DOI 10.1007/s10479-006-0035-6

Efficiency and total quality management in health care organizations: A dynamic frontier approach*

Diego Prior

Published online: 6 July 2006

Abstract This paper analyses hospital performance using Data Envelopment Analysis (DEA) and the Malmquist productivity index. We follow two approaches to quantify movements in productivity: (1) the traditional approach that only considers output and input variables; and (2) a more comprehensive approach that incorporates movements in quality and restricts some achievements, if quality is reduced. On the premise that the indicator for quality (*nosocomial infections*) is equivalent to a bad output, we explore the characteristics of, and compare the results of, the different technological ways to incorporate quality (good or bad attributes, strong or weak disposability technological assumptions). After discussing the virtues and limitations of the existing possibilities, the paper presents a better formulation that allows the preservation of TQM postulates. The decomposition in the Malmquist productivity index shows an improvement in productivity and a positive technical change, especially when quality is introduced.

Keywords Data Envelopment Analysis · Malmquist productivity index · Hospitals · Quality

Introduction

This paper focuses on the evaluation of efficiency and quality in a sample of Spanish hospitals. More specifically, the focus is on hospitals working within the public health care network in Catalonia, North-East of Spain. Taking into account the concerns expressed in the literature (see, for instance Rouse, 1997), we decided to follow an evaluation oriented towards effectiveness, so that efficiency and quality jointly establish the level of effectiveness obtained.

*This paper forms part of a more extensive research work, financed by the Spanish Science and Technology Ministry (ref. SEC2003-047707).

D. Prior (✉)
Department of Business Economics, Universitat Autonoma de Barcelona,
08193 Bellaterra, Barcelona, Spain
e-mail: diego.prior@uab.es

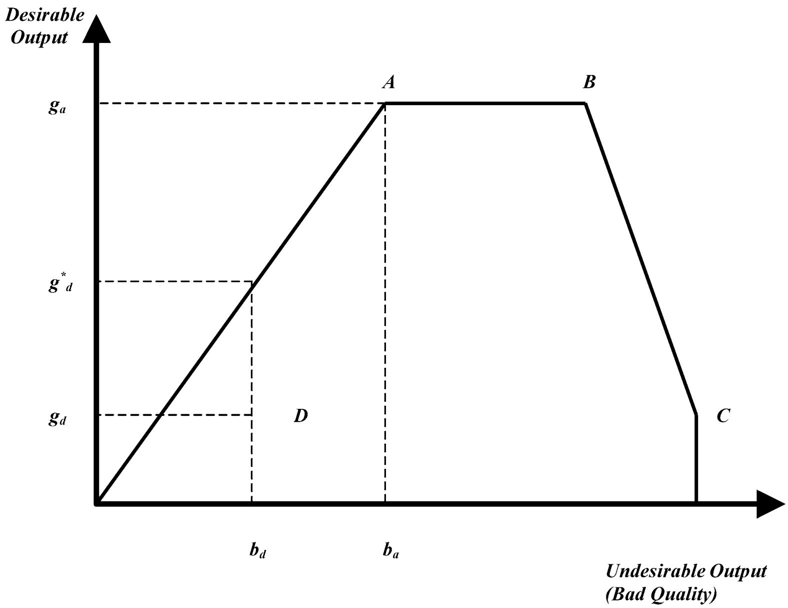

Fig. 1 Traditional approach between productivity and quality

Taken in isolation, efficiency objectives hardly serve as a final outcome. As is well known, in complex organizations productivity objectives are never the final goal to be achieved; although improvements resulting from realising productivity objectives help to meet other, more complex, objectives. The situation with quality is very different: Quality objectives can be considered as a final goal because their properties (such as zero defects or the safety of some processes) are desirable with regard to the performance objectives of health care organizations.

The relationship between efficiency and quality can follow two different paths. On the one hand, the so-called *traditional approach* assumes a negative rate of tradeoff between productivity and quality (positive rate between productivity gains and losses of quality). This follows in that, in most cases, improvements in quality require more input consumption. The situation supporting the traditional approach is presented in Fig. 1.

Figure 1 illustrates the situation of a production set with the desirable output (good output) on the y-axis and the undesirable output (bad quality) on the x-axis. A good example in health care can be, for instance, the rate of re-entrance. We assume that all four hospitals (A, B, C and D) consume the same inputs but obtain different quantities of output and quality. Hospitals A and B are the most productive. But A has better quality than B, and so it (A) dominates B. It is observable that the worst hospital, taking into account either productivity or quality, is C. Consider now the specific situation of Hospital D. D is inefficient and we can demand and expect an improvement in its productivity. Following a strictly isolated productivity analysis, it is possible to ask D for the increase in activity necessary to reach production g_a, an achievable goal but one that requires a drop in quality. The situation of Hospital D in Fig. 1 illustrates that the efficiency analysis has to be evaluated without ignoring quality. If we do this, then we can expect Hospital D to improve its efficiency without worsening quality by producing g_d^*. Obviously, from the point of view of effectiveness g_d^* is, globally, a better goal than g_a.

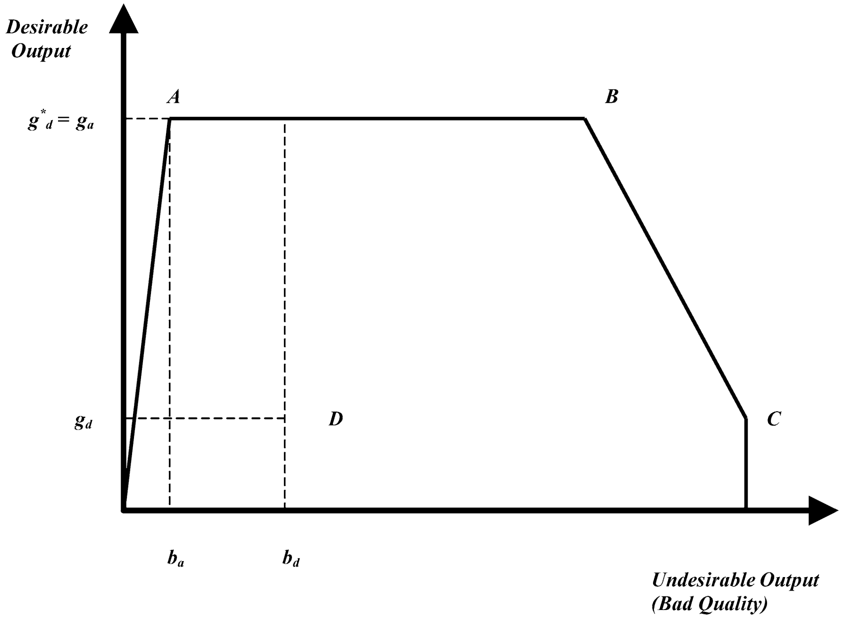

Fig. 2 Total Quality Management approach between productivity and quality

Now, let us look at the second type of relationship, known as the *Total Quality Management approach* (*TQM*). *TQM* reinforces the idea that improvements in quality lead to improvements in productivity. So, productivity and quality are related by means of a positive rate of transformation (negative rate between productivity gains and lack of quality). A situation supporting the *TQM* point of view is presented in Fig. 2. We can see that the most efficient hospital (Hospital *A*) also operates with the highest quality; so from the effectiveness point of view, it is fitting to refer to Hospital *A* as a peer for the inefficient hospitals *B, C* and *D*. In this case, the improvements in productivity and quality are aligned and there is no contradiction in the contemporaneous assumption of improvements in quality and productivity.

Knowing the different possible relationships between efficiency and quality, the empirical analysis requires a satisfactory definition of operational quality in hospitals. In this study a notion underpinned in the scientific-technical view of quality in health care is adopted: the level of health care quality is measured by a variable expressing the percentage of nosocomial infections (Section 3 outlines the reasons for this selection). We are aware that this evaluates only the technical aspect, and that we thereby set aside other important issues, such as the characteristics of personal care, the comfort of hospitalised patients, the use of commonly-shared areas, etc. Acceptance of this definition implies that it is possible to establish a rate of transformation between the physical level of production (the number of cases attended with the factors available) and the degree of technical quality in the cases treated. Consequently, we study the dominant trade-off that exists between productivity and technical quality in health care.

The methodology used to obtain the efficiency levels is based on *DEA* (Data Envelopment Analysis) evaluation. In *DEA* models a boundary is constructed with the most efficient observations without any formal production function being imposed on the data. Nor is it necessary to have access to output and input prices. All that is required are the data for inputs and outputs in physical units. There has been strong criticism about the use of *DEA* and, more generally, frontier analysis in health evaluation. For instance, Newhouse (1994) recognizes the difficulties in measuring output and adjusting for quality. This is an essential problem and more research is needed in order to derive conclusions from efficiency analysis that are 'useful for regulators'. However, it is also true that, over short periods of time, dynamic productivity indicates the improvements attained by hospitals.

In order to quantify the dynamic evolution of productivity, we have constructed the so-called Malmquist productivity index, which is very common in productivity research. The Malmquist index relates the movements between two time periods and establishes the specific position corresponding to each hospital in the sample. Our analysis adopts an integrative point of view and refers to: (a) the understanding of what is meant by evolution of productivity levels, (b) the quantification of the percentage by which each hospital should improve its levels of output and, (c) the perception of the sensitivity inherent in the best practice frontier when we consider productivity and quality variables jointly.

The remainder of this paper is organized as follows: first, in Section 1, the methodology of evaluation is presented. Section 2 analyses ways to introduce quality attributes in non-parametric frontier analysis and proposes the best way to preserve the postulates of *TQM* theories. Section 3 describes the sample of hospitals analysed and the variables used. The results are then presented in Section 4. We then conclude by outlining the considerations that we believe to be relevant and interpreting the results obtained.

1. Data envelopment analysis and Malmquist productivity index

DEA models appear in the seminal paper by Charnes, Cooper, and Rhodes (1978), adapting Farrell's (1957) proposal for treating the multiple outputs and multiple inputs technology. *DEA* models assume a convex technology; that is, they start from the belief that, if two units of production are efficient, it is possible to achieve another feasible unit by combining the two.

Linear programming models are needed to evaluate the *DEA* frontier. In the special case of output-oriented minimization and constant returns to scale technology, the so-called distance function is calculated by solving the following problem:

$$(D_i(x_i, u_i))^{-1} = \max \cdot \beta_i \tag{1}$$

$$\begin{aligned} \text{subject to:} \quad & x_{in} - \sum_{k=1}^{K} z_k \cdot x_{k,n} \geq 0 && n = 1, \ldots, N \\ & -\beta_i \cdot u_{im} + \sum_{k=1}^{K} z_k \cdot u_{k,m} \geq 0 && m = 1, \ldots, M \\ & z_k \geq 0 && k = 1, \ldots, K, \end{aligned}$$

where, $x_k = \lfloor x_{k,1}, x_{k,2}, \Lambda, x_{k,N} \rfloor \in R_+^N$ is the vector of the observed inputs corresponding to unit *k*, forming part of the sample containing *K* units; $u_k = \lfloor u_{k,1}, u_{k,2}, \Lambda, u_{k,M} \rfloor \in R_+^M$ is the vector of the observed outputs corresponding to unit *k*, forming part of the sample containing *K* units; and $z = \lfloor z_1, z_2, \Lambda, z_k \rfloor$ is the activity vector used to construct the linear segments of the frontier.

The coefficient β_i (or its inverse, the distance function $D_i(u_i, x_i)$) indicates the technical efficiency level of each of the units evaluated. If $\beta_i = 1$, the unit under evaluation is efficient in the Farrell-Debreu notion.[1] That is, no other peer has been found that yields the same output vector with a smaller consumption of inputs. Otherwise, $\beta_i > 1$ (or $D_i(u, x) < 1$), indicating the presence of technical inefficiency.

Generally, DEA models are applied to cross-sectional data. As a result, the evaluation of efficiency is obtained for a specific time-period but how efficiency varies over time is not known. To counter this weakness, there has been some dynamic approximations with the objective of quantifying the evolution of productivity over a period of time. The most widely used in frontier analysis is the Malmquist productivity index (Malmquist, 1953). Initially, Caves, Christensen, and Diewert (1982) adopted this index in order to evaluate productivity movements among different production units. More than ten years later, Färe et al. (1994) broke down the Malmquist index to recognize two sources of productivity change: (a) *efficiency change or 'catching-up effect' (EFF)*, and (b) *technical change (TCH)*.

The most widely known definition of the Malmquist productivity index, the so-called *adjacent period version*, takes the geometric mean of two Malmquist indices:

$$
\begin{aligned}
M_i^{t,t+1}\left(x_i^{t+1}, u_i^{t+1}, x_i^t, u_i^t\right) &= \left[\frac{D_i^t\left(x_i^{t+1}, u_i^{t+1}\right)}{D_i^t\left(x_i^t, u_i^t\right)} \cdot \frac{D_i^{t+1}\left(x_i^{t+1}, u_i^{t+1}\right)}{D_i^{t+1}\left(x_i^t, u_i^t\right)}\right]^{1/2} = EFF_i^{t,t+1} \cdot TCH_i^{t,t+1} \\
&= \frac{D_i^{t+1}\left(x_i^{t+1}, u_i^{t+1}\right)}{D_i^t\left(x_i^t, u_i^t\right)} \quad \text{(efficiency change)} \\
&\quad \cdot \left[\frac{D_i^t\left(x_i^{t+1}, u_i^{t+1}\right)}{D_i^{t+1}\left(x_i^{t+1}, u_i^{t+1}\right)} \cdot \frac{D_i^t\left(x_i^t, u_i^t\right)}{D_i^{t+1}\left(x_i^t, u_i^t\right)}\right]^{1/2} \quad \text{(technical change)}
\end{aligned} \tag{2}
$$

The $EFF^{t,t+1}$ index indicates whether or not a movement towards or away from the frontier has occurred between the two contiguous moments of time.

On the other hand, $TCH^{t,t+1}$ (technical change) indicates movements of the best-practice frontier between the periods t and $t+1$: It explains whether or not units belonging to the frontier have improved or worsened between these two periods. Contrary to the static DEA application, the Malmquist productivity index decomposition requires the computation of 4 linear programs for each unit under analysis: In Annex 1, we define the linear programming problems that yield the respective distance functions.

There have been subsequent proposals to improve the Malmquist productivity index. These extensions focus on three different aspects: (1) the underlying technology, (2) the notion of efficiency, and (3) the time-dimension computation of technical change. Referring to the underlying technology, Ray and Desli (1997) proposed the definition of $TCH^{t,t+1}$ using variable returns to scale as the reference technology. Additionally, Simar, and Wilson (1998), and Zofio and Lovell (1998), suggested the decomposition of $TCH^{t,t+1}$ by introducing a new component to indicate if the peers under variable returns to scale approach optimal scale under constant returns. Referring to how to compute efficiency, Grifell-Tatje, Lovell, and Pastor (1998) have proposed an alternative coefficient in order to introduce the slack variables in the

[1] Here, we follow the well-known Farrell-Debreu notion of efficiency; there is, however, another one more exigent: the Pareto-Koopmans definition of efficiency. See Färe, Grosskopf, and Lovell (1994) for a detailed study of their characteristics.

computation of the Malmquist indices- a proposal that has received criticism from Førsund (1998). Also of note is Lovell and Zofío's (1997) proposal of the use of 'graphyperbolic' efficiency coefficients in order to simultaneously increase outputs and reduce inputs. With respect to the time-dimension computation of technical change, Berg, Førsund, and Jansen (1992) introduced a *base period* (the base-period index is a two-period notion, but it uses an additional period to measure the technical change) Malmquist index alternative to the *adjacent period* (as indicated previously, the adjacent version is a two-period notion and measures the shift in the technology frontier as the shift in the frontier at time t and $t + 1$). Comparison of the characteristics of the adjacent and the base-period versions can be found in Althin (2001).

In our application, we follow the original Färe et al. (1994) version for a number of reasons. First, we do not compute technical change with variable returns to scale because of the reduced size of our sample (we only have 29 hospitals) and the unfeasible problems this technology can raise (problems already pointed out in Burguess and Wilson (1995). However, in order to observe scale inefficiency, following Førsund and Hjalmarsson (1979), we introduce three artificial hospitals with average values corresponding to three different sizes. Second, we apply distance functions following the Debreu-Farrell notion of efficiency because our main objective is to observe the impact on the Malmquist index of the inclusion of the quality variables, and the distance functions are the best way to compute this effect. Third, we decided to apply the adjacent-period version because of the reduced time period we analyze (3 years) and the lack of information about a third year necessary to apply the base-period version.

2. Introduction of quality in the Malmquist productivity index

As previously mentioned, in measuring productivity, it is necessary to consider quality in order to guarantee that the improvement in productivity would not be achieved at the expense of a reduction in service quality.

The introduction of attributes of quality in productivity indices was first proposed in Fixler and Zieschang (1992). Afterwards, Färe, Grosskopf and Roos (1995) redefined the Malmquist productivity index in order to incorporate attributes of quality into the technology. Färe, Grosskopf and Roos defined the technology comprising a set of feasible input and output vectors; among the outputs they distinguished between those that are marketable (u) and those that are desirable attributes (a). Given this specification of the technology, they redefined the Malmquist index in the following way (see the linear programming problems that give the respective distance functions, in Annex 2):

$$
\begin{aligned}
&\tilde{M}_i^{t,t+1}\left(x_i^{t+1}, a_i^{t+1}, u_i^{t+1}, x_i^t, a_i^t, u_i^t\right) \\
&\qquad = \left[\frac{D_i^t\left(x_i^{t+1}, a_i^{t+1}, u_i^{t+1}\right)}{D_i^t\left(x_i^t, a_i^t, u_i^t\right)} \cdot \frac{D_i^{t+1}\left(x_i^{t+1}, a_i^{t+1}, u_i^{t+1}\right)}{D_i^{t+1}\left(x_i^t, a_i^t, u_i^t\right)}\right]^{1/2} \\
&\qquad = \frac{D_i^{t+1}\left(x_i^{t+1}, a_i^{t+1}, u_i^{t+1}\right)}{D_i^t\left(x_i^t, a_i^t, u_i^t\right)} \cdot \quad \text{(efficiency change)} \\
&\qquad \cdot \left[\frac{D_i^t\left(x_i^{t+1}, a_i^{t+1}, u_i^{t+1}\right)}{D_i^{t+1}\left(x_i^{t+1}, a_i^{t+1}, u_i^{t+1}\right)} \cdot \frac{D_i^t\left(x_i^t, a_i^t, u_i^t\right)}{D_i^{t+1}\left(x_i^t, a_i^t, u_i^t\right)}\right]^{1/2} \text{(technical change)}, \quad (3)
\end{aligned}
$$

where $a_i^t = \lfloor a_{i,1}^t,\ a_{i,2}^t,\ \Lambda,\ a_{i,A}^t \rfloor \in R_+^A$ is the vector of the desired attributes of quality corresponding to the unit under evaluation (unit i).[2]

When quality variables refer to bad attributes (i.e., when quality variables are expressed in terms of the rate of infections or the number of unresolved treatments), the decomposition presented in (3) still holds, but the programs presented in Annex 2 require adaptation. In this way, Färe et al. (1989) propose to modify the technology assuming that the desirable outputs (u_i) are strongly disposable and that the undesirable or bad attributes (b_i) are weakly disposable. [3] This implies modifying the programs presented in Annex 2 so as to expand the desirable outputs while, at the same time, controlling the level of the bad attributes. So, the adaptation of program (A5) to this new specification of the technology is as follows:

$$\left(D_i^t\left(x_i^t, a_i^t, u_i^t\right)\right)^{-1} = \max \cdot \beta_i \tag{A5$'$}$$

$$\begin{aligned} \text{subject to :} \quad & x_{i,n}^t - \sum_{k=1}^{K} z_k \cdot x_{k,n}^t \geq 0 && n = 1, \ldots, N \\ & -\beta_i \cdot u_{i,m}^t + \sum_{k=1}^{K} z_k \cdot u_{k,m}^t \geq 0 && m = 1, \ldots, M \\ & -b_{i,b}^t + \sum_{k=1}^{K} z_k \cdot b_{k,b}^t = 0 && b = 1, \ldots, B \\ & z_k \geq 0 && k = 1, \ldots, K \end{aligned}$$

where the vector $b_i^t = \lfloor b_{i,1}^t,\ b_{i,2}^t, \ldots, b_{i,B}^t \rfloor \in R_+^B$ contains the attributes of bad quality of the unit under evaluation (unit i).[4] We illustrate this situation in Fig. 1. The maximum output achievable when the bad attributes are not taken into account is g_a (precisely the solution of program (A1). Taking care not to modify the attributes of bad quality, the maximum output is g_d^*. Consequently, part of the potential increase in the output level ($g_a - g_d^*$) is impossible to achieve without accepting any deterioration in the standards of quality. In other words, the control of bad outputs (provided that they are only weakly disposable) requires the consumption of resources that could otherwise be applied to the production of good outputs.

The consideration of the axiom of the weak disposability of bad outputs is the most general assumption in efficiency analysis with bad outputs (see, for instance, Färe, Grosskopf, and Hernandez-Sancho, 2000). However, in cases where the postulates from Total Quality Management (Crosby, 1979; Deming, 1982) prevail, the weak-disposability axiom could lead to a sub-optimal solution. This is illustrated in Fig. 3. In Fig. 3 we observe that unit *A,* which is dominant, produces the greatest quantity of good output and the lowest level of bad output, but, when evaluating unit *D*, the application of program (A5$'$) gives level g_d^* as the goal on the frontier, which, obviously, is worse than g_a. In order to face situations like these, two main

[2] Färe, Grosskopf and Roos (1995) also introduced a quality index change in the Malmquist index decomposition. The application of this quality index change to our sample of Spanish hospitals has been presented in a previous work (see Prior and Sola, 2001).

[3] This modification in the methodology permits an asymmetric treatment of inputs, desirable outputs and undesirable outputs. The difference between the weakly and the strongly disposable technologies enables the quantification of the desirable output loss due to the lack of strong disposability of undesirable outputs. See Färe et al. (1989).

[4] The rest of the programs included in Annex 2 have to be modified in the same way.

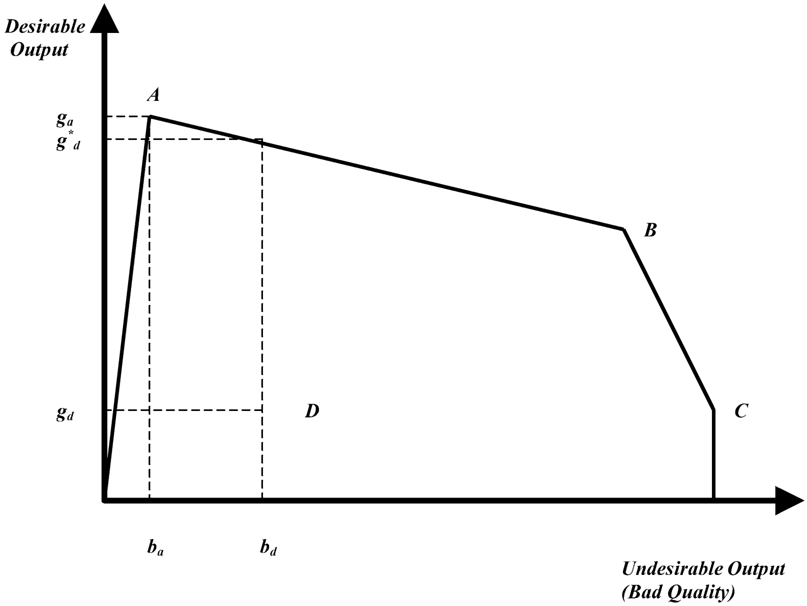

Fig. 3 Dominance of the postulates of the Total Quality Management approach

solutions have been put forward: the use of hyperbolic efficiency coefficients (Färe et al., 1989) or the more recent use of directional distance functions (Chung, Färe, and Grosskopf, 1997). Here, we follow an alternative path because, for the specific situation illustrated in Fig. 3, another perspective is preferred: the consideration of the undesirable attributes of bad quality as inputs. In fact, this concept has already been used in efficiency measurement (See Tyteca, 1997; Reinhard, Lovell, and Thijssen, 1999) but, to our knowledge, has never been used in the determination of Malmquist productivity indices. This option implies modifying program (A5) (and the rest of the programs presented in Annex 2) in the following way:

$$\left(D_i^t\left(x_i^t, a_i^t, u_i^t\right)\right)^{-1} = \max \cdot \beta_i \tag{A5''}$$

$$\begin{aligned} \text{subject to :} \quad & x_{i,n}^t - \sum_{k=1}^{K} z_k \cdot x_{k,n}^t \geq 0 & n = 1, \ldots, N \\ & -\beta_i \cdot u_{i,m}^t + \sum_{k=1}^{K} z_k \cdot u_{k,m}^t \geq 0 & m = 1, \ldots, M \\ & b_{i,b}^t - \sum_{k=1}^{K} z_k \cdot b_{k,b}^t \geq 0 & b = 1, \ldots, B \\ & z_k \geq 0 & k = 1, \ldots, K. \end{aligned}$$

Program (A5″) gives the maximal increment in desirable outputs while controlling the bad outputs. In the case presented in Fig. 3, this new program signals g_a as the target output. This is the most preferred solution according to our preferences structure and, at the same time, is perfectly aligned with the Total Quality Management postulates.[5]

3. Variables used for measuring technical quality in hospitals, outputs and inputs

Although the evaluation of quality in services is fundamentally a multidimensional task (Donabedian, 1980), here we concentrate on the technical dimension of quality. The technical quality of hospital treatment can be analyzed by means of different variables: (a) the intra-hospital mortality, after controlling for demographics, within a fixed period following admission (Gowrisankaran and Town, 1999); (b) the changes in functional and cognitive status, and living arrangements (Sloan et al., 2001); (c) indicators such as surgical complications; and (d) the number of re-admissions. There are some difficulties with these variables and, in the majority of the cases, they are important difficulties. On the one hand, information that is, at once, both homogenous and periodical is not easy to access. On the other hand, for some of the indicators, such as the level of re-admission, it is not entirely clear that they act as good quality indicators, since their connection to irregular processes is not always evident. Besides the four variables listed above, we might also include other indicators that measure user satisfaction with hospital services. However, the lack of periodical information does not facilitate their inclusion within our productivity indicators.

We do not use the aforementioned indicators here. Instead, the variable we use is the level of *nosocomial infections*,[6] which is traditionally accepted as a representative of the level of technical quality. Nosocomial infections are one of the principal causes of morbidity and mortality amongst hospitalised patients (Worning, 1994) and, clearly, directly affect the ultimate goal of all hospitals: to improve the health of their patients. In fact, these infections affect the greatest part of the aspects defined as components of health-care service quality. They have negative consequences on the effectiveness of a given program, since the infected patient develops complications that might result in death; and have similarly negative effects on efficiency, since the impact of the program will be reduced and costs will rise, given the increase in time spent in hospital and the use of resources. Furthermore, nosocomial infections have negative consequences on consumer satisfaction, as patients' expectations of recuperation fail to materialise. In fact, nosocomial infections are related to other indicators of bad quality, such as the proportion of residents with facility-acquired pressure sores (Grabowski, 2001), because they are preventable and treatable.

Even though there is, arguably, a minimum level of infection beyond which the infections cannot be eliminated, these infections can only be controlled with greater use of resources and, as a result, their presence implies a loss in apparent hospital productivity. In order to monitor and control the infections, a specifically trained medical and nursing staff, an

[5] As pointed out by one referee, following (Seiford and Zhu, 2002), another alternative is to translate the output. In our specific case study, our alternative is more flexible: it tries to improve desirable outputs while maintaining or, if possible, reducing undesirable outputs. The method of Seiford and Zhu requires radial movements in the desirable and also in the 'translated' undesirable outputs.

[6] Nosocomial infections are produced by hospital-acquired micro organisms that affect patients who have been admitted for a different illness. The infection is not present at the time of admission, nor is it in an incubational phase. Nosocomial infection statistics are extensively controlled in some other countries. For instance, for USA hospitals there is an organisation: National Nosocomial Infections Surveillance System (NNIS), which maintains a national nosocomial infections database. More information can be found on the web page: http://www.cdc.gov/ncidod/hip/Surveill/NNIS.HTM.

adoption of courses of action such as sterilization, disinfections etc. are needed (EPINE, 1995).

Summing up, on the outputs side, we have defined the following desirable and undesirable outputs:

u_1 (ACUTE): in-patient days spent in medical care, surgery, obstetrics, gynaecology and paediatrics.
u_2 (LONGSTAY): in-patient days spent in long-stay care and psychiatry.
u_3 (INTENSIVE): in-patient days spent in intensive care.

In accordance with the methodology used to collect the statistics, we understand 'in-patient days' to be the combination of night stay and the time corresponding to the serving of a main meal (lunch or supper).

u_4 (VISITS): medical care on an out-patient basis, for the diagnosis, treatment and monitoring of illness.
b_1 (INFECTIONS): The prevalence of nosocomial infections (percentage of nosocomial infections per patients treated): number of clinically active infections/number of patients treated.

As noted by Murray (1992), if we are interested in the final impact on hospital services, there are reasons to prefer outcome variables (i.e., the number of patients treated weighted by its Group Diagnostic Related classification) to throughput variables (number of hospital days). However, following Chilingerian and Sherman (1990), our application uses in-patient days as an output variable and concentrates on the efficiency of hospital management. Obviously, this analysis can be extended in a second stage by focusing on the efficiency of health staff in producing real health services (the value of the health added to the patients). Here our analysis concentrates exclusively on the first stage because our database does not contain the required information to complete the second stage.[7] The data we have is similar to the data used in Ozcan and Luke (1993). As in Ozcan (1995); and in O'Neill (1998), it is possible to define specific hospital variables because we have data about the case-mix of the hospitals.

Table 1 shows the descriptive statistics for the outputs. It can be seen that, on average, outputs related to in-patient days spent in hospitals have decreased. On the other hand, the number of visits has increased, indicating the inertia of replacing hospital stays by external visits since 1990. Our proxy to quality (infections) has, on average, decreased, which means that hospitals have had some success in controlling this dimension of quality.

Moving on to the inputs side, we have defined following variables:

(x_1) (PHYSICIANS): health care staff. This is made up of full-time medical staff.
(x_2) (OTHERSTAFF): other nursing personnel and non-health care staff, also full-time.
(x_3) (BEDS): number of beds assigned for the continuous care of admitted patients.
(x_4) (MATERIALS): Money spent on current purchases (Spanish pesetas in constant prices).

The descriptive statistics for inputs are presented in Table 2, which shows a trend towards reducing the number of physicians, the opposite of what to has happened to the other staff.

[7] In order to verify the impact on the results when comparing throughput and outcome variables in Spanish hospitals, there is very detailed research based on another sample of Spanish hospitals (see Calzado et al., 1998); this points out the differences and the similarities in the results of the two ways of specifying the outputs.

Table 1 Descriptive statistics for outputs

		1990	1993
(u_1) ACUTE	Arithmetic mean	88,835.17	87,891.37
	Standard deviation	59,223.62	59,051.92
	Maximum	274,400.00	271,937.00
	Minimum	25,038.00	26,360.00
(u_2) LONGSTAY	Arithmetic mean	6,083.79	6,181.31
	Standard deviation	13,750.96	14,754.60
	Maximum	72,812.00	78,197.00
	Minimum	0.00	0.00
(u_3) INTENSIVE	Average	3,750.20	3,050.65
	Standard deviation	5,689.52	5,279.35
	Maximum	22,512.00	22,284.00
	Minimum	0.00	0.00
(u_4) VISITS	Arithmetic mean	91,093.48	99,941.31
	Standard deviation	94,496.75	86,214.24
	Maximum	520,591.00	411,968.00
	Minimum	12,353.00	0.00
(b_1) INFECTIONS	Arithmetic mean	9.67	7.85
	Standard deviation	5.52	3.88
	Maximum	33.33	20.32
	Minimum	2.81	2.20

Table 2 Descriptive statistics for inputs

		1990	1993
(x_1) PHYSICIANS	Arithmetic mean	504.36	502.08
	Standard deviation	463.90	416.38
	Maximum	2,051.92	1,914.99
	Minimum	123.94	111.74
(x_2) OTHER STAFF	Arithmetic mean	133.00	244.53
	Standard deviation	118.13	492.88
	Maximum	529.33	2,743.63
	Minimum	35.74	28.16
(x_3) BEDS	Arithmetic mean	346.31	330.62
	Standard deviation	222.42	207.66
	Maximum	967.00	931.00
	Minimum	101.00	99.00
(x_4) MATERIALS	Arithmetic mean	728,350.20	1,383,560.72
	Standard deviation	865,050.93	1,493,460.23
	Maximum	3,766,932.00	6,743,263.00
	Minimum	87,108.00	231,033.00

With regard to beds, a slight decrease is seen, while the biggest increase is in current purchases, which have almost doubled in four years.

It is worth noting the size of the hospitals we are evaluating: It varies from 90 to 900 beds, although the majority have between 100 and 400 beds. The hospitals form part of the network of public-use hospitals, with twelve being publicly owned while the others are private. None of the hospitals in the sample is psychiatric. Nine of the hospitals in the sample are small (with less than 200 beds), 11 are medium-sized hospitals (with more than 200 but less than

400 beds); and nine are big (with more than 400 beds). The sample of hospitals used for this application was constrained by the data about the bad output (b_1), which was only available for 29 general hospitals (from a population of 54 hospitals operating in the Catalonian public health service network).

Physical data was obtained from the Catalonian *Statistics for Health Establishment Admissions* organization; data referring to nosocomial infections came from studies undertaken by the *EPINE* working group, co-ordinated by the *Spanish Society for Hygiene and Preventive Medicine in Hospitals*. Taking infections as undesirable outputs- in any case they are maximized: programs (A1) to (A4) do not take into account this output; program (A5′) does not maximize weak-disposable outputs and program (A5″) seeks to minimize bad outputs.

4. Results obtained from the proposed evaluation

We present in this section the results of the analysis. Let us start by examining the specific distance functions for each year under static frontiers. Table 3 shows the results corresponding to the following three models: (1) considering only inputs and desirable outputs, (2) introducing the level of infections (undesirable output) as a weak disposable output and (3) introducing the level of infections as a strong disposable input.

When comparing the values, it can be observed that the distance functions are sensitive to the specification of the technological reference. Thus, the lower the restrictions considered, the lower the level of the distance functions, which is the case for model 1 i.e., quality not considered. Model 2 is the most restrictive and the situation here is just the opposite: the higher the restrictions, the greater the distance functions. On average, an increase of 7.53% $(1/0.9299 - 1)$ in the 1990 outputs would have resulted in all the hospitals being on the frontier. Taking quality into account, the average potential increase in the output reduces to 4.52% $(1/0.9567 - 1)$ or 4.64% $(1/0.9556 - 1)$, depending on which evaluation we study.

Table 3 Static distance functions

Model		1990	1993
		$D_i^{90}\left(x_i^{90}, u_i^{90}\right)$	$D_i^{93}\left(x_i^{93}, u_i^{93}\right)$
1. Without the introduction of the quality (programs A1 and A2)	Arithmetic mean	0.9299	0.9074
	Standard deviation	0.0993	0.1097
	Maximum	1.0000	1.0000
	Minimum	0.6189	0.5544
		$D_i^{90}\left(x_i^{90}, b_i^{90}, u_i^{90}\right)$	$D_i^{93}\left(x_i^{93}, b_i^{93}, u_i^{93}\right)$
2. Considering quality as a weak disposable output (programs A5′ and A6′)	Arithmetic mean	0.9567	0.9239
	Standard deviation	0.0877	0.1012
	Maximum	1.0000	1.0000
	Minimum	0.6225	0.5915
		$D_i^{90}\left(x_i^{90}, b_i^{90}, u_i^{90}\right)$	$D_i^{93}\left(x_i^{93}, b_i^{93}, u_i^{93}\right)$
3. Considering quality as a strong disposable input (programs A5″ and A6″)	Arithmetic mean	0.9556	0.9201
	Standard deviation	0.0875	0.1062
	Maximum	1.0000	1.0000
	Minimum	0.6225	0.5544

The 1993 results show higher inefficiency, verifying that all the distance functions are lower, indicating more inefficiency than in 1990. The general pattern of change from efficient to inefficient of individual *DMU's* can also be verified by looking at the descriptive statistics presented in Tables 1 and 2. On average, two outputs (acute and intensive) have declined slightly while two inputs (physicians and materials) have increased substantially.

All of these differences are significant because, according to the Wilcoxon Signed Ranks Test, the differences in the distributions of the distance functions are statistically significant both within the evaluations and between the years (all at 1% significance, except for the distance functions of the second and the third evaluation for 1990, which are significant at 5%). The situation can be expressed in another way: In 1990 nine hospitals (31% of the total) registered changes in the distance function when applying programs (A5′) and (A5″). The rest remained the same. So, a third of the sample is in the situation illustrated in Fig. 3 while the rest is well represented by Fig. 2. The 1993 situation was more balanced: 12 hospitals (41% of the total) registered changes and 17 maintained the same distance function. So the Total Quality Management postulates of the domination in efficiency and quality holds for a sub-sample comprising between 30% and 40% of the total sample.

To summarize, the information presented in Table 3 shows how inefficiency (the distance separating the hospitals on the frontier and the hospitals below the frontier) has grown and, according to the Wilcoxon Test, we can reject the null hypothesis of the equality of the distributions of the distance functions among the different models defined. Since the only difference among the three models is based on the consideration of the variable for the quality, the obvious conclusion is that quality matters, even when we evaluate efficiency, and, equally important, the way we introduce quality (describing the different technological possibilities) has a significant effect on the distribution of the distance functions.

Now that we know what the situation is with respect to the specific frontier, let us turn to the Malmquist productivity index. As mentioned in the previous section, we have information for 29 hospitals, covering the years 1990 and 1993. The total number of hospitals is reduced and the years under analysis are too close to expect important movements in the Malmquist index. This is why, in order to add stability, we decided to quantify a *sequential frontier* by constructing the reference production set for 1993 with the observations not only of 1993 but also of 1990 (Tulkens and Vanden Eeckaut, 1995). This approach implies that knowledge does not get lost (technical regress is not possible with this specification); and past progress is accounted for in the determination of the contemporaneous frontier corresponding to year 1993. Table 4 presents the results for the Malmquist productivity index.

The general picture is that the Malmquist index grew due to the positive Technical Change, in spite of the retrogressive Efficiency Change. On average, all movements were in the same direction, independently of the technological approach assumed for the frontier evaluation. However, the Malmquist index reaches a higher value when we account for quality, the application notwithstanding. According to the Wilcoxon Signed Ranks Test, when model 1 is compared with models 2 and 3, the null hypothesis can be rejected at the 5% significance level for the distribution of the Malmquist index and at the 1% significance level for the Technical Change index, but not for the Efficiency index. These results confirm that applying the Malmquist productivity indices without considering the movements in quality, as in Model 1, is a poor way to model the real changes in effectiveness.

Given that the programs defined assume a constant return to scale technology, it is worth observing if there are differences in the results due to the size of the hospitals. To do this, we constructed the average hospital for three sub-samples for each year and included these artificial hospitals as additional observations. As Førsund and Hjalmarsson (1979) point out,

Table 4 Malmquist productivity index

Model		$M^{90,93}$	$EFF^{90,93}$	$TCH^{90,93}$
1. Without the introduction of the quality (programs A1, A2, A3 and A4)	Arithmetic mean	1.0046	0.9750	1.0300
	Standard deviation	0.0582	0.0436	0.0242
	Maximum	1.1199	1.0436	1.0855
	Minimum	0.8679	0.8497	1.0027
2. Considering quality as a weak disposable output (programs A5′, A6′, A7′ and A8′)	Arithmetic mean	1.0577	0.9652	1.0939
	Standard deviation	0.1489	0.0498	0.1247
	Maximum	1.4341	1.0291	1.4341
	Minimum	0.8963	0.8216	1.0105
3. Considering quality as a strong disposable input (programs A5″, A6″, A7″ and A8″)	Arithmetic mean	1.0547	0.9616	1.0945
	Standard deviation	0.1518	0.0546	0.1243
	Maximum	1.4341	1.0291	1.4341
	Minimum	0.8678	0.8215	1.0103

Table 5 Structural analysis of the Malmquist productivity index

Model	Size of hospitals	$M^{90,93}$	$EFF^{90,93}$	$TCH^{90,93}$
1. Without the introduction of the quality (programs A1, A2, A3 and A4)	Less than 200 beds	0.9516	0.9347	1.0180
	Between 200 and 400 beds	0.9836	0.9607	1.0238
	More than 400 beds	1.0377	1.0041	1.0334
2. Considering quality as a weak disposable output (programs A5′, A6′, A7′ and A8′)	Less than 200 beds	0.9514	0.9342	1.0184
	Between 200 and 400 beds	0.9861	0.9608	1.0262
	More than 400 beds	1.0281	0.9988	1.0293
3. Considering quality as a strong disposable input (programs A5″, A6″, A7″, and A8″)	Less than 200 beds	0.9514	0.9342	1.0184
	Between 200 and 400 beds	0.9861	0.9607	1.0264
	More than 400 beds	1.0281	0.9988	1.0293

this constitutes a more satisfactory measure of structural efficiency than the conventional approach of the weighted average (by output) of the efficiency scores of the individual units. The results are presented in Table 5.

It appears that big hospitals are the best performers, whatever application we look at. Once again we see how positive technical change improves the frontier, but this exacerbates inefficiency, separating the sectors from their best-practice frontier. In general, the exclusion of quality increases the indices and even changes the grading of good performer to bad performers; this is the case of the Efficiency Change index for big hospitals.

5. Conclusions

In this study, we have examined movements in productivity and quality in a sample of Spanish hospitals. From the methodological point of view, after a review of the literature, we studied the implications of the two existing proposals on how to introduce quality attributes in DEA models, namely: (a) as a desirable strongly disposable attribute and (b) as an undesirable weakly disposable output. However, to incorporate the postulates of Total Quality Management into DEA models, we conclude that it is possible to preserve the best of productivity and quality by taking another perspective (suggested and applied in the past by Tyteca (1997) and Reinhard, Lovell, and Thijssen (1999)). It involves looking at the undesirable attributes of quality (in our case the nosocomial infections) as a strongly disposable input. This technological option, which, as far as we know, has never been used in the Malmquist index literature, is of significant importance when the prevailing situation is similar to that presented in Fig. 3, illustrating a hospital that dominates both in the maximization of desired outputs and in the minimization of undesirable infections (just postulating that total quality helps productivity as the *TQM* theory defends).

In order to observe what is the preponderant situation in our specific case study, the applied part of the article presented the results of three applications: (1) excluding the attributes of quality in the calculation of efficiency, (2) introducing the level of infections (undesirable output) as a weak disposable output and (3) considering the level of infections as a strong disposable input. The results pointed out the importance of the defining the technology right, because the distributions of the distance functions are statistically different for all the applications. We also found that *TQM* holds for between 30% and 40% of the hospitals in the sample (say, the maximal improvement in desirable outputs induces, simultaneously, contraction in the undesirable attributes of bad quality). Summing up, it is not prudent to neglect the consideration of the attributes of quality when we evaluate the efficiency of the hospitals as we have a problem of misspecification and, equally important, the way we introduce the quality (describing the different technological possibilities) has a significant effect on the distribution of the distance functions.

When we came to the Malmquist productivity index, we once again applied the three previously mentioned specifications. As in the static application, significant differences appear indicating that the exclusion of quality affects the distribution of the indices (evaluating productivity indices without considering movements in quality—Application 1—is a poor model of real movements in effectiveness). In our specific case, the error reduces the productivity indices, but, depending on the evolution of the quality variable, bias of any sort can be expected.

Referring to structural scale efficiency, we found that big hospitals are the best performers, whatever application we look at. We also saw that positive technical change improves the frontier, but this exacerbates inefficiency, separating the sectors from their best-practice frontier.

Our final comment involves two considerations. The first is that, in health care, quality is important from the point of view of effectiveness. As such, the analysis of productivity and efficiency must take into consideration quality attributes; otherwise, there is a problem of misspecification with unpredictable effects. From the applied point of view, it is worth noting that, when *TQM* postulates are assumed as a target for hospitals working in the public sector environment, the methodological tools used have to be carefully defined in order not to favour sub-optimal behaviour. However, given the reduced number of hospitals analysed and the limitations of the variables used, we recommend that to deal with such issues future research should include all Spanish hospitals, or hospitals of similar type in the European Union countries.

Annex 1. Linear-programming problems (traditional formulation of distance functions)

$$\left(D_i^t\left(x_i^t, u_i^t\right)\right)^{-1} = \max .\beta_i \tag{A1}$$

$$\begin{aligned}
\text{subject to :}\quad & x_{i,n}^t - \sum_{k=1}^{K} z_k \cdot x_{k,n}^t \geq 0 && n = 1, \ldots, N\\
& -\beta_i \cdot u_{i,m}^t + \sum_{k=1}^{K} z_k \cdot u_{k,m}^t \geq 0 && m = 1, \ldots, M\\
& z_k \geq 0 && k = 1, \ldots, K
\end{aligned}$$

$$\left(D_i^{t+1}\left(x_i^{t+1}, u_i^{t+1}\right)\right)^{-1} = \max .\beta_i \tag{A2}$$

$$\begin{aligned}
\text{subject to :}\quad & x_{i,n}^{t+1} - \sum_{k=1}^{K} z_k \cdot x_{k,n}^{t+1} \geq 0 && n = 1, \ldots, N\\
& -\beta_i \cdot u_{i,m}^{t+1} + \sum_{k=1}^{K} z_k \cdot u_{k,m}^{t+1} \geq 0 && m = 1, \ldots, M\\
& z_k \geq 0 && k = 1, \ldots, K
\end{aligned}$$

$$\left(D_i^t\left(x_i^{t+1}, u_i^{t+1}\right)\right)^{-1} = \max .\beta_i \tag{A3}$$

$$\begin{aligned}
\text{subject to :}\quad & x_{i,n}^{t+1} - \sum_{k=1}^{K} z_k \cdot x_{k,n}^t \geq 0 && n = 1, \ldots, N\\
& -\beta_i \cdot u_{i,m}^{t+1} + \sum_{k=1}^{K} z_k \cdot u_{k,m}^t \geq 0 && m = 1, \ldots, M\\
& z_k \geq 0 && k = 1, \ldots, K
\end{aligned}$$

$$\left(D_i^{t+1}\left(x_i^t, u_i^t\right)\right)^{-1} = \max .\beta_i \tag{A4}$$

$$\begin{aligned}
\text{subject to :}\quad & x_{i,n}^t - \sum_{k=1}^{K} z_k \cdot x_{k,n}^{t+1} \geq 0 && n = 1, \ldots, N\\
& -\beta_i \cdot u_{i,m}^t + \sum_{k=1}^{K} z_k \cdot u_{k,m}^{t+1} \geq 0 && m = 1, \ldots, M\\
& z_k \geq 0 && k = 1, \ldots, K
\end{aligned}$$

Annex 2. Linear-programming problems (quality output-based distance functions)

$$\left(D_i^t\left(x_i^t, a_i^t, u_i^t\right)\right)^{-1} = \max .\beta_i \tag{A5}$$

$$\begin{aligned}
\text{subject to :}\quad & x_{i,n}^t - \sum_{k=1}^{K} z_k \cdot x_{k,n}^t \geq 0 && n = 1, \ldots, N \\
& -\beta_i \cdot u_{i,m}^t + \sum_{k=1}^{K} z_k \cdot u_{k,m}^t \geq 0 && m = 1, \ldots, M \\
& -a_{i,a}^t + \sum_{k=1}^{K} z_k \cdot a_{k,a}^t \geq 0 && a = 1, \ldots, A \\
& z_k \geq 0 && k = 1, \ldots, K
\end{aligned}$$

$$\left(D_i^{t+1}\left(x_i^{t+1}, a_i^{t+1}, u_i^{t+1}\right)\right)^{-1} = \max .\beta_i \tag{A6}$$

$$\begin{aligned}
\text{subject to :}\quad & x_{i,n}^{t+1} - \sum_{k=1}^{K} z_k \cdot x_{k,n}^{t+1} \geq 0 && n = 1, \ldots, N \\
& -\beta_i \cdot u_{i,m}^{t+1} + \sum_{k=1}^{K} z_k \cdot u_{k,m}^{t+1} \geq 0 && m = 1, \ldots, M \\
& -a_{i,a}^{t+1} + \sum_{k=1}^{K} z_k \cdot a_{k,a}^{t+1} \geq 0 && a = 1, \ldots, A \\
& z_k \geq 0 && k = 1, \ldots, K
\end{aligned}$$

$$\left(D_i^t\left(x_i^{t+1}, a_{i,a}^{t+1}, u_i^{t+1}\right)\right)^{-1} = \max .\beta_i \tag{A7}$$

$$\begin{aligned}
\text{subject to :}\quad & x_{i,n}^{t+1} - \sum_{k=1}^{K} z_k \cdot x_{k,n}^t \geq 0 && n = 1, \ldots, N \\
& -\beta_i \cdot u_{i,m}^{t+1} + \sum_{k=1}^{K} z_k \cdot u_{k,m}^t \geq 0 && m = 1, \ldots, M \\
& -a_{i,a}^{t+1} + \sum_{k=1}^{K} z_k \cdot a_{k,a}^t \geq 0 && a = 1, \ldots, A \\
& z_k \geq 0 && k = 1, \ldots, K
\end{aligned}$$

$$\left(D_i^{t+1}\left(x_i^t, a_i^t, u_i^t\right)\right)^{-1} = \max .\beta_i \tag{A8}$$

$$\begin{aligned}
\text{subject to :}\quad & x_{i,n}^t - \sum_{k=1}^{K} z_k \cdot x_{k,n}^{t+1} \geq 0 && n = 1, \ldots, N \\
& -\beta_i \cdot u_{i,m}^t + \sum_{k=1}^{K} z_k \cdot u_{k,m}^{t+1} \geq 0 && m = 1, \ldots, M \\
& -a_{i,a}^t + \sum_{k=1}^{K} z_k \cdot a_{k,a}^{t+1} \geq 0 && a = 1, \ldots, A \\
& z_k \geq 0 && k = 1, \ldots, K
\end{aligned}$$

Acknowledgments Early versions of this paper were presented at the 4th European Conference on Health Economics, Paris (France), July 2002, XXVI Congress of the European Accounting Association, Seville (Spain), April 2003 and in 8th European Workshop on Efficiency and Productivity Analysis, Oviedo (Spain), September 2003. Participants comments are gratefully acknowledged. This study also greatly benefited from comments made by three anonymous referees. The usual disclaimer applies.

References

Althin, R. (2001). "Measurement of Productivity Changes: Two Malmquist Index Approaches." *Journal of Productivity Analysis*, 16, 107–128.

Berg, S.A., F.R. Førsund, and E.S. Jansen. (1992). "Malmquist Indices of Productivity Growth during the Deregulation of Norwegian Banking, 1980–1989." *Scandinavian Journal of Economics* (Supplement), 211–228.

Burguess, J.F., Jr. and P.W. Wilson. (1995). "Decomposing Hospital Productivity Changes, 1985–1988: A Nonparametric Malmquist Approach." *The Journal of Productivity Analysis*, 6, 343–363.

Calzado, Y., T. García, J. Laffarga, and M. Larran. (1998). "Relación Entre Eficiencia y Efectividad en Los Hospitales del Servicio Andaluz de Salud. *Efficiency and Effectiveness in Andalusian Public Hospitals*." *Revista de Contabilidad*, 1(2), 49–83 (English summary).

Caves, D.W. et al. (1982). "The Economic Theory of Index Numbers and the Measurement of Input, Output, and Productivity." *Econometrica*, 50(6), 1393–1414.

Charnes, R.D. et al. (1978). "Measuring the Efficiency of Decision-Making Units." *European Journal of Operational Research*, 2, 429–444.

Chilingerian, J.A. and H.D. Sherman. (1990). "Managing Physician Efficiency and Effectiveness in Providing Hospital Services." *Health Services Management Research*, 3(1), 3–15.

Crosby, P.B. (1979). *Quality is free*. McGraw Hill.

Chung, Y.H., R. Färe, and S. Grosskopf. (1997). "Productivity and Undesirable Outputs: A Directional Distance Function Approach." *Journal of Environmental Management*, 51, 229–240.

Deming, W.E. (1982). *Quality, Productivity and Competitive Position*. MIT Center for Advanced Engineering Study.

Donabedian, A. (1980). *The Definition of Quality and Approaches to its Assessment*. The Regents of the University of Michigan, Ann Arbor, MI

EPINE. (1995). *Prevalencia de las Infecciones Nosocomiales en los Hospitales Españoles. Epine 1990–1994*. (Nosomial infections in Spanish hospitales, 1990–1994). Barcelona: Sociedad Española de Higiene y Medicina Preventiva Hospitalaria.

Fare, R., S. Grosskopf, and F. Hernandez-Sancho. (2000). "Environmental Performance: An Index Number Approach." *Working Paper*, Department of Economics, Oregon State University.

Färe, R., S. Grosskopf, B. Lindgreen, and P. Roos. (1994). "Productivity Developments in Swedish Hospitals: A Malmquist output index approach." In A. Charnes, W. Cooper, A.Y. Lewin, and L.M. Seiford (eds.), *Data Envelopment Analysis. Theory, Methodology and Applications*, Kluwer Academic Publishers.

Färe, R., S. Grosskopf, and C.A.K. Lovell. (1994). *Production Frontiers*. Cambridge University Press.

Färe, R., S. Grosskopf, C.A.K. Lovell, and C. Pasurka. (1989). “Multilateral Productivity Comparisons When Some Outputs are Undesirable: A Nonparametric Approach.” *The Review of Economics and Statistics*, 61(1), 90–98.

Färe, R., S. Grosskopf, and P. Roos. (1995). “Productivity and Quality Changes in Swedish Pharmacies.” *International Journal of Production Economics*, 39(1/2), 137–147.

Farrel, M.J. (1957). “The Measurement of Productive Efficiency.” *Journal of The Royal Statistical Society* Series A, 120, 253–351.

Fixler, D. and K. Zieschang. (1992). “Incorporating Ancillary Measures of Process and Quality Changes into a Superlative Productivity Index.” *Journal of Productivity Analysis*, 245–267.

Førsund, F. and L. Hjalmarsson. (1979). “Generalised Farrell Measures of Efficiency: An Application to Milk Processing in Swedish dairy plants.” *The Economic Journal*, 89, 294–315.

Førsund, F.R. (1998). “The Rise and Fall of Slacks: Comments on Quasi-Malmquist Productivity Indices.” *Journal of Productivity Analysis*, 10, 21–34.

Gowrinsankaran, G. and R.J. Town. (1999). “Estimating the Quality of Care in Hospitals using Instrumental Variables.” *Journal of Health Economics*, 18, 747–767.

Grabowbski, D.C. (2001). “Medicaid Reimbursement and the Quality of Nursing Home Care.” *Journal of Health Economics*, 20, 549–569.

Grifell-Tatje, E., C.A.K. Lovell, and J.T. Pastor. (1998). “A Quasi-Malmquist Productivity Index.” *Journal of Productivity Analysis*, 10, 7–20.

Lovell, C.A.K. and J.L. Zofío. (1997). “Graphyperbolic Efficiency and Productivity Measures.” Paper presented at the *Fifth European Workshop on Efficiency and Productivity Analysis*, Copenhagen, Denmark.

Malmquist, S. (1953). “Index Numbers and Indifference Surfaces.” *Trabajos de Estadística*, 4, 209–242.

Murray, R. (1992). “Measuring Public-Sector Output: The Swedish Report in Output Measurement.” In: Z. Griliches (ed.), *The Service Sectors*, National Bureau of Economic Research, The University of Chicago Press, pp. 517–542.

Newhouse, J.P. (1994). “Frontier Estimation: How Useful a Tool for Health Economics?” *Journal of Health Economics*, 13, 317–322.

O’Neill, L. (1998). “Multifactor Efficiency in Data Envelopment Analysis with an Application to Urban Hospitals.” *Health Care Management Science*, 1(1), 19–27.

Ozcan, Y.A. (1995). “Efficiency of Hospital Service Production in Local Markets: The Balance Sheet of U.S. Medical Armament.” *Socio Economic Planning Sciences*, 29(2), 139–150.

Ozcan, Y.A. and R.D. Luke. (1993). “A National Study of the Efficiency of Hospitals in Urban Markets.” *Health Services Research*, 28(6), 719–739.

Parasuraman, A., V. Zeithaml, and L. Berry. (1985). “A Conceptual Model of Service Quality and its Implications for Future Research.” *Journal of Marketing*, 49, 41–50.

Prior, D. and M. Sola. (2001). “Measuring Productivity and Quality Changes using Data Envelopment Analysis. An application to Catalan hospitals.” *Financial Accountability and Management*, 17(3), 219–245.

Ray, S. and E. Desli. (1997). “Productivity Growth, Technical Progress, and Efficiency Change in Industrialized Countries: Comment.” *American Economic Review*, 87(5), 1033–1039.

Reinhard, S., C.A.K. Lovell, and G. Thijssen. (1999). “Econometric Estimation of Technical and Environmental Efficiency: An Application to Dutch Dairy Farms.” *American Journal of Agricultural Economics*, 81(1), 44–60.

Rouse, J. (1997). “Resource and Performance Management in Public Service Organizations.” In K. Isaac-Henry, C. Painter, and C. Barnes (eds.), *Management in the Public Sector. Challenge and change*, 2nd edn., International Thomson Business Press.

Seiford, L.M. and J. Zhu. (2002). “Modeling Undesirable Factors in Efficiency Evaluation.” *European Journal of Operational Research*, 142, 16–20.

Simar, L. and P. Wilson. (1998). “Productivity Growth in Industrialized Countries.” *Working Paper,* Institut de Statistique and CORE, Université Catholique de Louvain, Louvain-la-Neuve.

Sloan, F.A., G.A. Picone, D.H. Taylor Jr. and S.Y. Chou. (2001). “Hospital Ownership and Cost and Quality of Care: Is there a Dime’s worth of Difference?” *Journal of Health Economics*, 20, 1–21.

Tulkens, H. and P. Vanden Eeckaut. (1995). “Non-parametric Efficiency, Progress and Measures for Panel Data: Methodological Aspects.” *European Journal of Operational Research*, 80(3), 474–499.

Tyteca, D. (1997). “Linear Programming Models for the Measurement of Environmental Performance of Firms: Concepts and Empirical Results.” *Journal of Productivity Analysis*, 8, 175–189.

Worning, A.M. (1994). “Stratégies de Réduction des Infections Nosocomiales. Un modèle Pour le Développement de la qualité.” (Reduction strategies for nosocomial infections. A quality development model). In: *la Santé: Qualité et choix*. Paris.

Zofío, J.L., and C.A.K. Lovell. (1998). “Yet Another Malmquist Productivity Index Decomposition.” *Mimeo,* Departamento de Economía, Universidad Autónoma de Madrid.