Efficiency and productivity change in Taiwan's hospitals: a non-radial quality-adjusted measurement

Chia-Hsuan Wu · Ching-Cheng Chang ·
Po-Chi Chen · Ken-Nan Kuo

Published online: 19 February 2012
© Springer-Verlag 2012

Abstract This study measures the quality-adjusted hospital efficiency and productivity index of a production unit. We propose a non-radial output-oriented directional distance function approach to analyze Taiwan's hospital productivity, which embeds the quality of care and environment variables simultaneously. There are two major advantages of this model. First, it considers all the radial and non-radial slacks that the model can identify, and hence is able to provide a more accurate performance measure and improve the discriminating power of the analysis. Second, it allows us to identify the source of the inefficiency. Our results show that the productivity indices of most of Taiwan's hospitals got worse during the 2002–2004 period, during which both technology and efficiency performance deteriorated, but divergence appeared among different types of hospitals. We confirmed the need to incorporate quality factors while measuring a hospital's efficiency and productivity. Nevertheless, there is no evidence to support the idea that healthcare quality is undermined by the cost-saving efforts by the care providers after the implantation of a global budget system.

C.-H. Wu
Taiwan WTO Center, Chung-Hua Institution for Economic Research, Taipei, Taiwan

C.-C. Chang
Institute of Economics, Academia Sinica, Taipei, Taiwan

C.-C. Chang
Socio-Economic Department, APEC Research Center for Typhoon and Society, Taipei, Taiwan

P.-C. Chen (✉)
Department of International Business, Chung Hua University,
707, Sec.2, WuFu Rd., Hsinchu 300, Taiwan, R.O.C.
e-mail: pochi@chu.edu.tw

K.-N. Kuo
Center for Health Policy Research and Development,
National Health Research Institutes, Taipei, Taiwan

Keywords Data envelopment analysis · Directional distance function · Russell measure · Slacks-based measures · Malmquist-Luenberger index · Hospitals

JEL Classification C5 · I12

1 Introduction

Rapid income growth and aging population have created new demands and high expectations on the health care systems. Many governments have allocated a considerable share of public resources to health care ([Barros et al. 2007](#)). Globally speaking, over 10% of GDP is spent on health care services ([Ozcan 2008](#)), and as time pass by, the portion is expected to be higher. Hence, it is important and necessary to address the issues of allocating limited resources and assessing the efficiency of health care services in meeting society's needs.

During the decades, a number of non-parametric and parametric methods have been applied to evaluate the performance of health care systems. In general, the measurement of changes of productivity and efficiency which provides a sensible frame for the distribution of resources in health care field is the primary method to examine the production of individual agencies ([Kontodimopoulos et al. 2006](#)). Sommersguter-Reichmann, for example, observed the changes in hospital productivity over four years by using input-based Malmquist index and its decomposed indices ([Sommersguter-Reichmann 2000](#)). Moreover, since it is difficult to assess directly the outcomes of medical and health services, some studies adopt the concept of Data Envelopment Analysis (DEA) and combine it with other gauges, for instance, the credit-points approach ([Hofmarcher et al. 2002](#)), while others may promote the technique with specific benchmarking tools for performance evaluation ([Ozcan 2008](#)). Meanwhile, [O'Neill et al. \(2007\)](#) and [Hollingsworth \(2003, 2008\)](#); [Hollingsworth and Street \(2006\)](#) conducted several comprehensive reviews on application and utility of DEA performance measures in health care systems worldwide.

Some progress has been made in developing quality indicators to compare quality across hospitals. For example, [Arocena and García-Prado \(2007\)](#) used the public hospitals data of Costa Rica to evaluate the productivity changes while controlling for quality of care. The results indicate that the main driving force of improvement is quality increase. On the other hand, several studies incorporate quality directly as an input or output into the efficiency calculations. [Färe et al. \(1995\)](#) regard quality as an input in the production process. By further imposing a separability assumption on the distance function, they establish the Malmquist productivity growth index for a sample of Swedish pharmacies and decompose the index into three components, namely quality change, technical change and efficiency change. [Maniadakis et al. \(1999\)](#) and [Chen \(2006\)](#) adopt this method to evaluate the reform of the national health services and the implementation of the national health insurance program in the UK, respectively. In a standard production model, inputs are only allowed to decrease to gain efficiency. However, in the health care sector, quality needs to be increased to improve its performance. Thus, treating quality as an input may not reflect the true production system where quantity and quality are supposed to increase simultaneously.

Alternatively, [Thanassoulis et al. \(1995\)](#) measure quality from the aspect of outcome. They treat quality as an output in gauging the performance of the district-level health authorities for providing perinatal care in the UK. [Dismuke and Sena \(2001\)](#) also define quality as an output. However, in order to take advantage of the Malmquist-Luenberger productivity index developed by [Chung et al. \(1997\)](#), in which hospitals are credited for improving the quality of service by reducing the production of undesirable outputs from a social standpoint, rather than define it as desirable output, these papers regard quality as undesirable outputs of treatment.

The purpose of this paper is twofold. First, a growing number of health care professionals around the world are increasingly frustrated by health care payment systems that do not reward efforts to improve health care quality. There is fairly widespread agreement that a good health care payment system should offer significant incentives to providers (hospitals, physicians, and others) who supply both quality and efficient care. However, methods adopted in the previous literature as mentioned above are mostly the radial approach, which has been shown to have the disadvantage of inability to account for the potential of slack in technological constraints and hence tends to overestimate the evaluated performance. Therefore, our first purpose is to establish a non-radial quality-based measurement which incorporates quality directly into the efficiency calculation model. The directional Russell measure of inefficiency proposed by [Fukuyama and Weber \(2009\)](#) will be modified into a generalized output-oriented measure which treats bad quality (e.g., the readmission rate in this study) as an undesirable output and gauges the performance of each hospital in terms of increasing good outputs while reducing the bad ones. In comparison to the traditional directional distance function approach, the proposed measure has the advantage of accounting for all slacks in desirable and undesirable output constraints and thus is able to provide a more accurate performance evaluation result.

Second, the hospital sector constitutes a particularly important ground for efficiency and productivity studies because it absorbs a large portion of public health expenditures. Beginning in 1995, Taiwan has implemented the National Health Insurance (NHI) program. At the beginning, the NHI paid medical fees to service providers on the basis of "fee-for-service". The health care expenditure grew rapidly and yielded a deficit in 1998 ([Bureau of National Health Insurance 2004](#)). It was revised to "case-payment", and the global budget (GB) system was introduced at the same time to contain medical costs while maintaining professional autonomy and service quality. The GB system involves regulations and controls on a macro level. On a micro level, there must be reform of payment and other standards to ensure reduction of waste and to implement appropriate changes in diagnostic and treatment behaviors.

The GB started in 1998 with dental care services, followed by the traditional Chinese medicine in 2000, small clinics in 2001 and finally hospitals in 2002. Under this system, a hospital's reimbursement is based on a predetermined payment for a particular diagnosis classification. Because every treated patient is awarded by a fixed payment, there is a tendency for hospitals to downgrade their service quality to lower their costs. To ensure the best possible outcomes for the GB, the NHI administrators have collaborated with the medical organizations to carry out quality

assurance programs to supervise medical care providers enrolled in the NHI system. However, as mentioned above, the healthcare quality may be undermined by the cost-saving efforts by the care providers. Consequently, our quality-based measurement is adopted to capture the effect of this payment system reform on hospitals' performance.

The remainder of this paper is organized as follows. The next section presents the proposed approach. For comparison reason, we first illustrate the traditional output-oriented directional distance function and then introduce our revised model of the quality-adjusted productivity index. Section three describes the data. Section four applies the model to evaluate the performance of Taiwan's hospitals over the period 2002–2004. The final section concludes.

2 Methods

2.1 Indicators of healthcare quality

The general indicators that are used to evaluate hospitals' quality can be divided into the clinical side and the performance side. On the clinical side, common indicators include mortality rate (or survival rate) (Thanassoulis et al. 1995; Maniadakis et al. 1999; Ho and Hamilton 2000; Dismuke and Sena 2001), readmission rate (Ashton and Wray 1996; Linna 1998; Ho and Hamilton 2000; Arocena and García-Prado 2007), average length of stay, other infection rate and complications. The performance-side indicators measure the quality of service delivery as perceived by the patients, which include satisfaction of patients (Thanassoulis et al. 1995), malpractice, medical injury, waiting or queuing time (Färe et al. 1995; Chen 2006), and so on.

Here, we use the number of readmissions¹ as a proxy for hospital quality. The control of hospital readmissions has significant implications for costs and quality because it is related to how well the physicians and hospital personnel apply their medical knowledge to the diagnosis and treatment of their patients (Epstein et al. 1991). On the cost side, Arocena and García-Prado (2007) point out that higher readmissions indicate higher medical costs because patients have to spend more time hospitalized and use more medical resources. On the other hand, it is also an indicator of quality because it reflects the fact that the ultimate goal of healthcare, which is to improve patients' well-being, has not been reached successfully (Arocena and García-Prado 2007). Therefore, it is defined as an undesirable output in this study.

2.2 The traditional directional distance function model (TDDFM)

Since we would like to take undesirable output (namely, readmissions) into account, we begin with Luenberger (1992)'s directional distance function. To specify the underlying production technology, let inputs be denoted by $x \in R_+^N$, good outputs by $y \in R_+^M$, and bad or undesirable outputs by $b \in R_+^J$. The directional distance function seeking

¹ To be more precise, this is the total number of readmissions of patients who were previously discharged from the hospital within a year.

to increase the desirable outputs while decreasing the undesirable ones directionally can be defined by the following formulation:

$$\vec{D}(x, y, b; g) = \sup \{ \beta : (x, y + \beta g_y, b - \beta g_b) \in P(x) \}, \quad (1)$$

where the nonzero vector $g = (g_y, -g_b)$ determines the “directions” in which desirable outputs and undesirable outputs are scaled, and the technology reference set $P(x) = \{(x, y) : x \text{ can produce } y\}$ satisfies the assumptions of constant returns to scale, strong disposability of desirable outputs and inputs, and weak disposability of undesirable outputs.²

Suppose at each time $t = 1, \dots, T$, there are $k = 1, \dots, K$ decision making units (DMUs) in the data set. Each DMU uses input $x^{k,t} = (x_1^{k,t}, x_2^{k,t}, \dots, x_N^{k,t}) \in R_+^N$ to jointly produce desirable outputs $y^{k,t} = (y_1^{k,t}, y_2^{k,t}, \dots, y_M^{k,t}) \in R_+^M$ and undesirable outputs $b^{k,t} = (b_1^{k,t}, b_2^{k,t}, \dots, b_J^{k,t}) \in R_+^J$. The data envelopment analysis (DEA) piecewise reference technology for time period t can be constructed as follows:

$$\begin{aligned} P(x) = \{(y, b) : & \sum_{k=1}^K z_k^t y_{mk}^t \geq y_m^t, \quad m = 1, \dots, M, \\ & \sum_{k=1}^K z_k^t b_{jk}^t = b_j^t, \quad j = 1, \dots, J, \\ & \sum_{k=1}^K z_k^t x_{nk}^t \leq x_n^t \quad n = 1, \dots, N, \\ & z_k^t \geq 0, \quad k = 1, \dots, K\}, \end{aligned} \quad (2)$$

where z_k^t are the intensity variables to shrink or expand the individual observed activities of DMU k at time period t for the purpose of constructing convex combinations of the observed inputs and outputs. The inequalities of the inputs in Eq. (2) make them freely disposable; and the same holds for the good outputs. The bad outputs are modeled with equalities due to their weak disposability property, namely, their not being freely disposable. Finally, the non-negativity constraints on the intensity variables z_k^t allow the model to exhibit constant returns to scale (Chung et al. 1997).

Relative to the reference technology $P(x)$ constructed in (2), traditionally, for each DMU $k' = 1, \dots, K$ at time period t , the directional output distance function can be obtained by solving the following linear programming problem with a constant-returns-to-scale technology:

² As pointed by Chung et al. (1997), the relationship between the Shephard's output distance function and output directional distance function is characterized as $\vec{D}_o(x, y, b; g) = (1/D_o(x, y, b)) - 1$, where $D_o(x, y, b) = \inf\{\lambda : ((y, b)/\lambda) \in P(x)\}$.

$$\begin{aligned}
\vec{D}_o^t(x^{k't}, y^{k't}, b^{k't}; g) &= \max \beta^{k't} \\
\text{s.t. } \sum_{k=1}^K z_k^t y_{mk}^t &\geq y_{mk'}^t + \beta^{k't} g_{y_m}, \quad m = 1, \dots, M, \\
\sum_{k=1}^K z_k^t b_{jk}^t &= b_{jk'}^t - \beta^{k't} g_{b_j}, \quad j = 1, \dots, J, \\
\sum_{k=1}^K z_k^t x_{nk}^t &\leq x_{nk'}^t, \quad n = 1, \dots, N, \\
z_k^t &\geq 0, \quad k = 1, \dots, K
\end{aligned} \tag{3}$$

where $\beta^{k'}$ measures the maximum expansion of desirable outputs and contraction of undesirable outputs and inputs that remain technically feasible and can serve as a measure of technical inefficiency. If $\beta^{k'} = 0$, then DMU k' operates on the frontier of $P(x)$ with technical efficiency. If $\beta^{k'} > 0$, then DMU k' operates inside the frontier of $P(x)$.

Other than being the generalization of Shephard's distance functions³, one of the important characteristics of the directional distance function is that the direction in which performance is scaled can be specified flexibly to accommodate different analysis purposes. For example, if we set $g = (g_y, -g_b) = (y^{k'}, -b^{k'})$, i.e., the direction is chosen based on the observed data, $\beta^{k'}$ represents the potential proportionate change in goods and bads. If instead we take $g = (g_y, -g_b) = (1, -1)$, then we can interpret the solution value as the net improvement in performance in terms of feasible increase in goods outputs and feasible decreases in bad outputs (Färe and Grosskopf 2004).

However, as Cooper et al. (2007a) pointed out, the traditional directional distance function, the same as many other radial measures, does not incorporate both desirable and undesirable output slacks in the output constraint when estimated using DEA. As a result, it fails to consider all the inefficiencies associated with non-zero slacks that may be identified by the model being employed. Therefore, it would have the problem of incorrectly regarding some DMUs as efficient units and leading to biased results.

2.3 The weighted Russell directional distance function model (WRDDFM)

Radial measures of efficiency overestimate technical efficiency when there are non-zero slacks in the constraints defining the piece-wise linear technology, leading researchers to construct alternative non-radial efficiency measures that account for slack (Fukuyama and Weber 2009). In the literature, the non-radial models can be roughly divided into two groups: (1) Russell measure, which was first presented by Färe and Lovell (1978) with an input-oriented form. Nevertheless, it only accounts for the slacks of inputs ignoring the inefficiencies associated with outputs. It was later revised by Färe et al. (1985) who use an additive approach to account for both input and output slacks. Later on, Pastor et al. (1999) used a ratio form to com-

³ Details of the relationship between directional distance functions and Shephard distance functions can be found in Chung et al. (1997) and Färe and Grosskopf (2000).

bine input and output slacks into a new measure called “Enhanced Russell Graph Measure” (ERGM). (2) Slacks-based model (SBM), which was proposed by [Tone \(2001\)](#) with the objective to maximize all the input and output slacks in a fractional programming form. [Cooper et al. \(2007b\)](#) showed that the SBM was equivalent to ERGM.

Recently, [Fukuyama and Weber \(2009\)](#) introduced the directional distance function technology into SBM to develop a generalized measure of technical inefficiency which also accounted for all slacks in input and output constraints. This new measure, which was referred to as the directional slacks-based inefficiency (SBI) measure, was shown to yield the same information on performance as Tone's SBM when the directional vectors for inputs and outputs are equal to the actual input and output vector, and could also be viewed as a generalization of the original Russell measure of efficiency. Because the directional-distance-function-based measure is evaluated in a linear form and hence has the advantage of easy computation and easy extension to incorporate additional undesirable outputs into the programming problem, we make the necessary modifications to incorporate the undesirable output into our Russell directional distance function model. Instead of assuming that inputs and outputs (desirable only) are freely variable as in the SBI model, we model technologies with the output (including desirable and undesirable) correspondence only as in the traditional directional distance function approach formulated above.

Consider again the inputs and outputs set described in the previous section. The proposed model can be expressed as follows:

$$\begin{aligned} \tilde{R}_o^t(x^{k't}, y^{k't}, b^{k't}; g) = \beta^R = \max w_y & \left(\sum_{m=1}^M w_m^y \beta_m^{k't} \right) + w_b \left(\sum_{j=1}^J w_j^b \beta_j^{k't} \right) \\ \text{s.t. } \sum_{k=1}^K z_k^t y_{mk}^t & \geq y_{mk'}^t + \beta_m^{k't} g_{y_m}, \quad m = 1, \dots, M, \\ \sum_{k=1}^K z_k^t b_{jk}^t & = b_{jk'}^t - \beta_j^{k't} g_{b_j}, \quad j = 1, \dots, J, \\ \sum_{k=1}^K z_k^t x_{nk}^t & \leq x_{nk'}^t, \quad n = 1, \dots, N, \\ z_k^t & \geq 0, \quad k = 1, \dots, K \end{aligned} \quad (4)$$

It is required that the directional vectors have the same units of measurement as the vectors of the observed data, so that the β 's in the objective function in (4) can be added.⁴ The measure $\tilde{R}_o^t(x^{k'}, y^{k'}, b^{k'}; g)$ given in (4) indicates that in order to become efficient, DMU k' is allowed to scale all its goods and bads hyperbolically by different rates. Therefore, $\beta_m^{k'}$ and $\beta_j^{k'}$ can serve as the individual inefficiency measure for each

⁴ [Fukuyama and Weber \(2009, p3\)](#) point out that the directional vectors should have same units of measurement as the vectors of the input and output slacks, and when $g = 1$, the role of the directional vectors is very similar to that of the e_l in [Färe and Grosskopf \(2010\)](#).

desirable output and each undesirable output. Accordingly, β^R provides us a weighted aggregate inefficiency measure of performance in a non-radial manner in which the component of $\sum_{m=1}^M w_m^y \beta_m^{k'} = \beta_y$ in the objective of (4) corresponds to the average desirable output mix inefficiencies and similarly, $\sum_{j=1}^J w_j^b \beta_j^{k'} = \beta_b$ corresponds to the average undesirable output mix inefficiencies. The weights w_m^y and w_j^b are denoted the priority or managerial preferences given to each good and bad exogenously, and their sums are assumed to be normalized to unity, i.e., $\sum_{m=1}^M w_m^y = 1$, $\sum_{j=1}^J w_j^b = 1$. Similarly, the coefficients w_y and w_b present the given priorities associated with both the desirable and undesirable outputs and their sum is also normalized to unity, i.e., $w_y + w_b = 1$.

According to Fukuyama and Weber (2009), it is easy to show that this output-oriented WRDDFM is equivalent to the output-oriented weighted directional SBI measure. Let

$$\beta_m^{k'} = \frac{s_{mk'}^+}{g_{y_m}}, \quad \beta_j^{k'} = \frac{s_{jk'}^-}{g_{b_j}}$$

where $s_{mk'}^+$, $s_{jk'}^-$ are the desirable and undesirable outputs slacks, respectively. Then, model (4) can be re-written as follows:

$$\begin{aligned} \bar{R}_o^t(x^{k't}, y^{k't}, b^{k't}; g) = \beta^R = \max w_y \left(\sum_{m=1}^M w_m^y \frac{s_{mk't}^+}{g_{y_m}} \right) + w_b \left(\sum_{j=1}^J w_j^b \frac{s_{jk't}^-}{g_{b_j}} \right) \\ \text{s.t. } \sum_{k=1}^K z_k^t y_{mk}^t \geq y_{mk'}^t + s_{mk't}^+, \quad m = 1, \dots, M, \\ \sum_{k=1}^K z_k^t b_{jk}^t = b_{jk'}^t - s_{jk't}^-, \quad j = 1, \dots, J, \\ \sum_{k=1}^K z_k^t x_{nk}^t \leq x_{nk'}^t, \quad n = 1, \dots, N, \\ s_{mk't}^+, s_{jk't}^- \geq 0, \quad m = 1, \dots, M, j = 1, \dots, J, \\ z_k^t \geq 0, \quad k = 1, \dots, K \end{aligned} \quad (5)$$

Consequently, it can be found that because the slacks for each variable are allowed to be different, the objective in (5) can help us identify all inefficiencies by calculating the maximum expansion of desirable outputs and contraction of undesirable outputs. In addition, it can be easily verified that β^R satisfies the properties of unit invariance and monotone. A smaller value of β^R indicates better performance. If there are no slacks in all outputs, i.e. $s_{mk'}^+ = 0$, $s_{jk'}^- = 0$, then β^R will attain the value of unity, indicating that k' is a technically efficient firm.

Compared to model (3), in which all the desirable and undesirable outputs are scaled by only one rate, our WRDDFM can differentiate the technical inefficiency among all outputs. This makes sense, since production efficiency for different products may be

different. Further, from a slacks accounting standpoint, our model is superior to the traditional one by being able to consider all the radial and non-radial slacks that the model can identify, and hence can provide a more accurate performance measure. On the other hand, from an economic standpoint, if one producer is inefficient, our model can tell us the source where we need to improve the most, whereas the traditional one can't. Furthermore, as [Zhou et al. \(2007\)](#) mentioned, when more information such as the preference of decision makers is available, our weighted model could incorporate the information by assigning different weights to w_y and w_b (or w_m^y and w_j^b) to reveal decision makers' production focus. For example, if quality is the main focus, then the reduction of the undesirable readmission rate is treated as more important than increasing the inpatient or outpatient visits. This information can be easily incorporated into the WRDDFM by imposing a larger w_b to the reduction of readmission.

Additionally, it is known that the uncontrollable variables (i.e., environmental variables) such as location characteristics, labour union power, government regulations, etc. ([Fried et al. 1999](#)) are not traditional inputs, but could influence the efficiency of a DMU. Therefore, if the DMUs operate in different environments, the following constraint could be added into the WRDDFM in (4) to incorporate such effects:

$$\sum_{k=1}^K z_k^t e_{hk}^t \leq e_{hk}^t, \quad h = 1, \dots, H \quad (6)$$

where e_{hk} , $h = 1, \dots, H$, are the environmental variables with a positive effect on the efficiency faced by DMU k .

2.4 Russell–Malmquist–Luenberger (RML) productivity index

Formulation (4) can help us compare the technical efficiency for each DMU under the same technology frontier. If we wish to identify their performance changes over time, by setting $g = (g_y, -g_b) = (y, -b)$, then we can extend the Malmquist–Luenberger (ML) productivity measurement to the Russell–Malmquist–Luenberger (RML) productivity index as follows:

$$\text{RML}_t^{t+1} = \left\{ \frac{(1 + \bar{R}_o^t(x^t, y^t, b^t; g))}{(1 + \bar{R}_o^t(x^{t+1}, y^{t+1}, b^{t+1}; g))} \frac{(1 + \bar{R}_o^{t+1}(x^t, y^t, b^t; g))}{(1 + \bar{R}_o^{t+1}(x^{t+1}, y^{t+1}, b^{t+1}; g))} \right\}^{1/2} \quad (7)$$

where the RML measure indicates productivity improvements if it is greater than one and productivity decrease if otherwise. Like the ML productivity index, the RML index can also be decomposed into two components, one accounting for efficiency change (EFCH), and the other technical change (TECH), as follows:

$$\text{EFCH}_t^{t+1} = \frac{1 + \bar{R}_o^t(x^t, y^t, b^t; g)}{1 + \bar{R}_o^{t+1}(x^{t+1}, y^{t+1}, b^{t+1}; g)}, \quad (8)$$

and

$$\text{TECH}_t^{t+1} = \left\{ \frac{(1 + \bar{R}_o^{t+1}(x^t, y^t, b^t; g)) \cdot (1 + \bar{R}_o^{t+1}(x^{t+1}, y^{t+1}, b^{t+1}; g))}{(1 + \bar{R}_o^t(x^t, y^t, b^t; g)) \cdot (1 + \bar{R}_o^t(x^{t+1}, y^{t+1}, b^{t+1}; g))} \right\}^{1/2}. \quad (9)$$

3 Data and variable specification

The sample used for this analysis consists of 286 hospitals in Taiwan for three consecutive years, 2002–2004. There are three outputs and four inputs. The outputs include the number of inpatients (Y1) and outpatients (Y2) as desirable outputs. The number of readmissions (B) is chosen as a proxy for the undesirable output following (Arocena and García-Prado 2007), mainly because patients' satisfaction will be decreased if they have to be readmitted. It is calculated by the number of previously discharged patients who returned to the same hospital or checked into another hospital with the same illness (based on an identification code ICD-9 in the National Health Insurance databank). Furthermore, the case-mix differences were calculated for the three outputs by using the weights based upon the Charlson Comorbidity Index (CCI).⁵

The four inputs are physicians (X1), allied personnel (X2), hospital beds as a proxy for capital (X3), and expenditures on medical supplies (X4). Table 1 shows the sample means and standard deviations of the variables. The hospitals are divided into three grades based on their appraisal by the hospital accreditation agency⁶ in Taiwan: medical centers (16 hospitals), which are usually large; regional hospitals (59 hospitals), which are medium-sized hospitals; and district hospitals (211 hospitals), which are even smaller. These data are obtained from the Taiwan National Health Insurance Databank at the National Health Research Institutes (NHRI).

Additionally, two external environmental variables relating to demographics and geographic location factors are specified. First, because Taiwan is a rapidly aging society, this situation will create an increasing demand for medical services. Thus, the concentration of the elderly population, age 65 and above, in the vicinity of the hospitals is used to control for aging discrepancies. Second, for geographic locations factor, hospitals located in rural areas may not have demand as strong as those located in urban areas due to differences in income level and lifestyle. Therefore, we also consider economic factors (*e.g.*, the average family income) to capture the local influences on hospitals' performance.

As shown in Table 1, the amounts of the inputs and outputs vary significantly not only among different grades but also among different hospitals in the same grade, especially for the district-level hospitals, because the coefficient of variances for this grade is the largest one among all three grades. This implies that the efficiency scores for these three grades may be highly diversified.

Observing the changes over time, on the input side, both of the total number of physicians (X1) and allied health personnel (X2) increased during 2002–2003, but then fell by a significant portion during 2003–2004. On the other hand, both the total

⁵ Please refer to Deyo et al. (1992) for the definition of CCI.

⁶ The agency is the Taiwan Joint Commission on Hospital Accreditation.

Table 1 Descriptive statistics for input and output variables by grade, 2002–2004

	2002		2003		2004		2002–	2003–
	Mean	SD	Mean	SD	Mean	SD	2003 (%)	2004 (%)
<i>X1 (physicians, no. of persons)</i>								
Medical center	595.6	305.0	612.4	302.5	516.4	256.2	2.8	−15.7
Regional hospital	106.6	52.8	111.8	57.0	78.3	53.3	4.9	−30.0
District hospital	15.6	15.5	15.9	15.5	12.2	13.3	2.0	−23.1
Total	66.8	153.7	69.0	157.2	54.0	132.6	3.4	−21.7
<i>X2 (allied health personnel, no. of persons)</i>								
Medical center	746.6	953.0	859.3	857.8	921.3	778.4	15.1	7.2
Regional hospital	231.3	195.7	286.5	196.0	264.2	189.2	23.9	−7.8
District hospital	50.6	59.0	60.8	64.5	52.3	55.8	20.1	−13.9
Total	128.8	293.7	152.0	296.0	144.6	290.9	19.9	−4.8
<i>X3 (no. of beds)</i>								
Medical center	1613.3	868.9	1613.3	868.9	2179.1	918.7	0.0	35.1
Regional hospital	631.2	359.3	631.2	359.3	1056.9	478.9	0.0	67.5
District hospital	140.0	160.1	140.0	160.1	225.9	278.4	0.0	61.4
Total	323.7	472.0	323.7	472.0	506.6	652.9	0.0	56.5
<i>X4 (expenditures, NT\$ millions)</i>								
Medical center	5605.1	3223.9	5439.7	3035.3	6151.6	3317.8	−3.0	13.1
Regional hospital	1087.9	635.0	1106.8	661.0	1266.6	764.6	1.7	14.4
District hospital	155.4	156.1	166.0	165.2	184.5	188.0	6.9	11.1
Total	652.6	1498.8	655.2	1448.6	741.5	1629.2	0.4	13.2
<i>Y1 (case-mix adjusted inpatients, no. of admissions)</i>								
Medical center	35745.0	22289.4	34083.8	19893.1	39817.5	22439.5	−4.6	16.8
Regional hospital	8027.8	5145.8	7791.5	5575.2	8624.4	6377.1	−2.9	10.7
District hospital	1239.1	1298.0	1181.2	1279.5	1288.2	1456.4	−4.7	9.1
Total	4570.0	9901.0	4385.6	9374.3	4957.1	10831.2	−4.0	13.0
<i>Y2 (case-mix adjusted outpatients, no. of visits)</i>								
Medical center	789656.3	482275.1	747218.8	473389.6	800906.3	429658.7	−5.4	7.2
Regional hospital	159389.8	101464.4	152661.0	98154.8	176194.9	118688.6	−4.2	15.4
District hospital	26857.8	26496.6	27002.4	26151.7	29753.6	29683.1	0.5	10.2
Total	96872.4	215027.6	93216.8	205545.6	103104.9	213676.7	−3.8	10.6
<i>B (case-mix adjusted hospital readmissions, no. of readmissions)</i>								
Medical center	906.3	880.3	840.0	617.9	1136.3	858.8	−7.3	35.3
Regional hospital	233.4	234.7	199.0	250.4	235.3	297.8	−14.7	18.2
District hospital	36.0	49.2	33.2	48.8	41.3	62.7	−7.8	24.5
Total	125.4	310.4	112.5	265.6	142.6	353.0	−10.3	26.7

number of beds (X3) and expenditures (X4) stayed almost the same over 2002–2003 and then increased substantially over 2003–2004. This suggests an input substitution during our sample period. While the changes of X3 and X4 show a similar pattern, the

reasons behind them are quite different. For X3, the increase in 2004 is the outcome of a new hospital survey and data realignment. But the stagnation of the expenditures (X4) during 2002–2003 is mainly due to the outbreak of severe acute respiratory syndrome (SARS) in 2003 in Taiwan, which kept many patients away from hospitals. The changes in output variables are somewhat different. Again, due to the outbreak of SARS in 2003, the outputs (both desirable and undesirable) decreased by 4–10% during 2003–2004, but rebounded significantly in 2004 and exceeded their previous levels.

4 Empirical results

Three modifications or specifications are made prior to our applications. First, as mentioned previously, the ratio of people aged 65 and above and the average family income are specified to reflect the external uncontrollable factors of demographics and geographic locations faced by hospitals. However, for the variable of aging population percentage, its impact is un-determined. Hospitals located in areas with higher ratios of aging population are more likely to face a higher demand for medical services and expected to result in positive effect on those hospitals, whereas, on the other hand, if aging people are sick, their readmission possibility would expected to be larger also and could cause negative effect on the hospitals. Therefore, the inequality sign in constraint (6) for this environmental variable is changed to an equal sign. As for the variable of the average family income, generally speaking, people with higher income may have stronger demand for health care since they may have more ability to afford the health care expenditure or care more about their health. In addition, they may usually be more concerned about their outpatient care services and hence not have readmission numbers as high as other people. Therefore, its impact is expected be positive on the performance of hospitals and its inequality sign in constraint (6) is kept.

Second, the weights in the objective function of the WRDDFM [i.e., w_y , w_b , w_m^y and w_j^b in Eq. (4)] are viewed as pre-specified parameters. For simplicity, we choose to weight various goods and bads equally. That is, we specify $w_y = w_b = 1/2$ and at the same time $w_m^y = 1/2$. Third, since rather than get the net improvement in performance in terms of feasible increase in goods outputs and feasible decreases in bad output, we would like to calculate the potential proportionate change, we set $g = (g_y, -g_b) = (y^{k'}, -b^{k'})$, i.e., the direction is chosen based on the observed data to solve the models.

For the following sub-sections, we first conduct two comparisons to present the need for using the WRDDFM with bad outputs for our application. Figure 1 depicts the research procedure. The software used for solving the linear programming problem is GAMS Version 2.0 and MINOS solver.

4.1 Comparison of TDDFM and WRDDFM

Using the panel data, we first compare the technical inefficiency scores obtained from the model based on the traditional direction distance function and WRDDFM to show whether the bias of not accounting for all the non-zero slacks is significant or not. Both of them include the constraints on the environmental variables. In addition, the t test

Fig. 1 The flowchart of research procedure

Table 2 Comparison of the inefficiency estimates—traditional model and WRDDFM

	2002		2003		2004	
	Traditional	WRDDFM	Traditional	WRDDFM	Traditional	WRDDFM
Mean	0.178	0.770	0.187	0.703	0.218	0.795
SD	0.209	1.528	0.222	1.583	0.247	1.389
No. of efficient units (%)	95 (33.2)	90 (31.5)	95 (33.2)	94 (32.9)	87 (30.4)	84 (29.4)
<i>T</i> test	6.871***		5.767***		7.435***	
Wilcoxon test	−12.140***		−12.016***		−12.324***	

* Significant at $Pr < 0.05$, ** Significant at $Pr < 0.01$, *** Significant at $Pr < 0.001$

and the Wilcoxon (rank sum) test are conducted to determine whether their differences in mean and ranking are significant, respectively.

As shown in Table 2, the inefficiency scores estimated from the WRDDFM are all significantly higher than those from the traditional model over the sample period. This implies that there are still a lot of slacks unrecognized under the traditional model, where the increase of good and bad outputs are scaled in a uniform manner and thus overestimate the performance of our sample hospitals. In addition, the numbers of efficient hospitals obtained from WRDDFM are less than those from the traditional model. Consequently, this means that WRDDFM not only provides us with more accurate performance evaluations, but also improves the discriminating power by screening out the inefficient hospitals which are mistakenly recognized as efficient by the traditional model.

4.2 Comparison of WRDDFM with and without bad output

The WRDDFM results obtained above have already taken the quality variable into account. Here we will compare them with the results where quality is ignored to show the importance of incorporating the quality factors. The comparisons are reported in

Table 3 Comparison of the inefficiency estimates—WRDDFM with and without bads

	2002		2003		2004	
	Without undesirable output	With undesirable output	Without undesirable output	With undesirable output	Without undesirable output	With undesirable output
Mean	1.933	0.770	2.233	0.703	2.448	0.795
SD	4.092	1.528	5.295	1.583	5.618	1.389
<i>T</i> test	−7.251***		−6.566***		−6.363***	
Wilcoxon test	−12.691***		−13.088***		−12.839***	

The same as Table 2

Table 3. For the hospitals providing better health care quality, we would expect them to have better performance ratings by crediting them for their effort to reduce the undesirable outputs.

In Table 3, we can find that, irrespective of which year, the *T* test indicates that failure to credit hospitals for reducing undesirable outputs tends to overestimate the degree of inefficiency. Hence the average inefficiency scores for the model without quality are statistically higher than those with quality. The Wilcoxon test also provides sufficient evidence that the performance rankings are significantly distorted when the quality index is ignored. Therefore, our results revealed that the failure to account for the undesirable outputs (or bad quality) would yield biases and thus unfair rankings of hospitals in Taiwan. Again, our approach provides an alternative method to deal with this measurement problem.

4.3 Overall inefficiency and its decomposition

One of our purposes is to see how Taiwan's hospitals perform after the implementation of the GB system. Table 4 shows that the average inefficiencies scores (β^R) of hospitals in Taiwan decreased from 0.770 in 2002 to 0.703 in 2003 and then increased to 0.795 in 2004. These mean that in order to become efficient, the average hospital needs to increase its numbers of inpatients and outpatients by at least 70% and reduce its readmissions by at least 70%, given its input usages over 2002–2004. This percentage looks large; nevertheless, to some extent, it reflects the existence of a great diversity within the dataset. It is worth noting that 2003 is the most efficient year among the 3 years over the sample periods. A possible explanation is the SARS outbreak in 2003. SARS is a highly infectious disease with significant in-hospital mortality, causing financial and operational disruptions in many hospitals in Taiwan. Healthcare workers involved in treating SARS cases and hospital management are particularly vulnerable. Therefore, a great deal of special assistance was allocated to improve the hospital management and emergency facilities during the later stage of the epidemic. Recall that DEA evaluates the performance of each DMU as compared to the best practice frontier constructed by the sample data. This means that if the productions for the DMUs get closer, the performance results will be more efficient and the calculated inefficiency scores will be lower. Hence, the low inefficiency score for 2003 also

Table 4 Summary statistics of overall inefficiency and its components

	β^R		β_y		β_{y1}		β_{y2}		β_b	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
286 hospitals										
2002	0.770	1.528	1.265	3.053	0.890	3.250	1.640	4.675	0.274	0.306
2003	0.703	1.583	1.173	3.178	0.816	2.621	1.530	5.494	0.234	0.292
2004	0.795	1.389	1.321	2.813	1.114	3.388	1.527	3.898	0.270	0.314
By grades										
2002										
Medical centers	0.085	0.144	0.054	0.121	0.004	0.015	0.103	0.243	0.117	0.197
Regional	0.548	1.478	0.873	2.999	0.223	0.838	1.524	5.293	0.223	0.236
District	0.883	1.582	1.467	3.157	1.144	3.728	1.789	4.660	0.300	0.324
K-W test	19.83***		20.35***		16.57***		17.23***		6.34*	
2003										
Medical centers	0.082	0.124	0.067	0.138	0.041	0.096	0.092	0.245	0.098	0.126
Regional	0.560	1.579	0.896	3.205	0.855	4.252	0.936	2.900	0.225	0.271
District	0.790	1.633	1.334	3.276	0.86	2.069	1.805	6.188	0.247	0.305
K-W test	16.47***		19.11***		5.91		15.52***		2.27	
2004										
Medical centers	0.085	0.149	0.048	0.123	0.025	0.060	0.071	0.193	0.122	0.207
Regional	0.571	1.811	0.883	3.662	0.791	4.535	0.976	3.910	0.258	0.298
District	0.912	1.282	1.539	2.610	1.287	3.122	1.791	4.010	0.284	0.323
K-W test	36.35***		40.61***		17.39***		29.45***		5.53	
By ownership										
2002										
Public	0.638	1.378	1.030	2.789	0.554	2.034	1.507	4.682	0.246	0.258
Private	0.817	1.580	1.350	3.145	1.012	3.587	1.688	4.683	0.284	0.321
M-W test	-1.13		-1.44		-2.50*		-1.28		-0.74	
2003										
Public	0.633	1.428	1.051	2.889	1.013	3.863	1.088	2.769	0.216	0.273
Private	0.729	1.638	1.217	3.282	0.744	1.998	1.689	6.189	0.241	0.299
M-W test	-0.67		-0.96		-0.37		-1.57		-0.28	
2004										
Public	0.677	1.664	1.093	3.358	1.137	4.460	1.050	3.533	0.261	0.312
Private	0.838	1.276	1.403	2.592	1.106	2.918	1.699	4.016	0.273	0.316
M-W test	-2.77**		-3.34**		-1.94		-3.33**		-0.68	

The same as Table 2

reflects the fact that most people try to stay away from hospitals as much as possible at that time, especially the larger medical centers where most SARS cases were treated. If we overlook this, then the efficiencies over 2002–2004 would turn out to show a slightly declining trend.

The overall inefficiencies score, β^R , for each hospital is solved by aggregation of the inefficiency of each good and bad outputs. Therefore it allows us to examine where

the inefficiency comes from. According to the β_y and β_b in Table 4, the inefficiency is mainly from the desirable outputs instead of the ability to control quality, because the average inefficiency scores of β_y are greater than β_b for each year. In addition, the pattern of their changes over the period is very similar to that of the change of overall inefficiencies. This indicates that there is no evidence to support the notion that healthcare quality is undermined by the cost-saving efforts by the care providers since the implantation of the GB system. Due to the larger impact of good outputs on overall inefficiency, hospitals on average should make more efforts to expand their good outputs (especially hospitals with lower healthcare demand). Otherwise they should devote more effort to decrease their input usages.

Further, if we further look at the two components of β_y , the inefficiency scores for outpatients (β_{y2}) are larger than those of inpatients (β_{y1}). This indicates that improving the efficiency of outpatients would be more effective for the overall performance than improving the efficiency of inpatients. However, β_{y2} declines slightly over time while β_{y1} increases so that they are converging. The worsening of the inpatient efficiency is probably related to the increase in the number of beds, which may reduce the occupancy rate.

Finally, we proceed with the Kruskal-Wallis (K-W) test for three or more samples and the Mann-Whitney (M-W) test for two samples to look for statistic differences in efficiency among different accreditation grades (medical centers, regional, and district level) and ownership (public and private) of hospitals, respectively. Note that hospitals with different grades often correspond to different sizes. Therefore, the comparison among different grades can be viewed as the comparison among different operational scales. On the other hand, the performance difference across ownership groups is an important issue in the literature. The property rights theory suggests that public enterprises usually perform less efficiently than private enterprises do (Boardman and Vining 1989).⁷ Sueyoshi and Goto (2012) argue that a firm under public ownership can easily access public capital so that it can reduce capital cost (or incur less operation cost). In contrast, private ownership often faces high capital cost. Therefore, the the impact of ownership on efficiency is undetermined. Our comparison will enable us to verify the linkage of ownership with performance in light of the above hypotheses. The results are also summarized in Table 4.

It can be found that the performance in terms of good outputs of the regional and district hospitals is significantly worse than that of the medical centers, while the performance in terms of readmissions is not statistically different by operational scale. This result is consistent with the patients' preferences for larger hospitals with higher grade of appraisal. The substantial discrepancy between medical centers and the other two smaller hospitals also suggests that it is unfair to compare them without controlling for the size effect. Since the medical centers are usually endowed with more modern

⁷ According property rights theory, owners of private firms have an incentive to incur the cost of monitoring the actions of management, since concentrated ownership makes monitoring of management economical and close monitoring induces managers to maximize the wealth of the private owners (Hinds et al. 2005). Whereas the owners (i.e. citizens) of a public firm have less incentive to incur monitoring costs since any additional wealth spawned by that investment will be dispersed among all owners. Thus, public managers will not act to maximize the wealth of the owners.

Table 5 Tobit regression results of overall inefficiency and its components

Explanatory variables	β^R	β_y	β_{y1}	β_{y2}	β_b
CONSTANT	-1.177***	-2.6614***	-6.656***	-5.722***	-0.151*
REG	1.227***	1.839**	2.214*	3.373**	0.261***
DIS	1.599***	3.032***	4.325***	5.210***	0.303***
OWN	0.076	0.259	0.202	0.889	0.024
Y2003	-0.097	-0.282	0.297	-0.947	-0.068
Y2004	0.051	-0.020	0.767	-0.652	-0.017

REG and DIS are the two dummy variables applied to reflect the effects of accreditation grades, in which REG takes the value of one if a hospital is regional and zero otherwise, and DIS takes the value of one if a hospital is district and zero otherwise. OWN is also a dummy variable in which it takes the value of one if a hospital is private and zero otherwise. Y2003 and Y2004 are two period dummies for periods 2003 and 2004

The same as Table 2

health care facilities and resources, their ability to attract patients is naturally superior to other hospitals. Therefore, we divide the sample into two sub-samples by grade to examine whether it is reasonable to combine the sample into one group. The details are described in the Appendix.

On the aspect of ownership, although the inefficiency scores for the public hospitals are lower than those for the private hospitals, the differences are not significant except for the year 2004. Previous literature (e.g., Wang and Lee 2004) indicate that private hospitals in Taiwan are in general more efficient than the public ones. Our finding, however, echoes Chen (2006)'s finding in that the efforts to improve efficiency in public hospitals cannot be overlooked and that the performance of public hospitals has gradually caught up with that of private hospitals. Observing the components of β^R over time, we find the discrepancy between the private and public hospitals is mainly from good outputs, especially the inpatient services in 2003 and outpatient services in 2004. This indicates that the public hospitals have evolved from providing more inpatient services into more outpatient services.

We also employ a multivariate approach based on Tobit regression to identify the influences of these two factors, accreditation grades and ownership. The results are reported in Table 5. Note that the dependent variables in the regression model are inefficiency measures, in which a greater value of the measure indicates a larger degree of inefficiency, a positive (negative) sign of the estimated coefficient represents a negative (positive) effect on efficiency. It can be found that the results are basically the same as those obtained from the K-W and M-W tests.

4.4 Productivity changes

In order to analyze the performance change over periods, we further calculate the RML productivity index and its decomposition components of efficiency change (EFCH) and technical change (TECH), according to Eqs. (7)–(9). Table 6 shows the geometric mean results for the whole sample and the results classified by accreditation grades

Table 6 Summary statistics of Malmquist index and its components over 2002–2004

	2002–2003						2003–2004					
	RML		TECH		EFFI		RML		TECH		EFFI	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
286 hospitals	0.956	0.299	0.934	0.131	1.024	0.365	1.000	0.335	1.060	0.234	0.943	0.315
<i>By grade</i>												
Medical centers	1.019	0.093	1.018	0.068	1.001	0.094	0.994	0.104	0.995	0.065	0.999	0.097
Regional	0.976	0.129	0.979	0.089	0.997	0.132	0.997	0.301	0.981	0.147	1.016	0.286
District	0.946	0.340	0.916	0.140	1.033	0.416	1.001	0.356	1.088	0.254	0.920	0.331
<i>By ownership</i>												
Public	0.963	0.222	0.964	0.147	0.998	0.262	0.990	0.314	0.987	0.179	1.004	0.337
Private	0.954	0.323	0.923	0.123	1.033	0.395	1.003	0.344	1.088	0.246	0.922	0.304

and ownership. Recall that when the index is greater than 1, it indicates productivity growth, while less than 1 means a regress.

According to Table 6, overall productivity growth declined during 2002–2003 but then rebound during 2003–2004. These results are consistent with what we observe in Table 4, which indicates that 2003 is the most efficient year. The 2003 SARS epidemic kept many patients away from hospitals, and almost all hospitals' outputs shrank simultaneously, causing the production frontier to move backwards. However, the 2003 SARS outbreak had different influences on different groups of hospitals. For example, the productivity of medical centers improved, but the district level hospitals suffered from deterioration in both technology and productivity growth. This also implies that the GB may have contributed to the concentration in the provision of hospital services. Larger hospitals or medical centers can reduce their costs through the operation of economies of scale and at the same time improve patient outcomes by increasing average volumes of activity by clinicians.

To set aside the interruptions from the 2003 SARS epidemic⁸, we drop the abnormal data from the sample data for this end. That is, we directly compare the productivity change between 2002 and 2004 and ignore the year 2003. Table 7 shows that the RML and its two components, TECH and EFFI, are all less than 1. This indicates that the productivity growth of health care provision for most of the hospitals in Taiwan disappeared during 2002–2004. As for different groups of hospitals, the productivity changes still differ. The productivity of medical centers improves due to the advancement of technology. In contrast, the productivity of regional hospitals benefits from improvement in efficiency. However, for district hospitals, both of their technology and efficiency deteriorate and hence their RMLs turn out to be the worst among all hospitals. The K-W test indicates that there is no significant difference in all the RML and its two components among these three groups. The difference in the productivity

⁸ One of the referees reminds us that SARS might not only have impacts on the year of 2003 but also 2004, in spite that SARS epidemic occurred only in 2003. Therefore, the results of 2004 may be biased to some extent.

Table 7 Summary statistics of Malmquist index and its components over 2002 and 2004

	RML		TECH		EFFI	
	Mean	SD	Mean	SD	Mean	SD
286 hospitals	0.945	0.327	0.978	0.185	0.966	0.393
No. of >1	131		127		107	
No. of <1	149		152		129	
No. of = 1	6		7		50	
<i>By grade</i>						
Medical centers	1.016	0.092	1.016	0.090	1.000	0.022
Regional	1.002	0.208	0.990	0.122	1.013	0.195
district	0.924	0.364	0.972	0.204	0.950	0.446
K-W test	2.701		2.216		2.384	
<i>By ownership</i>						
Public	0.957	0.278	0.955	0.187	1.002	0.413
Private	0.941	0.344	0.987	0.184	0.953	0.386
M-W test	-0.838		-0.618		-1.271	

The same as Table 2

Table 8 OLS regression results of Malmquist Index and its components over 2002 and 2004

Explanatory variables	RML	TECH	EFFI
CONSTANT	1.014 * **	1.005 * **	1.032 * **
REG	0.005	-0.015	0.018
DIS	-0.032	-0.024	0.033
OWN	0.008	0.021	-0.046

The same as Table 5

of private and public ownership from the results of M-W test is also insignificant. The regression results reported in Table 8 give the same conclusions.

5 Discussion and conclusions

In this article, we establish a non-radial DEA approach based on output-oriented directional distance function WRDDFM to analyze Taiwan's hospital efficiency and productivity growth. We also incorporate the service quality and environment variables simultaneously into the WRDDFM. Compared with the traditional directional distance function approach, our model dominates the traditional one by considering all the radial and non-radial slacks. Thus our model can provide more accurate performance measures and better discriminating power than the traditional model

In 2002, Taiwan's NHI started the GB payment system, with all hospitals participating. Up to now, we do not have sufficient evidence or longitudinal data to support the notion that the GB system is related to a hospital's efficiency and productivity. The empirical findings conclude that, on average, the hospital productivity performance

has decreased over the sample period of 2002–2004. However, some variations exist. First, we find that efficiency deteriorated after the overall implementation of the GB reform. This result corresponds, to some extent, to what [Feldman and Lobo \(1997\)](#) have found: that the GB allows hospital administrators to control the allocation of their budget as they decide service quantity and resource intensity to maximize their own utility. The quantity of care provided may be less than the quantity demanded by the patients because many hospitals may begin to limit their services as a strategic reaction. Previous literature ([Dredge 2004](#)) has also pointed out that the GB alone may not be sufficient to ensure that efficiency will be improved. However, the GB still has the advantage of controlling total spending in a hospital. It can also encourage changes in service delivery patterns and create incentives to reward quality care and appropriate clinical practices. To some extent, the GB system can combine administrative simplicity with strong incentives for performance enhancements. As such, it can deliver real progress in a cost-effective way. Second, according to the source of the inefficiency, improving the outpatient care delivery is suggested to be a more effective way than improving the inpatient care provisions for the sake of overall efficiency improvement. Third, our results indicate that the hospital's performance is very sensitive to the quality. Nevertheless, there is no evidence to support the idea that health-care quality is undermined by the cost-saving efforts by the care providers since the implantation of the GB system. Fourth, our results demonstrate that accreditation and ownership of hospitals affect the efficiency but not the productivity change. Finally, the impact of SARS on the performance of hospitals is obvious.

We also want to mention a few limitations of Taiwan's National Health Insurance database. First, we need to obtain data over a longer period to support our analysis. However, the data are not publicly available. So far, we can only obtain the database from 2002 to 2004. Second, [Donabedian \(1980\)](#) distinguishes at least three components of health care quality: technical aspects, interpersonal aspects, and the amenities of care. However, due to data limitations, we can only use the number of readmissions as a proxy for the quality of service, which is very similar to the "technical" dimension of medical quality, ignoring the other components of service quality. To conclude, our study confirms the need to incorporate quality factors when measuring a hospital's productivity and the potential for incorporating quality as an output into the productivity measurement. Future studies can be extended to incorporate the other two components of quality.

Appendix

We divide the sample into two sub-samples by grade, in which medical centers and regional hospitals are put together in one group, since the number of the medical centers is only 16 and it is too few as compared to the total dimensions of the inputs and outputs to be classified as one group alone, and district hospitals alone are treated as another group. We solve these two sub-samples separately to examine whether it is suitable for us to use a single DEA to assess the performance for the hospitals altogether. The results are displayed in Table 9. It can be seen that the average inefficiency scores for medical centers and district hospitals do not exhibit too much difference as

Table 9 Summary statistics of overall inefficiency and its components by separating DEA problems

	β^R		β_y		β_{y1}		β_{y2}		β_b	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
<i>2002</i>										
Medical centers	0.081	0.137	0.048	0.119	0.004	0.018	0.090	0.239	0.114	0.193
Regional	0.072	0.106	0.031	0.072	0.017	0.064	0.045	0.129	0.112	0.169
District	0.852	1.591	1.423	3.163	1.140	3.728	1.706	4.656	0.281	0.327
<i>2003</i>										
Medical centers	0.082	0.124	0.067	0.138	0.041	0.096	0.092	0.245	0.098	0.126
Regional	0.112	0.163	0.057	0.160	0.048	0.134	0.066	0.215	0.168	0.249
District	0.761	1.641	1.303	3.281	0.856	2.065	1.751	6.192	0.219	0.305
<i>2004</i>										
Medical centers	0.077	0.141	0.046	0.127	0.026	0.066	0.066	0.194	0.109	0.191
Regional	0.500	1.707	0.803	3.444	0.645	3.589	0.961	4.045	0.197	0.280
District	0.866	1.279	1.475	2.582	1.257	3.101	1.693	3.932	0.257	0.325

compared to results reported in Table 4 for every three years, while the performance values obtained for regional hospitals show substantial improvement after the removal of the district hospitals from the sample for 2002 and 2003 years. This suggests that the results of medical centers and district hospitals are not affected significantly by the grouping of the sample; nevertheless, the peers of many of the regional hospitals are very likely to belong to district hospitals and hence their performance assessment results are distorted by the exclusion of district hospitals from the sample. Consequently, we still choose to include all the hospitals together in a DEA problem.

Acknowledgments The authors would like to thank the Editor-in-Chief Rudolf Vetschera and two anonymous reviewers for their constructive comments and suggestions that resulted in improvements to this paper.

References

- Arocena P, García-Prado A (2007) Accounting for quality in the measurement of hospital performance: evidence from Costa Rica. *Health Econ* 16:667–685
- Ashton CM, Wray NP (1996) A conceptual framework for the study of early readmission as an indicator of quality of care. *Soc Sci Med* 43(11):1533–1541
- Barros CP, Menezes A, Vieira JAC, PeyPOCH N, Solonandrasana B (2007) An analysis of hospital efficiency and productivity growth using the Luenberger productivity indicator. IZA Discussion Paper No. 2689
- Boardman AE, Vining AR (1989) Ownership and performance in competitive environments: a comparison of the performance of private, mixed, and state-owned enterprises. *J Law Econ* 32(1):1–33
- Bureau of National Health Insurance (2004) National Health Insurance annual statistical report. Bureau of National Health Insurance, Taipei, Taiwan, pp 94–99
- Chen SN (2006) Productivity changes in Taiwanese hospitals and the national health insurance. *Serv Ind J* 26(4):459–477
- Chung YH, Färe R, Grosskopf S (1997) Productivity and undesirable outputs: a directional distance function approach. *J Environ Manag* 51:229–240
- Cooper WW, Huang Z, Li SX, Parker BR, Pastor JT (2007a) Efficiency aggregation with enhanced Russell measures in data envelopment analysis. *Socio-Econ Plan Sci* 41:1–21

- Cooper WW, Seiford LM, Tone K (2007b) Some models and measures for evaluating performances with DEA: past accomplishments and future prospects. *J Product Anal* 28:151–163
- Deyo RA, Cherklin DC, Ciol MA (1992) Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. *J Clin Epidemiol* 45:613–619
- Dismuke CE, Sena V (2001) Is there a trade-off between quality and productivity? The case of diagnostic technologies in Portugal. *Ann Oper Res* 107:101–116
- Donabedian A (1980) The definition of quality and approaches to its assessment. Health Administration Press, Ann Arbor
- Dredge R (2004) Hospital global budgeting. HNP discussion paper. Available from: <http://siteresources.worldbank.org/HEALTHNUTRITIONANDPOPULATION/Resources/281627-1095698140167/DredgeHspGblBdgtFinal.pdf>. Accessed May 2008
- Epstein AM, Bogen J, Dreyer P, Thorpe KE (1991) Trends in length of stay and rates of readmission in Massachusetts: implications for monitoring quality of care. *Inquiry* 28:19–28
- Färe R, Lovell CAK (1978) Measuring the technical efficiency of production. *J Econ Theory* 19(1):150–162
- Färe R, Grosskopf S (2000) Network DEA. *Socio-Econ Plan Sci* 34:35–49
- Färe R, Grosskopf S (2004) Modeling undesirable factors in efficiency evaluation: comment. *Eur J Oper Res* 157:242–245
- Färe R, Grosskopf S (2010) Directional distance functions and slacks-based measures of efficiency. *Eur J Oper Res* 200(1):320–322
- Färe R, Grosskopf S, Lovell CAK (1985) The measurement of efficiency of production. Kluwer, Boston
- Färe R, Grosskopf S, Roos P (1995) Productivity and quality changes in Swedish pharmacies. *Int Prod Econ* 39:137–147
- Feldman R, Lobo F (1997) Global budgets and excess demand for hospital care. *Health Econ* 6:187–196
- Fried HO, Schmidt SS, Yaisawarng S (1999) Incorporating the operating environment into a nonparametric measure of technical efficiency. *J Product Anal* 12:249–267
- Fukuyama H, Weber WL (2009) A directional slacks-based measure of technical inefficiency. *Socio-Econ Plan Sci* 43(4):274–287
- Hinds S, Sanchez N, Schap D (2005) Public enterprise: retrospective review and prospective theory. In: Backhaus JG, Wagner RE (eds) *Handbook of public finance*. Springer, Berlin, pp 277–300
- Ho V, Hamilton BH (2000) Hospital mergers and acquisitions: does market consolidation harm patients? *J Health Econ* 19:767–791
- Hofmarcher MM, Paterson I, Riedel M (2002) Measuring hospital efficiency in Austria—a DEA approach. *Health Care Manag Sci* 5:7–14
- Hollingsworth B (2003) Non-parametric and parametric applications measuring efficiency in health care. *Health Care Manag Sci* 6:203–218
- Hollingsworth B (2008) The measurement of efficiency and productivity of health care delivery. *Health Econ* 17:1107–1128
- Hollingsworth B, Street A (2006) The market for efficiency analysis of health care organizations. *Health Econ* 15:1055–1059
- Kontodimopoulos N, Nanos P, Niakas D (2006) Balancing efficiency of health services and equity of access in remote areas in Greece. *Health Policy* 76:49–57
- Linna M (1998) Measuring hospital cost efficiency with panel data models. *Health Econ* 7:415–427
- Luenberger DG (1992) Benefit function and duality. *J Math Econ* 21:461–481
- Maniadamakis N, Hollingsworth B, Thanassoulis E (1999) The impact of the internal market on the hospital efficiency, productivity and service quality. *Health Care Manag Sci* 2:75–85
- O'Neill L, Rauner M, Heidenberger K, Kraus M (2007) A cross-national comparison and taxonomy of DEA-based hospital efficiency studies. *Socio-Econ Plan Sci* 42(3):158–189
- Ozcan YA (2008) Health care benchmarking and performance evaluation using data envelopment analysis (DEA). Springer, Newton. ISBN: 978-0-387-75447-5
- Pastor JT, Ruiz JL, Sirvent I (1999) An enhanced DEA Russell graph efficiency measure. *Eur J Oper Res* 115:596–607
- Sommersguter-Reichmann M (2000) The impact of the Austrian hospital financing reform on hospital productivity: empirical evidence on efficiency and technology changes using a non-parametric input-based Malmquist approach. *Health Care Manag Sci* 3(4):309–321
- Sueyoshi T, Goto M (2012) Data envelopment analysis for environmental assessment: comparison between public and private ownership in petroleum industry. *Eur J Oper Res* 216:668–678

- Thanassoulis E, Boussofiane A, Dyson RG (1995) Exploring output quality targets in the provision of perinatal care in England using data envelopment analysis. *Eur J Oper Res* 80:588–607
- Tone K (2001) A slack based measure of efficiency in data envelopment analysis. *Eur J Oper Res* 130: 498–509
- Wang YH, Lee WF (2004) Technical efficiency of district hospitals in Taiwan: nonparametric data envelopment analysis. *Taipei Econ Inq* 40:61–95
- Zhou P, Poh KL, Ang BW (2007) A non-radial DEA approach to measuring environmental performance. *Eur J Oper Res* 178:1–9