MEASURING THE PRODUCTIVITY GROWTH AND QUALITY CHANGES OF TEACHING AND NON-TEACHING HOSPITALS IN TAIWAN

Kuan-Chen Chen[1], Chung-I Lin[2] and Yu-Jen Hsiao[3],*

Abstract. The objective of this study was to compare productivity growth and quality changes in teaching and non-teaching hospitals in Taiwan using a quality-based Malmquist productivity index under a metafrontier framework. Researchers have argued that teaching hospitals are more costly than non-teaching hospitals. Whether Taiwanese teaching hospitals operate with higher productivity than non-teaching hospitals is important to determine given increasing financial pressure on hospitals. The results of this study show that, among the Taiwanese teaching and non-teaching hospitals in the study sample, both types of institutions increased in productivity and improved in quality from 2008 to 2014. However, non-teaching hospitals saw much greater positive developments in productivity and quality than teaching hospitals. Therefore, additional teaching and research duties may influence the performance of teaching hospitals in provision of direct patient care.

Mathematics Subject Classification. 62, 65, 90.

1. Introduction

Teaching hospitals are medical institutions that provide internships, practical experience, and research opportunities to medical and nursing students, as well as to students in other medicine-related fields. There are many rules governing the establishment of teaching hospitals, and the standards are different from those for non-teaching hospitals. The prerequisites for the establishment of a teaching hospital are equipment and sites for teaching and research, a good clinical training environment, and a complete teacher training system with a discrimination index for assessing the quality of teaching [27,29]. Because of these additional standards, the operating costs of teaching hospitals are usually higher than those of non-teaching hospitals.

Cameron [5] assessed the direct and indirect costs of teaching programs. The direct costs included higher salaries and additional benefits for residents, as well as the costs of teaching- and research-related equipment and activities. The indirect costs included the increased use of ancillary services and increased lengths of stay, as well as the overhead allocated to teaching and administration units. Holding patient case mix constant, the study showed that university hospitals are 33% more costly than their non-teaching counterparts. Furthermore,

Rich *et al.* [26] found that teaching hospitals are 9%–30% more costly than non-teaching hospitals. The general consensus is thus that teaching activities add to hospital costs. Burke *et al.* [4] investigated the costs of hospitalizations among U.S. Medicare beneficiaries 65 years and older and found that major teaching hospitals had higher initial hospitalization costs than non-teaching hospitals, but the total costs of care at 30 days were much lower at major teaching hospitals. In Taiwan, teaching hospitals sometimes receive government subsidies to help cover the expenses of training programs. The present study further examined whether the additional responsibilities of teaching hospitals affect their health care provision.

Teaching hospitals in Taiwan are usually affiliated with a university's college of medicine or are commissioned by the government to partner with a university's college of medicine. To ensure the quality of medical education, teaching hospitals must pass an accreditation process. Accreditation requires that faculty meet a high standard for medical abilities and demonstrate intent to cultivate outstanding medical personnel through their teaching, and thus to further improve the quality of medical services. Therefore, to a certain extent, teaching hospitals have better conditions, *i.e.*, more qualified staff and environments more conducive to learning, than non-teaching hospitals, an advantage that should have a positive effect on medical services. However, the "master-apprentice system" through which teaching hospitals provide clinical training may have a negative effect on the quality of medical services. During training, medical interns or interns of other medicine-related departments shadow physicians or other medical personnel as they take care of patients. Although such a model provides medical interns with a complete clinical training environment, the ability of full-time medical personnel to provide efficient, high-quality care may be influenced by their concurrent responsibility to teach interns. The training system may also negatively affect service hours. Therefore, the aim of this study was to perform an in-depth investigation into the differences in health care quality between teaching hospitals and non-teaching hospitals.

A few previous studies have explored the impact of teaching programs on hospital performance. Grosskopf *et al.* [16] found an average inefficiency score of 0.80 for sample teaching hospitals, indicating that these hospitals could have reduced inputs by 20% while maintaining output levels. The congestion of residents accounted for 20% of the total inefficiency score. Grosskopf *et al.* [17] found that only approximately 10% of the teaching hospitals in their study could effectively "compete" with the non-teaching hospitals in the provision of patient services. Sloan *et al.* [28] and Jensen and Morrisey [19] have suggested that teaching medical residents may negatively affect the productivity of other hospital inputs, resulting in increased costs or reduced revenue. For example, teaching responsibilities among physicians and nurses may detract from their roles in direct patient care. By contrast, in a study of Veterans Affairs hospitals, Campbell *et al.* [6] found no difference in the financial performance of teaching and non-teaching hospitals. Nuti *et al.* [23] also found that performance was not affected by being in the university hospitals rather than the general hospitals. Therefore, the effect of teaching programs on hospital efficiency and productivity is still worthy of in-depth investigation.

The outputs of medical education are difficult to determine when the education is in conjunction with patient care [25]. Most previous studies have focused on analyzing the production efficiency of teaching hospitals rather than the differences in performance between teaching and non-teaching hospitals. This study aims to fill the gap in hospital research by investigating the performance differences in patient care services. Proponents of teaching hospitals frequently claim that care quality differentiates teaching from non-teaching hospitals. Shahian *et al.* [27] found that teaching intensity was favorably associated with Surgical Care Improvement Project (SCIP) performance, acute myocardial infarction (AMI) and heart failure process scores, and mortality. Burke *et al.* [3] noted that among hospitalizations for U.S. Medicare beneficiaries, the status of major teaching hospital was associated with lower mortality rates for common conditions compared with non-teaching hospitals. However, given resource constraints in medical care, teaching programs may force hospitals to make trade-offs between the quality and quantity of care provided [14, 31]. Therefore, in this study we incorporated measures of quality into overall measures of productivity, and we then analyzed levels of productivity and quality over time among teaching and non-teaching hospitals.

In Taiwan, teaching and non-teaching are the two main hospital types. The conventional Malmquist productivity index (MPI) is not suitable for comparing the performance of a hospital operating with one type of technology to that of another hospital using a different type of technology [11]. Therefore, in this study, we used

an MPI incorporating quality attributes under a metafrontier framework, as recently proposed by Chen *et al.* [10]. This technique enabled us to evaluate the determinants of productivity growth and to compare teaching and non-teaching hospitals with regard to quality improvements and productivity growth.

Section 2 of this article introduces the MPI with quality attributes under the concept of a metafrontier model. Section 3 briefly describes our sample data. Section 4 presents the main findings based on the empirical results, and the last section offers conclusions.

2. Methodology

In this study, a metafrontier MPI (MMPI) including a quality index was used to measure the productivity growth of the teaching and non-teaching hospitals [10]. Suppose that there are K groups in the sample and that each group k at time t has I hospitals, which use an input vector $x_t^k \in R_+^N$ to produce an output vector $y_t^k \in R_+^M$ and a quality vector $a_t^k \in R_+^J$ under the production technology set S_t^k. An input-oriented data envelopment analysis (DEA) model was formulated as it focuses on minimizing the use of inputs for producing the given amount of outputs. This model fits with the context of hospitals since these hospital managers cannot seek output increment (*i.e.*, service volume or cases), but rather reduced input usage (*i.e.*, staff or beds) [8,15,22]. Therefore, an input-oriented DEA model is appropriate. The following is an input-oriented DEA model in which the inputs are minimized and the outputs are kept at their current levels:

$$D_t^k \left(x_t^k, y_t^k, a_t^k \right) = \sup \left\{ \lambda : \left(x_t^k / \lambda, y_t^k, a_t^k \right) \in S_t^k \right\} \tag{2.1}$$

where D_t^k represents the distance function of group k measuring the maximal feasible reduction in x_t^k given $\left(y_t^k, a_t^k \right)$.

Färe *et al.* [12] defined an MPI incorporating quality attributes as the geometric mean of two ratios of distance functions based on the benchmark period t and $t+1$. The MPI of group k is as follows:

$$\mathrm{MPI}_{t,t+1}^k = \left[\frac{D_t^k \left(x_t, y_t, a_t \right)}{D_t^k \left(x_{t+1}, y_{t+1}, a_{t+1} \right)} \frac{D_{t+1}^k \left(x_t, y_t, a_t \right)}{D_{t+1}^k \left(x_{t+1}, y_{t+1}, a_{t+1} \right)} \right]^{\frac{1}{2}}. \tag{2.2}$$

The quality change (QC) or quality index of group k from period t to $t+1$ is defined as

$$\mathrm{QC}_{t,t+1}^k = \left[\frac{D_t^k \left(x_t, y_t, a_t \right)}{D_t^k \left(x_t, y_t, a_{t+1} \right)} \frac{D_{t+1}^k \left(x_{t+1}, y_{t+1}, a_t \right)}{D_{t+1}^k \left(x_{t+1}, y_{t+1}, a_{t+1} \right)} \right]^{\frac{1}{2}}. \tag{2.3}$$

$\mathrm{QC}_{t,t+1}^k \gtreqqless 1$ is determined by whether quality improvement, stagnation, or decline occurs between two periods.

As hospitals seek to maximize their quality of care, constrained resources may force them to make some tradeoffs between outputs of the quantity and quality of care [14]. Therefore, it would be more reasonable to identify quality measures as dependent attributes with inputs/outputs in the analytical structure when decomposing MPI. In this respect, distance functions are assumed to not be multiplicatively separable in attributes and inputs/outputs. Assuming that production technology exhibits constant returns to scale, equation (2.3) can be decomposed as

$$\mathrm{MPI}_{t,t+1}^k = \mathrm{TEC}_{t,t+1}^k \times \mathrm{TC}_{t,t+1}^k \tag{2.4}$$

where

$$\mathrm{TEC}_{t,t+1}^k = \left[\frac{D_t^k \left(x_t, y_t, a_t \right)}{D_{t+1}^k \left(x_{t+1}, y_{t+1}, a_{t+1} \right)} \right] \quad \text{and} \tag{2.5}$$

$$\mathrm{TC}_{t,t+1}^k = \left[\frac{D_{t+1}^k \left(x_{t+1}, y_{t+1}, a_{t+1} \right)}{D_t^k \left(x_t, y_t, a_t \right)} \frac{D_{t+1}^k \left(x_t, y_t, a_t \right)}{D_t^k \left(x_{t+1}, y_{t+1}, a_{t+1} \right)} \right]^{\frac{1}{2}}. \tag{2.6}$$

Equation (2.5) measures technical efficiency change (TEC) and equation (2.6) measures technical change (TC). $\text{TEC}_{t,t+1}^k > 1 (<1)$ indicates enhancement (deterioration) in the efficiency of the production unit. TC captures the shift in the production frontier, meaning that technical progress (regress) is indicated as $\text{TC}_{t,t+1}^k > 1 (<1)$. Accordingly, MPI, the quality index, and the corresponding components for DMUi can be determined by the related input distance functions, which can be obtained by solving the linear programming problems with the CRS technology. The related input distance functions can be calculated as follows:

$$\left[D_l^k\left(x_p, y_p, a_s\right)\right]^{-1} = \min \left\{\theta_i \left| \begin{array}{ll} \sum\limits_{j=1}^{J} \lambda_{l,j}^k x_{l,nj} \leq \theta_i x_{p,ni}, & n = 1, \ldots, N; \\ \sum\limits_{j=1}^{J} \lambda_{l,j}^k y_{l,mj} \geq y_{p,mi}, & m = 1, \ldots, M; \\ \sum\limits_{j=1}^{J} \lambda_{l,j}^k a_{l,hj} \geq y_{s,hi}, & h = 1, \ldots, H; \\ \lambda_{l,j}^k \geq 0, & j = 1, \ldots, J \end{array} \right. \right\},$$

where l, p, $s = t$, $t + 1$.

The metafrontier is then assumed to take the same technology as the individual stochastic frontiers in the various groups [1]. The upper boundary then faced by all the hospitals is the metafrontier rather than the individual group frontier. Chen *et al.* [10] developed a quality-based MMPI that includes the dynamic aspects of the metafrontiers in efficiency measurement. The model is formulated as

$$\text{MMPI}_{t,t+1} = \left[\frac{D_t^*\left(x_t, y_t, a_t\right)}{D_t^*\left(x_{t+1}, y_{t+1}, a_{t+1}\right)} \frac{D_{t+1}^*\left(x_t, y_t, a_t\right)}{D_{t+1}^*\left(x_{t+1}, y_{t+1}, a_{t+1}\right)}\right]^{\frac{1}{2}}. \tag{2.7}$$

All distances involved in equation (2.7) are input-oriented distance functions based on the metafrontier and denoted by " * ". The MMPI can be decomposed as

$$\text{MMPI}_{t,t+1} = \text{TEC}_{t,t+1}^* \times \text{TC}_{t,t+1}^* \tag{2.8}$$

where

$$\text{TEC}_{t,t+1}^* = \left[\frac{D_t^*\left(x_t, y_t, a_t\right)}{D_{t+1}^*\left(x_{t+1}, y_{t+1}, a_{t+1}\right)}\right] \quad \text{and} \tag{2.9}$$

$$\text{TC}_{t,t+1}^* = \left[\frac{D_{t+1}^*\left(x_{t+1}, y_{t+1}, a_{t+1}\right)}{D_t^*\left(x_t, y_t, a_t\right)} \frac{D_{t+1}^*\left(x_t, y_t, a_t\right)}{D_t^*\left(x_{t+1}, y_{t+1}, a_{t+1}\right)}\right]^{\frac{1}{2}}. \tag{2.10}$$

When equations (2.5) and (2.6) are substituted into equation (2.7), the MMPI then has the following decomposition:

$$\text{MMPI}_{t,t+1} = \text{TEC}_{t,t+1}^k \times \text{TC}_{t,t+1}^k \times \text{TGRC}_{t,t+1}^k \tag{2.11}$$

where

$$\text{TGRC}_{t,t+1}^k = \left[\frac{\text{TGR}_t^k\left(x_{t+1}, y_{t+1}, a_{t+1}\right)}{\text{TGR}_{t+1}^k\left(x_t, y_t, a_t\right)} \frac{\text{TGR}_t^k\left(x_{t+1}, y_{t+1}, a_{t+1}\right)}{\text{TGR}_{t+1}^k\left(x_t, y_t, a_t\right)}\right]^{\frac{1}{2}}. \tag{2.12}$$

The technology gap ratio (TGR) refers to the ratio of the distance function based on the metafrontier to that based on the group frontier (*i.e.*, $D_t^*(\cdot)/D_t^k(\cdot)$) (O'Donnell *et al.* [24]). The TGR can also be applied to other period distance functions to measure how close the group k frontier is to the metafrontier. Specifically, increases in the TGR imply decreases in the gap between the group frontier and the metafrontier. The geometric mean of the TGR on the right-hand side of equation (2.12) can be defined as the TGR change (TGRC), which signifies a change in the TGR from period t to $t + 1$ and also denotes the technological catch-up [9, 20, 21].

TABLE 1. Sample statistics for the entire sample hospitals.

Variables name	Mean	Std. dev.	Minimum	Maximum
Panel A: inputs				
Number of physicians (x_1)	199	248	7	1366
Number of registered nurses (x_2)	627	633	43	3691
Number of other registered medical personnel (x_3)	167	152	14	839
Number of patient beds (x_4)	747	636	47	3807
Panel B: outputs				
Number of outpatient visits and emergency room visits (y_1)	693 208	600 660	44 670	3 359 220
Number of surgeries (y_2)	12 894	14 664	112	89 538
Total inpatient days (y_3)	180 510	176 943	2587	1 100 000
Panel C: quality attributes				
Number of net inpatient mortality (a_1)	373	390	1	2286
Number of referrals (a_2)	911	901	22	7472
Number of nosocomial infection (a_3)	451	620	2	4529

The quality index based on the metafrontier can be defined as

$$\text{QC}^*_{t,t+1} = \left[\frac{D^*_t(x_t, y_t, a_t)}{D^*_t(x_t, y_t, a_{t+1})} \frac{D^*_{t+1}(x_{t+1}, y_{t+1}, a_t)}{D^*_{t+1}(x_{t+1}, y_{t+1}, a_{t+1})} \right]^{\frac{1}{2}}. \tag{2.13}$$

In accordance with the definition of the TGR and by inserting equation (2.3) into equation (2.13), we have

$$\text{QC}^*_{t,t+1} = \text{QC}^k_{t,t+1} \times \text{QGRC}^k_{t,t+1} \tag{2.14}$$

where

$$\text{QGRC}^k_{t,t+1} = \left(\frac{\text{TGR}^k_t(x_t, y_t, a_{t+1})}{\text{TGR}^k_t(x_t, y_t, a_t)} \frac{\text{TGR}^k_{t+1}(x_{t+1}, y_{t+1}, a_{t+1})}{\text{TGR}^k_{t+1}(x_{t+1}, y_{t+1}, a_t)} \right)^{\frac{1}{2}}. \tag{2.15}$$

$\text{QGRC}^k_{t,t+1} > 1\,(<1)$ indicates a reduction (enlargement) in the gap when the quality attributes are evaluated between period t and $t+1$. That is, $\text{QGRC}^k_{t,t+1} > 1\,(<1)$ reflects improvement (deterioration) in the technology gap due to change in quality.

3. DATA

The sample hospital data used in this paper were obtained between 2008 and 2014 from the Collaboration Center of Health Information Application under Taiwan's Ministry of Health and Welfare. The sample was limited to general hospitals to meet the homogeneity requirement of the DEA. After hospitals with either missing or extreme values were eliminated, 106 general hospitals with data relevant to our 7-year study period remained. Of these hospitals, 87 were teaching hospitals and 19 were non-teaching hospitals. In accordance with the literature [7, 8, 13, 14, 18], we identified four inputs: the number of physicians and full-time equivalents (x_1), the number of registered nurses (x_2), the number of other registered medical personnel (x_3), and the number of patient beds (x_4). We then identified three outputs: the number of outpatient visits and emergency room visits (y_1), the number of surgeries including outpatient and inpatient surgeries (y_2), and total inpatient days (y_3). Finally, we identified three quality attributes: net inpatient mortality (a_1), the number of referrals in and referrals out (a_2), and the number of nosocomial infections (a_3).

Table 1 provides descriptive statistics for all of the aforementioned variables. Note that all of the variables have large standard deviations, because the sample contained small- to large-scale hospitals. Table 2 presents

TABLE 2. Sample statistics for teaching and non-teaching hospitals.

Variable name	Teaching hospitals ($N = 87$)		Non-teaching hospitals ($N = 19$)	
	Mean	Std. dev.	Mean	Std. dev.
Panel A: inputs				
Number of physicians (x_1)	236	259	29	12
Number of registered nurses (x_2)	739	646	114	50
Number of other registered medical personnel (x_3)	194	155	42	17
Number of patient beds (x_4)	854	649	258	201
Panel B: outputs				
Number of outpatient and emergency room visits (y_1)	804 100	608 058	185 436	76 252
Number of surgeries (y_2)	15 277	15 163	1981	1396
Total inpatient days (y_3)	207 297	183 072	57 854	53 989
Panel C: quality attributes				
Number of net inpatient mortality (a_1)	433	405	99	75
Number of referrals (a_2)	959	957	690	528
Number of nosocomial infection (a_3)	530	657	87	82

statistics by hospital type. Teaching hospitals provide a public good in terms of education and research, and their direct patient care is often more sophisticated and specialized than that available in non-teaching hospitals. Moreover, teaching hospitals often operate on a larger scale than non-teaching hospitals, possibly resulting in teaching hospitals' outperformance of non-teaching hospitals in patient care. As expected, the teaching hospitals represented in this study had larger input, output, and quality quantities than the non-teaching hospitals.

4. RESULTS

Panels A and B of Table 3 show the respective empirical results of productivity analysis of teaching and non-teaching hospitals. Table 3 shows that the average annual productivity growth rates of teaching and non-teaching hospitals were 0.02% and 1.49%, respectively, and quality increased on average by 1.74% and 5.83% per year, respectively. The results show that both the teaching and non-teaching hospitals increased in productivity and improved in quality over the study period of 2008–2014. The non-teaching hospitals showed much greater positive developments in productivity and quality than the teaching hospitals.

We then examined the components of the MMPI and identified the main source for productivity growth. Table 3 shows that the TEC for the teaching hospitals remained stagnant, while the TEC for the non-teaching hospitals exhibited an average annual reduction rate of 0.23% between 2008 and 2014. As for the TGRC, teaching hospitals demonstrated weak catch-up dynamics. The convergence rate toward the potential technology over the whole period was a mere 0.5%. The technology adopted by the non-teaching hospitals diverged from the potential technology, resulting in an annual change rate of -0.14%. The TC for the teaching and non-teaching hospitals had average annual growth rates of 0.39% and 1.96%, respectively. After decomposition of the MMPI, overall productivity growth in the teaching and non-teaching hospitals was found to stem mainly from technological progress. Moreover, the TC for the non-teaching hospitals was much higher than that for the teaching hospitals. Consequently, the MMPI with quality attributes for the non-teaching hospitals was 1.49%, compared with the 0.02% change rate of the teaching hospitals.

Table 3 further illustrates the dynamics of the QC and its decomposition for the teaching and non-teaching hospitals. Over the study period, the growth rates of quality for the teaching and non-teaching hospitals were positive. However, the QC for the non-teaching hospitals was 5.83%, which is higher than the 1.74% change rate

TABLE 3. The mean of the MMPI and its decompositions for teaching and non-teaching hospitals.

Time period	MMPI	TEC	TC	TGRC	QC	QGRC
	Panel A: teaching hospital					
2008–2009	0.9839	0.9784	1.0137	0.9962	1.0100	0.9886
2009–2010	1.0082	1.0061	0.9983	1.0069	1.0194	0.9906
2010–2011	0.9932	0.9953	1.0031	0.9966	1.0114	0.9891
2011–2012	0.9990	1.0190	0.9928	0.9937	1.0253	0.9842
2012–2013	1.0010	1.0078	0.9938	1.0046	1.0172	0.9892
2013–2014	1.0159	0.9941	1.0222	1.0053	1.0213	0.9924
2008–2014 [a]	*1.0002*	*1.0000*	*1.0039*	*1.0005*	*1.0174*	*0.9890*
	Panel B: non-teaching hospital					
2008–2009	1.0757	0.9716	1.1203	0.9883	1.1488	0.9839
2009–2010	1.0051	1.0165	0.9854	1.0040	1.0431	0.9926
2010–2011	1.0073	0.9600	1.0463	1.0067	1.0366	0.9946
2011–2012	0.9830	0.9846	1.0065	0.9918	1.0456	0.9892
2012–2013	0.9970	1.0621	0.9449	0.9947	1.0426	0.9912
2013–2014	1.0241	0.9947	1.0231	1.0062	1.0376	0.9988
2008–2014 [a]	*1.0149*	*0.9977*	*1.0196*	*0.9986*	*1.0583*	*0.9917*

Notes. [a] Numbers are geometric means.

for the teaching hospitals. As for the QGRC, the teaching and non-teaching hospitals dropped by an annual average of 1.1% and 0.83%, respectively. These results reflect that, during the study period, the quality attribute gap grew, causing a reduction in the technology gap.

Teaching hospital programs aim to cultivate outstanding medical personal by facilitating education and research. Therefore, in general, these hospitals maintain high-quality hardware and software; high levels of professionalism among doctors, nurses, and other personnel; and strong management. However, among the hospitals in this study, the rates of technological change and QC for the non-teaching hospitals were not inferior to those of the teaching hospitals. Furthermore, the non-teaching hospitals showed greater productivity growth than the teaching hospitals. Previous studies have focused on the performance of medical services, leading researchers to conclude that the medical services of non-teaching hospitals are not inferior to those of teaching hospitals; for example, Grosskopf et al. [16] stated that only approximately 10% of teaching hospitals could effectively "compete" with non-teaching hospitals. In teaching hospitals, education and research programs increase institutional costs [5,26], divert medical and nursing personnel's energy and attention, adversely affect the output of medical services and activities, and result in the inefficient provision of patient services. As Grosskopf et al. [17] showed, 20% of the inefficiency in teaching hospitals was due to congestion of residents. However, in terms of performance of medical services alone, the gap between teaching hospitals and non-teaching hospitals is not significant. Furthermore, if the teaching and research achievements of teaching hospitals are considered, the overall operating performance of teaching hospitals should be regarded as superior to that of non-teaching hospitals.

In this study, we examined the productivity changes of teaching and non-teaching hospitals under the assumption that these two types of hospitals use different production technologies. As previously stated, the performance evaluation of a hospital operating with one type of technology should not be compared with that of another hospital operating with a different type of technology. Therefore, the use of a metafrontier technique is recommended for such analyses [1,2]. In our application of this method, we calculated the TGR of various groups to represent the technical differences between the group boundary and common boundary. Table 4 shows that the average TGRs of teaching and non-teaching hospitals were 0.9164 and 0.9555, respectively. These results indicate that the non-teaching hospital data are closer to the metafrontier than the teaching hospital data. We examined the difference in TGRs between these two groups using a Mann–Whitney nonparametric test.

TABLE 4. Mean difference test for the TGR.

Time period	Teaching hospitals [a]	Nonteaching hospitals [a]	p-value [b]
2008	0.9221	0.9764	0.0012
2009	0.9026	0.9459	0.0525
2010	0.9196	0.9441	0.2407
2011	0.9052	0.9485	0.0323
2012	0.9095	0.9694	0.0011
2013	0.9257	0.9497	0.1488
2014	0.9304	0.9546	0.1608
2008–2014 [a]	*0.9164*	*0.9555*	*0.0000*

Notes. [a] Numbers are arithmetic means. [b] The p-value is for testing whether TGR exists significantly difference between two groups by Mann–Whitney test.

The mean TGR in the teaching and non-teaching hospitals significantly differed from 0 over the study period. Therefore, technical differences between the teaching and non-teaching hospitals justified the application of the metafrontier approach in this study.

As to the measure of TGR in each group, the arithmetic average has been used for this work. Using the arithmetic average implies that the same weights are assigned to all DMUs in the group. Walheer [30] suggested to find an aggregation scheme, which takes the relative importance of the DMUs in the group into account. Constrained by data collection from the National Health Insurance Research Database (NHIRD), we do not have price information for the service provided. Therefore, we retain this new way to obtain the group indicator for TGR as a future research direction.

5. CONCLUSIONS

The main finding of this study is that teaching hospitals in Taiwan cannot straightforwardly be associated with better results in terms of productivity and quality than non-teaching hospitals do. These findings appear to raise doubts that teaching and research tasks may affect teaching hospitals' performance in health care. Therefore, policymakers and managers are suggested to define specific policies and actions in order to take care of both teaching tasks and patient care in teaching hospitals, where teaching hospitals play an important role as training clinicians for the future.

REFERENCES

[1] G.E. Battese and D.S.P. Rao, Technology gap, efficiency, and a stochastic metafrontier function. *Int. J. Bus. Econ.* **1** (2002) 87–93.

[2] G.E. Battese, D.S.P. Rao and C.J. O'Donnell, A metafrontier production function for estimation of technical efficiencies and technology gaps for firms operating under different technologies. *J. Prod. Anal.* **21** (2004) 91–103.

[3] L.G. Burke, A.B. Frakt, D. Khullar, E.J. Orav and A.K. Jha, Association between teaching status and mortality in U.S. hospitals. *JAMA* **317** (2017) 2105–2113.

[4] L.G. Burke, D. Khullar, J. Zheng, A.B. Frakt, E.J. Orav and A.K. Jha, Comparison of costs of care for medicare patients hospitalized in teaching and nonteaching hospitals. *JAMA Netw Open.* **2** (2019) e195229.

[5] J.M. Cameron, The indirect costs of graduate medical education. *New England J. Med.* **312** (1985) 1233–1238.

[6] C. Campbell, K. Gillespie and J. Romeis, The effects of residency training programs on the financial performance of veterans affairs medical centers. *Inquiry* **28** (1991) 288–299.

[7] K. Carey and J.F. Burgess, Jr., On measuring the hospital cost-quality trade-off. *Health Econ.* **8** (1999) 509–520.

[8] S.J. Chang, H.C. Hsiao, L.H. Huang and H. Chang, Taiwan quality indicator project and hospital productivity growth. *Omega* **39** (2011) 14–22.

[9] K.H. Chen and H.Y. Yang, A cross-country comparison of productivity growth using the generalised metafrontier Malmquist productivity index: with application to banking industries in Taiwan and China. *J. Prod. Anal.* **35** (2011) 197–212.

[10] K.C. Chen, L.N. Chien, Y.H. Hsu and M.M. Yu, Metafrontier frameworks for studying hospital productivity growth and quality changes. *Int. J. Qual. Health Care* **28** (2016) 650–656.

[11] R. Färe, S. Grosskopf, M. Norris and Z. Zhang, Productivity growth, technical progress, and efficiency change in industrialized countries. *Am. Econ. Rev.* **84** (1994) 66–83.

[12] R. Färe, S. Grosskoph and P. Roos, Productivity and quality changes in Swedish pharmacies. *Int. J. Prod. Econ.* **39** (1995) 137–144.

[13] M. Farsi, Changes in hospital quality after conversion in ownership status. *Int. J. Health Care Finance Econ.* **4** (2004) 211–230.

[14] G.D. Ferrier and J.S. Trivitt, Incorporating quality into the measurement of hospital efficiency: a double DEA approach. *J. Prod. Anal.* **40** (2013) 337–355.

[15] A. Flokou, V. Aletras and D. Niakas, A window-DEA based efficiency evaluation of the public hospital sector in Greece during the 5-year economic crisis. *PLoS One* **12** (2017) e0177946.

[16] S. Grosskopf, D. Margaritisb and V. Valdmanis, The effects of teaching on hospital productivity. *Soc.-Econ. Planning Sci.* **35** (2001) 189–204.

[17] S. Grosskopf, D. Margaritisb and V. Valdmanis, Comparing teaching and non-teaching hospitals: a frontier approach (teaching vs. non-teaching hospitals). *Health Care Manage. Sci.* **4** (2001) 83–90.

[18] A.J. Hartz, M.S. Gottlieb, E.M. Kuhn and A.A. Rimm, The relationship between adjusted hospital mortality and the results of peer review. *Health Serv. Res.* **27** (1993) 765–777.

[19] G. Jensen and M. Morrisey, The role of physicians in hospital production. *Rev. Econ. Stat.* **68** (1986) 432–441.

[20] S. Kumar and R.R. Russell, Technological change, technological catch-up, and capital deepening: relative contributions to growth and convergence. *Am. Econ. Rev.* **92** (2002) 527–548.

[21] S.C. Kumbhakar and H.J. Wang, Estimation of growth convergence using a stochastic production frontier approach. *Econ. Lett.* **88** (2005) 300–305.

[22] P. Nayar, Y.A. Ozcan, F. Yu, A.T. Nguyen, Benchmarking urban acute care hospitals. *Health Care Manage. Rev.* **38** (2013) 137–145.

[23] S. Nuti, T. Grillo Ruggieri and S. Podetti, Do university hospitals perform better than general hospitals? A comparative analysis among Italian regions. *BMJ Open.* **6** (2016) e011426.

[24] C. J. O'Donnell, D.S.P. Rao and G.E. Battese, Metafrontier frameworks for the study of firm-level efficiencies and technology ratios. *Empir. Econ.* **34** (2008) 231–255.

[25] H. Pardes, The future of medical schools and teaching hospitals in the era of managed care. *Acad. Med.* **72** (1997) 97–102.

[26] E. Rich, G. Gifford, M. Luxenberg and B. Dowd, The relationship of house staff experience to the cost and quality of inpatient care. *J. Am. Med. Assoc.* **263** (1990) 953–957.

[27] D.M. Shahian, P. Nordberg, G.S. Meyer, B.B. Blanchfield, E.A. Mort, D.F. Torchiana and S.L. Normand, Contemporary performance of U.S. teaching and nonteaching hospitals. *Acad Med.* **87** (2012) 701–708.

[28] F. Sloan, R. Feldman and B. Steinwald, Effects of teaching on hospital productivity. *J. Health Econ.* **2** (1983) 1–23.

[29] V. Valencia, V.M. Arora, S.R. Ranji, C. Meza and C. Moriates, A comparison of laboratory testing in teaching vs nonteaching hospitals for 2 common medical conditions. *JAMA Intern Med.* **178** (2018) 39–47.

[30] B. Walheer, Aggregation of metafrontier technology gap ratios: the case of European sectors in 1995–2015. *Eur. J. Oper. Res.* **269** (2018) 1013–1026.

[31] J. Yang and W. Zeng, The trade-offs between efficiency and quality in the hospital production: some evidence from Shenzhen, China. *Chin. Econ. Rev.* **31** (2014) 166–184.