

# Measuring Hospital Efficiency with Data Envelopment Analysis: Nonsubstitutable vs. Substitutable Inputs and Outputs

Darold T. Barnum · Surrey M. Walton ·  
Karen L. Shields · Glen T. Schumock

Received: 3 September 2009 / Accepted: 2 December 2009 / Published online: 15 December 2009  
© Springer Science+Business Media, LLC 2009

**Abstract** There is a conflict between Data Envelopment Analysis (DEA) theory's requirement that inputs (outputs) be substitutable, and the ubiquitous use of *nonsubstitutable* inputs and outputs in DEA applications to hospitals. This paper develops efficiency indicators valid for nonsubstitut-

able variables. Then, using a sample of 87 community hospitals, it compares the new measures' efficiency estimates with those of conventional DEA measures. DEA substantially overestimated the hospitals' efficiency on the average, and reported many inefficient hospitals to be efficient. Further, it greatly overestimated the efficiency of some hospitals but only slightly overestimated the efficiency of others, thus making any comparisons among hospitals questionable. These results suggest that conventional DEA models should not be used to estimate the efficiency of hospitals unless there is empirical evidence that the inputs (outputs) are substitutable. If inputs (outputs) are not substitutes, efficiency indicators valid for nonsubstitutability should be employed, or, before applying DEA, the non-substitutable variables should be combined using an appropriate weighting scheme or statistical methodology.

---

D. T. Barnum (✉)  
Department of Managerial Studies,  
University of Illinois at Chicago,  
601 South Morgan Street,  
Chicago, IL 60607-7123, USA  
e-mail: dbarnum@uic.edu

D. T. Barnum  
Department of Information & Decision Sciences,  
University of Illinois at Chicago,  
Chicago, IL, USA

D. T. Barnum  
Department of Pharmacy Administration,  
University of Illinois at Chicago,  
Chicago, IL, USA

S. M. Walton  
Center for Pharmacoeconomic Research,  
Department of Pharmacy Administration,  
University of Illinois at Chicago,  
Chicago, IL, USA

S. M. Walton  
Department of Economics, University of Illinois at Chicago,  
Chicago, IL, USA

K. L. Shields  
Sisters of St. Francis Health Services, Inc.,  
Mishawaka, IN, USA

G. T. Schumock  
Center for Pharmacoeconomic Research,  
Department of Pharmacy Practice,  
University of Illinois at Chicago,  
Chicago, IL, USA

**Keywords** Data envelopment analysis · Efficiency · Hospitals · Fixed proportion technology

## Introduction

The importance of healthcare efficiency is extremely high, given the rapid growth in healthcare costs and the increasing numbers of people covered by publicly-financed programs. To identify useful healthcare productivity improvements, efficiency must be validly measured. On the other hand, if healthcare efficiency is incorrectly measured, then governmental policy makers and hospital managers may respond in ineffective and even counterproductive ways.

In his latest review of healthcare efficiency studies, Hollingsworth [1] reports that there has been a rapid growth in the number of publications using Data Envelopment Analysis (DEA), and that over half of all healthcare DEA

publications involve hospitals. The growing list of publications using DEA to measure hospital efficiency is mirrored in the *Journal of Medical Systems*, which has published nine hospital DEA articles since 2000, with four of the nine coming in the last 2 years [2–10].

If hospital DEA studies are to inform effective practice, we need to assure policy makers and hospital managers that DEA is being correctly applied [11]. Studies confirming the validity of DEA applications to hospitals would raise the confidence of both academic scientists and real-world practitioners in the analytical results. For any incorrect aspects of hospital applications that are discovered, DEA models could be adapted to deal with the problems, or DEA could be replaced with more appropriate efficiency indicators.

This paper considers a heretofore overlooked problem in DEA's application to hospitals, which nevertheless has important consequences for the validity of DEA estimates. The issue at hand is the conflict between DEA theory's requirement that inputs and outputs be substitutable, and the ubiquitous use of *non*substitutable inputs and outputs in DEA hospital applications.

#### Input and output substitutability: Definitions and DEA theory

When inputs are nonsubstitutable, then they cannot replace each other in the production of a constant amount of output. Such inputs must be utilized in a fixed proportion to produce their output, and any quantity of an input in excess of the required ratio is wasted. Production systems using nonsubstitutable inputs are well-known in economics, and are called "Fixed Factor Proportion Technologies" [12]. If outputs are nonsubstitutable, then, for a fixed amount of input, production of one output cannot be increased by producing less of another. Such production systems are "Fixed Product Proportion Technologies" [12].

When inputs are substitutable for each other in the production of output, a fixed amount of output can be produced with varying combinations of the inputs. When outputs are substitutable for each other, the amount of one output can be increased and the amount of another decreased for a fixed amount of input by changing the proportion of the input that each output receives.

DEA mathematical models and the economic theory underlying them require substitutability. Farrell [13] and Charnes, Cooper and Rhodes [14] assume substitutability, as does Banker, Charnes and Cooper [15], Färe, Grosskopf and Lovell [16] and more recent work [17, 18].

#### Input and output substitutability in DEA applications

The issue of substitutability in DEA application papers has rarely been addressed. Although none of them involved

healthcare, we do know of two recent articles that purposely selected inputs and outputs that avoided non-substitutability [19, 20], and another article that used Multi-Directional Efficiency Analysis instead of DEA partly because of the substitutability issue [21].

In DEA applications to hospitals, Hollingsworth [1, p. 1110] reports that inputs "are mainly measures of staff and capital employed," and most of the studies use output measures "such as inpatient days or discharges." Recent hospital DEA articles in the *Journal of Medical Systems* are consistent with the widespread use of such inputs and outputs. The nine hospital DEA papers published in the journal since 2000 all included staffed beds as a proxy for capital, and the number of employees (in various categories) as a proxy for labor. Also, all of the nine papers included outputs separately measuring the numbers of inpatients and outpatients [2–10].

As discussed later, these labor and capital input proxies cannot be substituted for each other in the production of a fixed amount of output. And, although it would seem that outputs such as inpatients and outpatients would always be substitutes, they were not substituted for each other in our sample hospitals. In short, the hospitals that we studied employed fixed proportion technologies.

#### DEA theory vs. DEA applications to hospitals

In sum, DEA applications measuring hospital efficiency have employed inputs and outputs that conflict with DEA theory. This inconsistency could result in trivial effects on DEA efficiency scores, or could cause substantial and significant errors in DEA efficiency estimates.

The purpose of this paper is to examine the impact of DEA theory's substitutability requirements on its applications to hospitals. To analyze the effects of the conflict between DEA theory and hospital applications, we developed efficiency indicators that assume nonsubstitutability rather than substitutability. Next, using hospital-wide data supplied by the pharmacy departments of US community hospitals, we ascertained empirically that their inputs and outputs were indeed nonsubstitutable, and then we compared their DEA scores with the scores from our new efficiency measures.

For our sample hospitals, DEA resulted in severely biased and imprecise estimates of efficiency. All hospitals were less efficient in truth than estimated by DEA, and DEA reported many inefficient hospitals to be efficient. Further, the efficiency scores of some hospitals were only slightly affected while the efficiency scores of others showed large biases, thereby making comparisons among hospitals unattainable. Of course, we don't know if other DEA hospital efficiency studies suffer to the same extent, but we do suggest that future studies restrict DEA inputs and outputs to substitutable variables or use efficiency indicators not requiring substitutability.

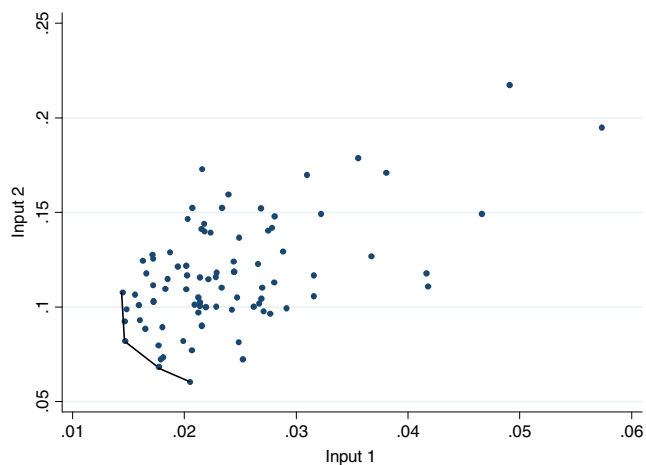
**Methods**

Location of the production frontier when inputs are substitutable and nonsubstitutable

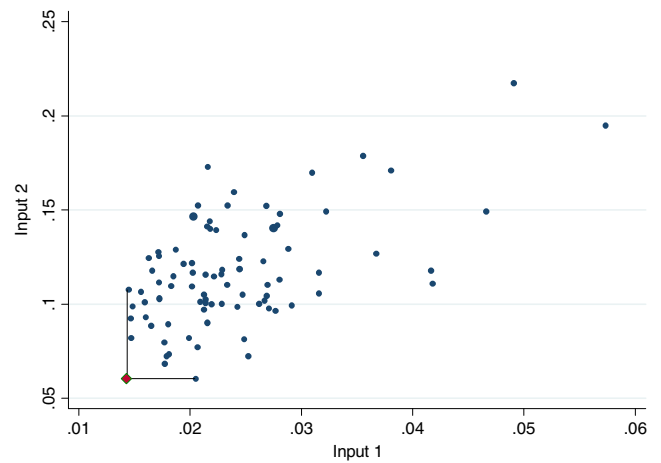
In DEA, the organizations being analyzed often are called Decision Making Units (DMUs). In this paper, the organizations being analyzed are individual hospitals, so each hospital is a DMU. In order to be consistent with the DEA literature, we often use the term DMUs to refer to the hospitals.

Consider a hypothetical case of one output and two inputs (a) when the inputs are substitutable, and (b) when the inputs are not substitutable. Suppose one unit of output is produced by each organization being analyzed with various combinations of the two inputs (Fig. 1). If the two inputs used to produce one unit of output are substitutable inputs, a representation of the production frontier is shown by the inner-most piecewise isoquant. That is, if less of one input is used, some amount more of the other input must be used in its place to hold output constant. Conventional DEA models would report that the four DMUs defining the isoquant are efficient, because, though having different mixes of the inputs, they are all on the same inner-most isoquant.

If the two inputs are *not* substitutable, then an efficient DMU must use them in a fixed proportion. Suppose that the inputs are truly nonsubstitutable and one unit of output is produced, as is shown in Fig. 2 (using the same data as Fig. 1). The production frontier now consists of a single point that is, a *point frontier*. This frontier is estimated by the composite DMU in the south-west corner of the graph. The frontiers of the reference set increase vertically and horizontally from this point, forming a right-angle or L-shaped reference set frontier. However, the *only* Pareto–Koopmans efficient subset of the reference set frontier is the point frontier, because only at that point is the requisite output achieved without the overuse of one of the inputs [22].



**Fig. 1** Substitutable inputs best-practice (isoquant) frontier



**Fig. 2** Nonsubstitutable inputs best-practice (point) frontier

For substitutable inputs, the minimum level of each input is conditioned on the level of the other input. However, for *non*substitutable inputs, that is, fixed factor proportion technologies, the minimum level of each input needed to produce a given amount of output is not influenced by the other input [12]. So, if all DMUs’ outputs are equal, when inputs are nonsubstitutable it is only necessary to find the minimum level of each input. As is true for conventional DEA, this frontier estimation method envelops the data, and assumes that a composite DMU can be used to identify a point on the efficient frontier that is attainable by an actual DMU. As also is true for DEA, this deterministic measure estimates efficiency based on the most extreme observations, ignoring any stochastic variation that might exist.

Note that all DMUs’ reported efficiencies will be quite different when the point frontier is used in place of the isoquant frontier as the efficient reference. We return to the issue of efficiency indices for the point frontier after discussing methods for identifying whether or not inputs (outputs) are substitutable.

Method for identifying the presence or absence of substitutability

Because we know that the DMUs shown in Figs. 1 and 2 each produced one unit of output, it might appear that the empirical evidence suggests that these two inputs are substitutable for each other, as shown by the isoquant in Fig. 1. But, substitutability is not necessarily present because a DMU is unlikely to be equally efficient in its use of both inputs [23]. In Fig. 1, for example, if inputs are truly nonsubstitutable, the supposed piecewise isoquant frontier may be the result of one DMU being the most efficient of all DMUs in the use of the first input but less efficient in the use of the second, and another DMU being the most efficient of all DMUs in the use of the second input but less efficient in the use of the first. Substitutability, or the lack

thereof, can be identified by logic and statistical testing, but *cannot* be identified by a deterministic estimation of an alleged best-practice frontier.

One simple method for assessing whether or not inputs are substitutable is to regress each input on the remaining inputs and all outputs. If any two inputs are substitutes, then the relationship between them must be negative (with statistical significance). Because the remaining inputs and all outputs are held constant, a decrease in any one input would have to be compensated for by an increase in the other input. If the two inputs are not substitutes, then there will be no statistically significant relationship between them if their inefficiencies are uncorrelated. There will be a statistically significant positive relationship if there is a high degree of correlation between the two inputs' inefficiency levels.

Empirically estimating whether outputs are substitutable or nonsubstitutable follows the same methodology used for inputs. Each output is regressed on the remaining outputs and all inputs. If any two outputs are substitutes, then the relationship between them must be negative to a statistically significant degree, because an increase in any one output would have to be compensated for by a decrease in another output. And, if two outputs are not substitutes, then, as with inputs, there will be no statistically significant relationship between them if their degree of inefficiencies are uncorrelated, and there will be a statistically significant positive relationship if there is a high degree of correlation between their inefficiency levels.

In truth, the inputs in Figs. 1 and 2 are not substitutable for each other. One of the inputs is staffed beds and the other is number of staff used for one unit of inpatient output. Logically, it is not possible to serve a fixed number of inpatients by decreasing one of these inputs and making up for the decrease by increasing the other. Statistically, because there is one unit of output and two inputs, we can regress one of the inputs on the other to determine whether there is a statistically significant negative relationship (indicating the inputs are substitutable) or not. In fact, the two inputs show a positive relationship, confirming that they are nonsubstitutable.

The next task is to develop efficiency measures for cases of nonsubstitutable inputs. Then, we can compare the new efficiency measures with DEA efficiency estimates.

#### Additive efficiency measure when inputs are nonsubstitutable

In order to estimate each DMU's efficiency relative to the point frontier in Fig. 2, one simple possibility would be a variation on the DEA Additive (ADD) model [22]. With DEA's ADD model, the efficient point for an assessed DMU is the furthest point on the piecewise isoquant frontier where neither of its inputs has increased and its output has not

decreased. The rectilinear distance between the assessed DMU and that point measures the DMU's inefficiency.

We call the variation of the DEA Additive model the "Fixed Proportion Additive" (FPA) model, because it assumes that the inputs and outputs occur in fixed proportions. Like the ADD model, the degree of inefficiency is measured by the rectilinear distance between the target DMU and the efficient point. But, for the FPA model, the efficient point is the *point frontier* rather than a point on a piecewise isoquant frontier.

Significantly, the only difference between the two models is the location of the point from which inefficiency is measured. This can be seen in Fig. 3. We can estimate the point frontier for one unit of output when inputs are nonsubstitutable from the DMUs using the minimum amounts of each input, that is, from the DMUs establishing the boundaries of the right-angle reference set frontier. DMU A uses the least of input capital (1.16) and DMU E uses the least of input labor 2 (0.13), so a fully efficient composite DMU would use 1.16 units of capital and 0.13 units of labor, as shown by point F on the graph. Of course, if a particular DMU were the most efficient in the use of both inputs, then that one DMU alone would determine the point frontier. For example, if point F represented an actual DMU instead of a composite DMU, then that actual DMU would reflect the point of maximum efficiency.

The FPA efficiency score for the assessed DMU<sub>k</sub> can be obtained for each DMU<sub>k</sub> from a set of  $j=1,2,\dots,J$  DMUs with one output  $y_{j1}$  and  $M$  inputs  $x_{jm}$  for  $m=1,2,\dots,M$  by the use of Eq. 1.

$$FPA_k = \sum_{m=1}^M \left[ \left| (x_{km}/y_{k1}) - \text{Min}_j (x_{jm}/y_{j1}) \right| \right] \quad (1)$$

So that the scores of the ADD and the FPA models will be directly comparable, we divide each DMU's inputs by its output. Therefore, the input and output values used in the FPA model (Eq. 1) and the ADD model (Eqs. 2–5) are identical, so the resulting sums of the slacks for the target DMU  $k$  are directly comparable.

$$ADD_k = \max \sum_{m=1}^M s_m^- \quad (2)$$

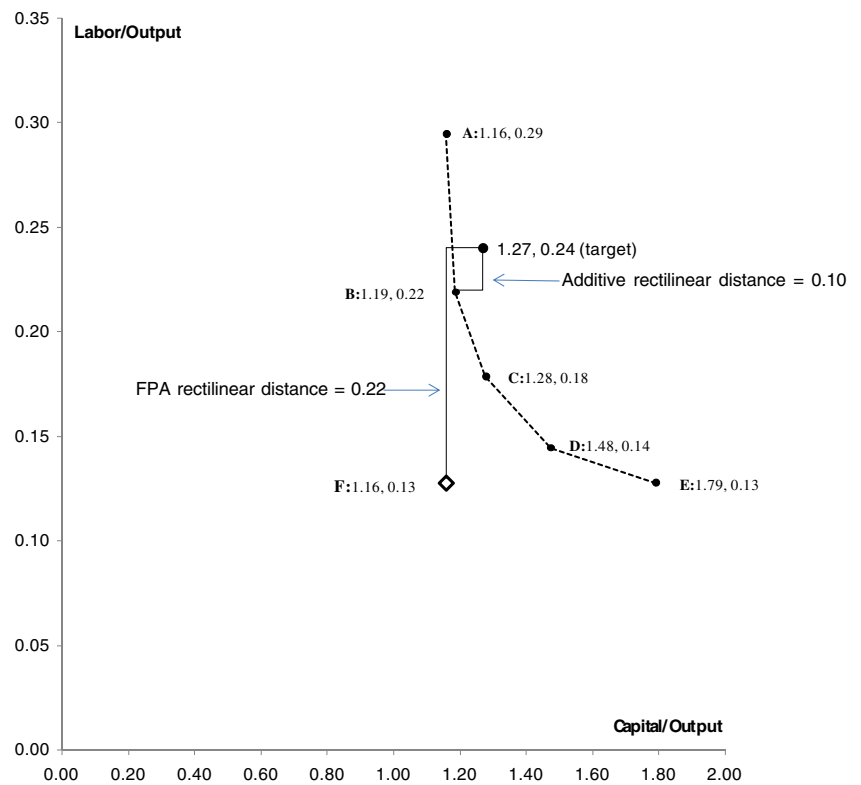
Subject to

$$\sum_{j=1}^J \lambda_j (x_{jm}/y_{j1}) + s_m^- = x_{km}/y_{k1} \quad m = 1, 2, \dots, M \quad (3)$$

$$\hat{y}_{j1} = 1 = y_{j1}/y_{j1} \quad j = 1, 2, \dots, J \quad (4)$$

$$\lambda_j \geq 0 \quad (5)$$

**Fig. 3** Graphical representation of ADD and FPA measures



Ratio efficiency index when inputs are nonsubstitutable

The primary value of the two preceding additive models is that because they use the same metric, their inefficiency scores are directly comparable. However, because both yield absolute measures of inefficiency rather than indices, their inefficiency values have no intuitive meaning and they are not units-invariant [24]. A more useful measure would be an index in [0, 1], because it would identify the proportional efficiency of target DMUs, as do conventional DEA radial measures such as the Charnes–Cooper–Rhodes (CCR) model. In this section, we develop such an index, the *Fixed Proportion Ratio* (FPR) measure, to deal with nonsubstitutability.

In order to measure a DMU’s relative degree of inefficiency in the use of an input to produce an output, the indicator needs to be normalized by some base. Thus, for *each* output/input combination, we compute the normalized output/input ratio by dividing the target DMU’s output/input ratio by that of the DMU *j* that is the most efficient for that particular output/input ratio. Equation 6 illustrates the efficiency of DMU *k*’s input *m* and output *n*. The input and output in the numerator are from the assessed DMU *k*, and the input and output in the denominator are from the DMU that has the maximum output/input ratio for that specific output/input combination. The range of efficiency scores for each output/input

pair is [0, 1], and at least one DMU will achieve an efficiency score of 1.

$$eff_{kmn} = \left[ \frac{y_{kn}/x_{km}}{\text{Max}_j (y_{jn}/x_{jm})} \right] \tag{6}$$

Because a DMU’s efficiency would usually be different for each output/input combination, its average efficiency can be computed as the mean of its individual efficiencies. Thus, for each of DMU *k*’s output/input ratios, Eq. 7 is used to compute the normalized efficiency measure for that ratio. If there are *m* inputs and *n* outputs, then there will be *m* × *n* efficiency measures of the form *eff<sub>kmn</sub>*. So, for each DMU *k* in a set of *J* DMUs, the mean of its *m* × *n* efficiency measures is computed, which yields a partially normalized efficiency measure for that DMU. Then, each DMU *k*’s partially normalized efficiency measure is divided by the maximum partially normalized efficiency measure, which yields a normalized efficiency measure in [0, 1]. This is the Fixed Proportion Ratio (FPR) index:

$$FPR_k = \frac{(1/MN) \sum_{m=1}^M \sum_{n=1}^N eff_{kmn}}{\text{Max}_j \left( (1/MN) \sum_{m=1}^M \sum_{n=1}^N eff_{jmn} \right)} \tag{7}$$



**Table 1** Summary statistics for 87 community hospitals

Input and output variables	Mean	Median	Smallest	Largest
FTE employees	1,231	892	85	4,973
Staffed beds	243	189	22	861
Annual inpatients	11,318	7,919	436	39,948
Annual outpatients	168,128	128,954	14,536	1,078,423

### Comparing DEA models with fixed proportion models

Now we have indicators for comparing the efficiency estimates of the two additive models, DEA's ADD model that assumes substitutable inputs and the new FPA model that assumes nonsubstitutable inputs. And, we can also compare the efficiency estimates of the two proportional indices, DEA's Charnes–Cooper–Rhodes (CCR) model [22, p. 94] that assumes substitutable inputs and outputs, and the FPR model that assumes nonsubstitutable inputs and outputs. All four of these models incorporate both technical efficiency and any scale effects that may exist. (For our hospital sample, the relationship between a weighted patient dependent variable and labor and capital independent variables was linear, so there were no scale effects involved in this case.)

### Materials: Sample, inputs and outputs

Our sample consisted of data from 87 community hospitals in the United States that were members of a national group purchasing organization. The data were collected for use in an earlier study of community hospital pharmacy productivity [25], from an online questionnaire that was completed by pharmacy directors at the hospitals. Herein, we used hospital-wide data from the hospitals that included all of the inputs and outputs that we needed for our computations.

For tests comparing the FPA and ADD additive models, the one output was total inpatients, and the two inputs were staffed beds and full-time-equivalent (FTE) employment. It was only possible to use one output in comparisons of these additive models, so we chose the output that had by far the strongest impact on the levels of inputs needed.

For tests comparing the FPR and the CCR models, multiple inputs and outputs are possible. We used the two

most common outputs, total inpatients and total outpatients, and, as before, the two inputs were staffed beds and FTE employment. Summary values for the 87 hospitals are shown in Table 1.

### Results

#### Input and output substitutability

One input was regressed on the other input, with the outputs included as control variables. As Table 2 shows, the number of employees was positively related to the number of beds with statistical significance of 0.059, which, based on our earlier logical argument, would be as hypothesized. More important, there was *not* a negative relationship, statistically significant or otherwise, and a negative relationship would be necessary if the factors could be substituted for each other. Using a different proxy for capital *might* result in a different conclusion, but we used the proxy that has been empirically validated and is common to most published hospital DEA articles [26]. Therefore, a fixed factor proportion technology was present.

One output was regressed on the other, with the inputs included as control variables. As Table 3 shows, the relationship between outpatients and inpatients was positive, a somewhat surprising finding.

However, looking again at Table 2, it can be seen that the number of outpatients had relatively minor influence on the number of employees. This apparent lack of influence may have resulted from the narrow range within which the outpatient and inpatient proportions occurred for our sample. Except for a very few hospitals, the proportion of outpatients clustered between 90% and 97%, out of a possible range from 0 to 100. So, this appears to be a case of our community hospital sample all having about the

**Table 2** Regression of hospital employees on beds, holding inpatients and outpatients constant

Total FTE employees	Coefficient	Robust Std. Err.	<i>t</i>	<i>P</i> > <i>t</i> (1-tail)
Staffed beds	0.8950035	0.5668977	1.58	0.059
Annual inpatients	0.0817770	0.0115093	7.11	0.000
Annual outpatients	0.0007314	0.0004787	1.53	0.065
Constant	-35.29337	62.81318	-0.56	0.576

*R*-square=0.9106. One-tail probabilities because we predicted outcomes

**Table 3** Regression of inpatients on outpatients, holding beds and hospital employees constant

Annual inpatients	Coefficient	Robust Std. Err.	<i>t</i>	<i>P</i> > <i>t</i> (2-tail)
Annual outpatients	0.0033999	0.0022096	1.54	0.128
Staffed beds	24.42506	5.547089	4.40	0.000
Total FTE employees	3.897101	1.088807	3.58	0.001
Constant	14.62555	345.3861	0.04	0.966

*R*-square=0.9447. Two-tail probabilities because we did not predict outcomes

same ratio of inpatients to outpatients, rather than a case of true technical nonsubstitutability. However, from the viewpoint of modeling choice, the reason for the empirical lack of substitutability does not matter and we have to honor the data. Therefore, a fixed proportion efficiency model was also applicable for these outputs as well as the inputs. It may be worthwhile to note that substitutability or lack thereof can be caused either by strict technical constraints, by other constraints such as regulations or norms, or simply by the environment. But, whatever the reason, if outputs (inputs) are *not* substituted for each other, then a *de facto* fixed proportion technology is present.

Efficiencies reported by the ADD and FPA models

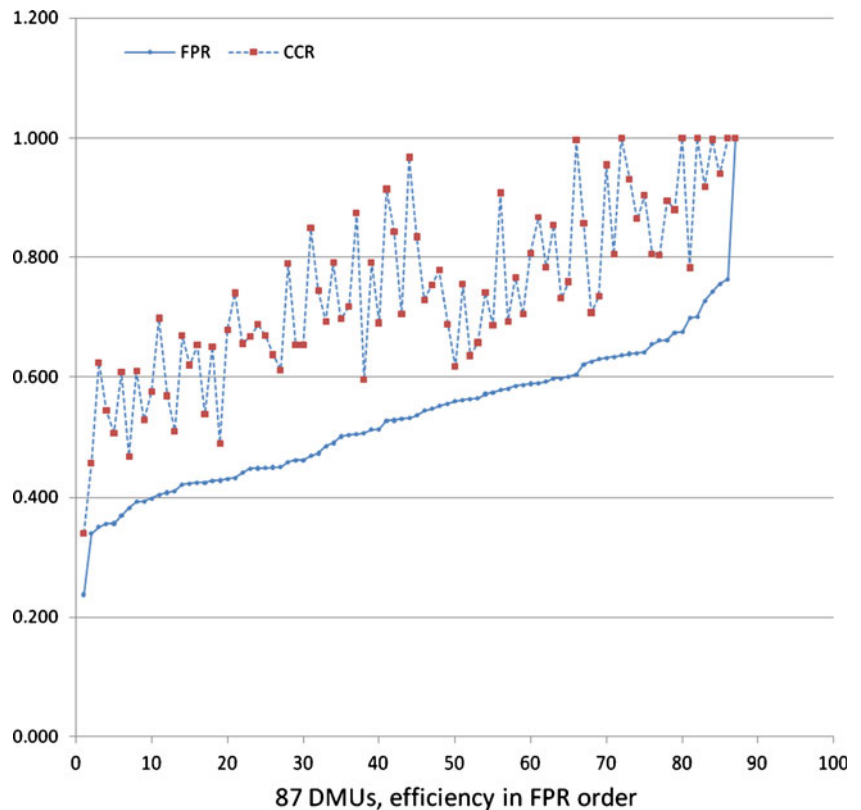
Using the FPA scores as the base, the ADD model reported efficiencies that were 42.4% greater on the average, ranging from 3.6% greater to 100% greater. The two models

measure efficiency the same way and only differ on their identification of efficient points based on whether or not the inputs were substitutable. We know that the FPA model was correct because the inputs are not substitutable. Thus, if the ADD model were (inappropriately) applied to these data, it would greatly overestimate mean efficiency. Moreover, the efficiency of some DMUs would be only slightly overestimated and the overestimation would be substantial for others. In short, in the presence of nonsubstitutable inputs, the conventional DEA additive model efficiency estimates were remarkably biased and showed strikingly low precision.

Efficiencies reported by the CCR and FPR models

Next, we compared scores of the FPR efficiency indicator with those of the CCR measure. As can be seen in Fig. 4, the CCR scores were much higher than the FPR scores, at all of the reported efficiency levels except for the highest

**Fig. 4** Comparison of CCR and FPR efficiency scores



one. Moreover, the difference between the CCR estimate and the FPR estimate varied substantially. The R-square value of FPR and CCR was 0.83 for all 87 DMUs, but only 0.33 for the 24 DMUs with highest efficiencies. Also, for the highest 24, the Spearman rank coefficients between FPR and CRR was 0.63, with the Kendall rank coefficients was 0.50. Further, the rank order of some DMUs' FPR scores was substantially different from their CRR scores. The CCR ranks ranged from 29 higher to 24 lower than the FPR ranks. Finally, six hospitals were reported efficient by DEA but inefficient by FPR. So, using nonsubstitutable inputs and nonsubstitutable outputs, the conventional DEA radial model's efficiency estimates had a very large upward bias and low precision.

### Summary of results

For our sample of 87 community hospitals, empirical testing showed that staff and bed inputs were not substitutable, nor were inpatient and outpatient outputs, thus violating DEA's substitutability requirements. Comparison of the DEA additive model with an additive model that assumed nonsubstitutability showed the DEA efficiency estimates to be highly biased upward on the average, but some DMUs showed little bias while others showed huge bias. Similar results occurred in a comparison of a DEA radial model and a new ratio model that assumed nonsubstitutability, with the DEA scores showing a large upward bias and low precision.

### Discussion

In some hospital efficiency studies, the effects of using DEA with nonsubstitutable inputs (outputs) may be less severe than they were with our sample. But, the effects in other studies might be even worse than ours. Thus, if DEA is used in hospital efficiency studies without having addressed the issue of input and output substitutability, then the efficiency estimates would be open to question.

Although inpatients and outpatients were not substitutable for each other in our sample of community hospitals, this lack of substitutability may not always be the case. However, our sample shows that it should never be assumed that inpatient and outpatient substitutability exists without empirical testing to justify the assumption. In the case of inputs, it seems unlikely that staffed beds and employment could be substitutable under any circumstances.

Therefore, we suggest that conventional DEA hospital efficiency applications should never include both employment and beds as separate inputs, and DEA should include both inpatients and outpatients as separate outputs only if it has been shown that they are substitutes in the dataset being used.

If all inputs and all outputs are nonsubstitutable, then one alternative efficiency measure would be the FPR indicator that we presented in this paper. Using this measure, employment and beds could be included as separate inputs, as could inpatients and outpatients for cases where they are not substitutes. However, use of the FPR measure would *not* be appropriate if some of the inputs (outputs) were substitutable and others were not, or if all of the inputs and outputs were substitutable.

There are several methods that permit the use of conventional DEA models without suffering the bias and precision problems illustrated in this paper. One method is to aggregate nonsubstitutable variables using their prices (or some other logical choice) as weights. We believe that this solution is a good one for inputs if prices are available and can be adjusted for price differences over time and among DMUs. For hospitals, this solution might lead to using total operating costs (perhaps adding depreciation) as the sole input variable, and using some reasonable weighting scheme to aggregate inpatients and outpatients into one output variable if the two are not substituted in the sample at hand.

A second solution is to use conventional DEA models but utilize only one of the nonsubstitutable inputs and one of the nonsubstitutable outputs. Because nonsubstitutable variables occur in a fixed proportion for efficient DMUs, they will increase and decrease together. So, one can serve as a rough proxy for all. The problem with this approach is that it does not account for differences in a DMU's efficiency in producing different outputs or in using different inputs. But, in the absence of comparable prices or other acceptable weighting schemes, it may be the best choice available if one wishes to use conventional DEA models. For hospitals, it would seem to us that the best input variable would be FTE employment, and the best output variable would be number of inpatients, because employment seems to be driven primarily by the inpatient load with other factor inputs seeming to have little effect.

A third way of using conventional DEA models is to combine nonsubstitutable variables with methodologies such as Factor Analysis or Principle Components Analysis, [27], or a variation of two-stage regression analysis [28, 29]. There undoubtedly are other applicable statistical methodologies.

### Conclusions

This paper identified the effects on efficiency estimates when conventional DEA models are applied to hospitals that employ a fixed proportion technology. For our sample of community hospital data, the inputs and outputs both occurred in fixed proportions. As a result, the DEA efficiency estimates were substantially biased and provided little precision.



We suggest that when DEA models are to be used, all potential inputs (outputs) be empirically tested to assure that substitutability exists. If any inputs (outputs) are not substitutable for each other, then, before applying DEA, the nonsubstitutable variables should be combined using an appropriate weighting scheme or statistical methodology, or only one of the nonsubstitutable inputs (outputs) should be included. If the analyst wishes to include nonsubstitutable variables, then efficiency models allowing nonsubstitutability should be used.

## References

- Hollingsworth, B., The measurement of efficiency and productivity of health care delivery. *Health Econ.* 17(10):1107–1128, 2008.
- Chern, J. Y., and Wan, T. T. H., The impact of the prospective payment system on the technical efficiency of hospitals. *J. Med. Syst.* 24(3):159–172, 2000.
- Sahin, I., and Ozcan, Y. A., Public sector hospital efficiency for provincial markets in Turkey. *J. Med. Syst.* 24(6):307–320, 2000.
- Kirigia, J. M., Emrouznejad, A., and Sambo, L. G., Measurement of technical efficiency of public hospitals in Kenya: using data envelopment analysis. *J. Med. Syst.* 26(1):39–45, 2002.
- Wan, T. T. H., Lin, B. Y. -J., and Ma, A., Integration mechanisms and hospital efficiency in integrated health care delivery systems. *J. Med. Syst.* 26(2):127–143, 2002.
- Harrison, J. P., Coppola, M. N., and Wakefield, M., Efficiency of federal hospitals in the United States. *J. Med. Syst.* 28(5):411–422, 2004.
- Hajialiafzali, H., Moss, J., and Mahmood, M., Efficiency measurement for hospitals owned by the Iranian social security organisation. *J. Med. Syst.* 31(3):166–172, 2007.
- Kirigia, J., Emrouznejad, A., Cassoma, B., Asbu, E., and Barry, S., A performance assessment method for hospitals: the case of municipal hospitals in Angola. *J. Med. Syst.* 32(6):509–519, 2008.
- Nayar, P., and Ozcan, Y., Data envelopment analysis comparison of hospital efficiency and quality. *J. Med. Syst.* 32(3):193–199, 2008.
- Lee, K. H., Yang, S. B., and Choi, M., The association between hospital ownership and technical efficiency in a managed care environment. *J. Med. Syst.* 33:307–315, 2009. doi:10.1007/s10916-008-9192-2.
- Hollingsworth, B., and Street, A., The market for efficiency analysis of health care organizations. *Health Econ.* 15:1055–1059, 2006.
- Beattie, B. R., and Taylor, C. R., *The economics of production*. Wiley, New York, 1985.
- Farrell, M. J., The measurement of productive efficiency. *J. R. Stat. Soc., A.* 120(3):253–290, 1957.
- Charnes, A., Cooper, W. W., and Rhodes, E., Measuring the efficiency of decision making units. *Eur. J. Oper. Res.* 2(6):429–444, 1978.
- Banker, R. D., Charnes, A., and Cooper, W. W., Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Manag. Sci.* 30(9):1078–1092, 1984.
- Färe, R., Grosskopf, S., and Lovell, C. A. K., *Production frontiers*. Cambridge University Press, Cambridge, 1994.
- Cooper, W. W., Seiford, L. M., Tone, K., and Zhu, J., Some models and measures for evaluating performances with DEA: past accomplishments and future prospects. *J. Prod. Anal.* 28(3):151, 2007.
- Fried, H. O., Lovell, C. A. K., and Schmidt, S. S., *The measurement of productive efficiency and productivity growth*. Oxford University Press, Oxford, 2008.
- Casu, B., and Thanassoulis, E., Evaluating cost efficiency in central administrative services in UK universities. *Omega.* 34(5):417–426, 2006.
- Banker, R. D., Chang, H., Janakiraman, S. N., and Konstans, C., A balanced scorecard analysis of performance metrics. *Eur. J. Oper. Res.* 154(2):423–436, 2004.
- Holvad, T., Hougaard, J. L., Kronborg, D., and Kvist, H. K., Measuring inefficiency in the Norwegian bus industry using multi-directional efficiency analysis. *Transportation.* 31(3):349–369, 2004.
- Cooper, W. W., Seiford, L. M., and Tone, K., *Data envelopment analysis: a comprehensive text with models, applications, references and DEA-Solver software*. Springer, New York, 2007.
- Bogetoft, P., and Hougaard, J. L., Efficiency evaluations based on potential (non-proportional) improvements. *J. Prod. Anal.* 12(3):233–247, 1999.
- Charnes, A., Cooper, W. W., Lewin, A. Y., and Seiford, L. M., *Data envelopment analysis: theory, methodology and application*. Kluwer Academic, Boston, 1994.
- Gupta, S. R., Wojtynek, J. E., Walton, S. M., Botticelli, J. T., Shields, K. L., Quad, J. E., and Schumock, G. T., Pharmacy staffing and productivity. *Am. J. Health. Syst. Pharm.* 64(22):2323–2324, 2006.
- Sikka, V., Luke, R. D., and Ozcan, Y. A., The efficiency of hospital-based clusters: evaluating system performance using data envelopment analysis. *Health Care Manage. Rev.* 34(3):251–261, 2009.
- Pedhazur, E. J., and Schmelkin, L. P., *Measurement, design, and analysis: an integrated approach*. Lawrence Erlbaum Associates, Hillsdale, 1991.
- Wooldridge, J. M., *Econometric analysis of cross section and panel data*. MIT, Cambridge, 2002.
- Greene, W. H., *Econometric analysis*, 6th edition. Prentice Hall, Upper Saddle River, 2008.