

1  
2  
3  
4 **Measuring Hospital Efficiency with Data Envelopment Analysis:**  
5 **Nonsubstitutable vs. Substitutable Inputs and Outputs**  
6  
7

8  
9 by

10  
11 Darold T Barnum, Surrey M Walton, Karen L Shields, and Glen T Schumock  
12  
13

14  
15 **Darold T Barnum**, PhD, MBA. Departments of Managerial Studies, Information & Decision  
16 Sciences, and Pharmacy Administration, University of Illinois at Chicago, Chicago, IL, USA.  
17  
18

19  
20 **Surrey M Walton**, PhD. Center for Pharmacoeconomic Research, Department of Pharmacy  
21 Administration, and Department of Economics, University of Illinois at Chicago, Chicago, IL, USA.  
22  
23

24  
25 **Karen L Shields**, BSP Pharm. Sisters of St. Francis Health Services, Inc., Mishawaka, IN, USA.  
26  
27

28  
29 **Glen T Schumock**, PharmD, MBA, FCCP. Center for Pharmacoeconomic Research, and  
30 Department of Pharmacy Practice, University of Illinois at Chicago, Chicago, IL, USA.  
31  
32

33 **Corresponding author information:** Darold T Barnum, 601 South Morgan Street, Department of  
34 Managerial Studies (M/C 243), University of Illinois at Chicago, Chicago, IL 60607-7123 USA.  
35 Voice +1-312-996-3073, Fax +1-312-996-3559, Email [dbarnum@uic.edu](mailto:dbarnum@uic.edu).  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4  
5  
6  
7 **Measuring Hospital Efficiency with Data Envelopment Analysis:**  
8 **Nonsubstitutable vs. Substitutable Inputs and Outputs**  
9

10  
11  
12 **Abstract:** There is a conflict between Data Envelopment Analysis (DEA) theory's requirement that inputs  
13 (outputs) be substitutable, and the ubiquitous use of nonsubstitutable inputs and outputs in DEA applications to  
14 hospitals. This paper develops efficiency indicators valid for nonsubstitutable variables. Then, using a sample  
15 of 87 community hospitals, it compares the new measures' efficiency estimates with those of conventional DEA  
16 measures. DEA substantially overestimated the hospitals' efficiency on the average, and reported many  
17 inefficient hospitals to be efficient. Further, it greatly overestimated the efficiency of some hospitals but only  
18 slightly overestimated the efficiency of others, thus making any comparisons among hospitals questionable.  
19 These results suggest that conventional DEA models should not be used to estimate the efficiency of hospitals  
20 unless there is empirical evidence that the inputs (outputs) are substitutable. If inputs (outputs) are not  
21 substitutes, efficiency indicators valid for nonsubstitutability should be employed, or, before applying DEA, the  
22 nonsubstitutable variables should be combined using an appropriate weighting scheme or statistical  
23 methodology.  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46

47 **Keywords:** Data envelopment analysis; Efficiency; Hospitals; Fixed proportion technology.  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

## 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 two 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 nonsubstitutable 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

1  
2  
3  
4 quantity of an input in excess of the required ratio is wasted. Production systems using nonsubstitutable  
5  
6 inputs are well-known in economics, and are called “Fixed Factor Proportion Technologies” [12]. If  
7  
8 outputs are nonsubstitutable, then, for a fixed amount of input, production of one output cannot be  
9  
10 increased by producing less of another. Such production systems are “Fixed Product Proportion  
11  
12 Technologies” [12].  
13

14  
15 When inputs are substitutable for each other in the production of output, a fixed amount of output can  
16  
17 be produced with varying combinations of the inputs. When outputs are substitutable for each other, the  
18  
19 amount of one output can be increased and the amount of another decreased for a fixed amount of input  
20  
21 by changing the proportion of the input that each output receives.  
22

23  
24 DEA mathematical models and the economic theory underlying them require substitutability. Farrell  
25  
26 [13] and Charnes, Cooper and Rhodes [14] assume substitutability, as does Banker, Charnes and Cooper  
27  
28 [15], Färe, Grosskopf and Lovell [16] and more recent work [17, 18].  
29

### 30 31 32 Input and output substitutability in DEA applications 33

34 The issue of substitutability in DEA application papers has rarely been addressed. Although none of  
35  
36 them involved healthcare, we do know of two recent articles that purposely selected inputs and outputs  
37  
38 that avoided nonsubstitutability [19, 20] , and another article that used Multi-Directional Efficiency  
39  
40 Analysis instead of DEA partly because of the substitutability issue [21].  
41

42  
43 In DEA applications to hospitals, Hollingsworth [1, p. 1110] reports that inputs “are mainly measures  
44  
45 of staff and capital employed,” and most of the studies use output measures “such as inpatient days or  
46  
47 discharges.” Recent hospital DEA articles in the *Journal of Medical Systems* are consistent with the  
48  
49 widespread use of such inputs and outputs. The nine hospital DEA papers published in the journal since  
50  
51 2000 all included staffed beds as a proxy for capital, and the number of employees (in various categories)  
52  
53 as a proxy for labor. Also, all of the nine papers included outputs separately measuring the numbers of  
54  
55 inpatients and outpatients [2-10].  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4 As discussed later, these labor and capital input proxies cannot be substituted for each other in the  
5  
6 production of a fixed amount of output. And, although it would seem that outputs such as inpatients and  
7  
8 outpatients would always be substitutes, they were not substituted for each other in our sample hospitals.  
9  
10 In short, the hospitals that we studied employed fixed proportion technologies.

#### 11 12 13 14 DEA theory vs. DEA applications to hospitals

15  
16 In sum, DEA applications measuring hospital efficiency have employed inputs and outputs that  
17  
18 conflict with DEA theory. This inconsistency could result in trivial effects on DEA efficiency scores, or  
19  
20 could cause substantial and significant errors in DEA efficiency estimates.  
21  
22

23 The purpose of this paper is to examine the impact of DEA theory's substitutability requirements on  
24  
25 its applications to hospitals. To analyze the effects of the conflict between DEA theory and hospital  
26  
27 applications, we developed efficiency indicators that assume nonsubstitutability rather than  
28  
29 substitutability. Next, using hospital-wide data supplied by the pharmacy departments of US community  
30  
31 hospitals, we ascertained empirically that their inputs and outputs were indeed nonsubstitutable, and then  
32  
33 we compared their DEA scores with the scores from our new efficiency measures.  
34  
35

36 For our sample hospitals, DEA resulted in severely biased and imprecise estimates of efficiency. All  
37  
38 hospitals were less efficient in truth than estimated by DEA, and DEA reported many inefficient hospitals  
39  
40 to be efficient. Further, the efficiency scores of some hospitals were only slightly affected while the  
41  
42 efficiency scores of others showed large biases, thereby making comparisons among hospitals  
43  
44 unattainable. Of course, we don't know if other DEA hospital efficiency studies suffer to the same extent,  
45  
46 but we do suggest that future studies restrict DEA inputs and outputs to substitutable variables or use  
47  
48 efficiency indicators not requiring substitutability.  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4 **Methods**  
5

6 Location of the production frontier when inputs are substitutable and nonsubstitutable  
7

8 In DEA, the organizations being analyzed often are called Decision Making Units (DMUs). In this  
9 paper, the organizations being analyzed are individual hospitals, so each hospital is a DMU. In order to  
10 be consistent with the DEA literature, we often use the term DMUs to refer to the hospitals.  
11  
12

13 Consider a hypothetical case of one output and two inputs (a) when the inputs are substitutable, and  
14 (b) when the inputs are not substitutable. Suppose one unit of output is produced by each organization  
15 being analyzed with various combinations of the two inputs (Figure 1). If the two inputs used to produce  
16 one unit of output are substitutable inputs, a representation of the production frontier is shown by the  
17 inner-most piecewise isoquant. That is, if less of one input is used, some amount more of the other input  
18 must be used in its place to hold output constant. Conventional DEA models would report that the four  
19 DMUs defining the isoquant are efficient, because, though having different mixes of the inputs, they are  
20 all on the same inner-most isoquant.  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31

32 -Put Figure 1 about here-  
33  
34

35 If the two inputs are not substitutable, then an efficient DMU must use them in a fixed proportion.  
36 Suppose that the inputs are truly nonsubstitutable and one unit of output is produced, as is shown in  
37 Figure 2 (using the same data as Figure 1). The production frontier now consists of a single point that is,  
38 a *point frontier*. This frontier is estimated by the composite DMU in the south-west corner of the graph.  
39 The frontiers of the reference set increase vertically and horizontally from this point, forming a right-  
40 angle or L-shaped reference set frontier. However, the only Pareto-Koopmans efficient subset of the  
41 reference set frontier is the point frontier, because only at that point is the requisite output achieved  
42 without the overuse of one of the inputs [22].  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52

53 -Put Figure 2 about here-  
54  
55

56 For substitutable inputs, the minimum level of each input is conditioned on the level of the other  
57 input. However, for nonsubstitutable inputs, that is, fixed factor proportion technologies, the minimum  
58  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4 level of each input needed to produce a given amount of output is not influenced by the other input [12].  
5  
6 So, if all DMUs' outputs are equal, when inputs are nonsubstitutable it is only necessary to find the  
7  
8 minimum level of each input. As is true for conventional DEA, this frontier estimation method envelops  
9  
10 the data, and assumes that a composite DMU can be used to identify a point on the efficient frontier that  
11  
12 is attainable by an actual DMU. As also is true for DEA, this deterministic measure estimates efficiency  
13  
14 based on the most extreme observations, ignoring any stochastic variation that might exist.  
15  
16

17  
18 Note that all DMUs' reported efficiencies will be quite different when the point frontier is used in  
19  
20 place of the isoquant frontier as the efficient reference. We return to the issue of efficiency indices for the  
21  
22 point frontier after discussing methods for identifying whether or not inputs (outputs) are substitutable.  
23  
24

#### 25 Method for identifying the presence or absence of substitutability 26

27  
28 Because we know that the DMUs shown in Figures 1 and 2 each produced one unit of output, it might  
29  
30 appear that the empirical evidence suggests that these two inputs are substitutable for each other, as  
31  
32 shown by the isoquant in Figure 1. But, substitutability is not necessarily present because a DMU is  
33  
34 unlikely to be equally efficient in its use of both inputs [23]. In Figure 1, for example, if inputs are truly  
35  
36 nonsubstitutable, the supposed piecewise isoquant frontier may be the result of one DMU being the most  
37  
38 efficient of all DMUs in the use of the first input but less efficient in the use of the second, and another  
39  
40 DMU being the most efficient of all DMUs in the use of the second input but less efficient in the use of  
41  
42 the first. Substitutability, or the lack thereof, can be identified by logic and statistical testing, but cannot  
43  
44 be identified by a deterministic estimation of an alleged best-practice frontier.  
45  
46

47  
48 One simple method for assessing whether or not inputs are substitutable is to regress each input on the  
49  
50 remaining inputs and all outputs. If any two inputs are substitutes, then the relationship between them  
51  
52 must be negative (with statistical significance). Because the remaining inputs and all outputs are held  
53  
54 constant, a decrease in any one input would have to be compensated for by an increase in the other input.  
55  
56 If the two inputs are not substitutes, then there will be no statistically significant relationship between  
57  
58  
59  
60  
61  
62  
63  
64  
65

1  
2  
3  
4 them if their inefficiencies are uncorrelated. There will be a statistically significant positive relationship  
5  
6 if there is a high degree of correlation between the two inputs' inefficiency levels.  
7

8  
9 Empirically estimating whether outputs are substitutable or nonsubstitutable follows the same  
10 methodology used for inputs. Each output is regressed on the remaining outputs and all inputs. If any  
11 two outputs are substitutes, then the relationship between them must be negative to a statistically  
12  
13 significant degree, because an increase in any one output would have to be compensated for by a decrease  
14  
15 in another output. And, if two outputs are not substitutes, then, as with inputs, there will be no  
16  
17 statistically significant relationship between them if their degree of inefficiencies are uncorrelated, and  
18  
19 there will be a statistically significant positive relationship if there is a high degree of correlation between  
20  
21 their inefficiency levels.  
22  
23  
24  
25

26  
27 In truth, the inputs in Figures 1 and 2 are not substitutable for each other. One of the inputs is staffed  
28  
29 beds and the other is number of staff used for one unit of inpatient output. Logically, it is not possible to  
30  
31 serve a fixed number of inpatients by decreasing one of these inputs and making up for the decrease by  
32  
33 increasing the other. Statistically, because there is one unit of output and two inputs, we can regress one  
34  
35 of the inputs on the other to determine whether there is a statistically significant negative relationship  
36  
37 (indicating the inputs are substitutable) or not. In fact, the two inputs show a positive relationship,  
38  
39 confirming that they are nonsubstitutable.  
40  
41

42 The next task is to develop efficiency measures for cases of nonsubstitutable inputs. Then, we can  
43  
44 compare the new efficiency measures with DEA efficiency estimates.  
45  
46

#### 47 Additive efficiency measure when inputs are nonsubstitutable

48  
49 In order to estimate each DMU's efficiency relative to the point frontier in Figure 2, one simple  
50  
51 possibility would be a variation on the DEA Additive (ADD) model [22]. With DEA's ADD model, the  
52  
53 efficient point for an assessed DMU is the furthest point on the piecewise isoquant frontier where neither  
54  
55 of its inputs has increased and its output has not decreased. The rectilinear distance between the assessed  
56  
57 DMU and that point measures the DMU's inefficiency.  
58  
59



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 Figure 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.

-Put Figure 3 about here-

The FPA efficiency score for the assessed  $DMU_k$  can be obtained for each  $DMU_k$  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 equation 1.

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

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 (equation 1) and the ADD model (equations 2-5) are identical, so the resulting sums of the slacks for the target DMU  $k$  are directly comparable.

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

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

$$\hat{y}_{j1} = 1 = y_{j1} / y_{j1} \quad j = 1, 2, \dots, J \quad (3)$$

$$\lambda_j \geq 0 \quad (4)$$

### 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.

$$eff_{kmn} = \left[ \frac{y_{kn} / x_{km}}{\text{Max}_j (y_{jn} / x_{jm})} \right] \quad (5)$$

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, equation 7 is used to compute the normalized efficiency measure for that ratio. If there are  $m$  inputs and  $n$  outputs, then there will be  $m \times n$  efficiency measures of the form  $eff_{kmn}$ . So, for each DMU  $k$  in a set of  $J$  DMUs, the mean of its  $m \times n$  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:

$$FPR_k = \frac{(1 / MN) \sum_{m=1}^M \sum_{n=1}^N eff_{kmn}}{\text{Max}_j \left( (1 / MN) \sum_{m=1}^M \sum_{n=1}^N eff_{jmn} \right)} \quad (6)$$

#### 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.)

1  
2  
3  
4 **Materials: Sample, inputs and outputs**  
5

6  
7 Our sample consisted of data from 87 community hospitals in the United States that were members of  
8  
9 a national group purchasing organization. The data were collected for use in an earlier study of  
10  
11 community hospital pharmacy productivity [25], from an online questionnaire that was completed by  
12  
13 pharmacy directors at the hospitals. Herein, we used hospital-wide data from the hospitals that included  
14  
15 all of the inputs and outputs that we needed for our computations.  
16  
17

18 For tests comparing the FPA and ADD additive models, the one output was total inpatients, and the  
19  
20 two inputs were staffed beds and full-time-equivalent (FTE) employment. It was only possible to use one  
21  
22 output in comparisons of these additive models, so we chose the output that had by far the strongest  
23  
24 impact on the levels of inputs needed.  
25  
26

27 For tests comparing the FPR and the CCR models, multiple inputs and outputs are possible. We used  
28  
29 the two most common outputs, total inpatients and total outpatients, and, as before, the two inputs were  
30  
31 staffed beds and FTE employment. Summary values for the 87 hospitals are shown in Table 1.  
32  
33

34 --put Table 1 about here--  
35

36 **Results**  
37

38 Input and output substitutability  
39

40 One input was regressed on the other input, with the outputs included as control variables. As Table  
41  
42 2 shows, the number of employees was positively related to the number of beds with statistical  
43  
44 significance of 0.059, which, based on our earlier logical argument, would be as hypothesized. More  
45  
46 important, there was not a negative relationship, statistically significant or otherwise, and a negative  
47  
48 relationship would be necessary if the factors could be substituted for each other. Using a different proxy  
49  
50 for capital *might* result in a different conclusion, but we used the proxy that has been empirically  
51  
52 validated and is common to most published hospital DEA articles [26]. Therefore, a fixed factor  
53  
54 proportion technology was present.  
55  
56

57  
58 --put table 2 about here--  
59  
60

1  
2  
3  
4 One output was regressed on the other, with the inputs included as control variables. As Table 3  
5  
6 shows, the relationship between outpatients and inpatients was positive, a somewhat surprising finding.

7  
8 --put table 3 about here--  
9

10 However, looking again at Table 2, it can be seen that the number of outpatients had relatively minor  
11  
12 influence on the number of employees. This apparent lack of influence may have resulted from the  
13  
14 narrow range within which the outpatient and inpatient proportions occurred for our sample. Except for a  
15  
16 very few hospitals, the proportion of outpatients clustered between 90 and 97 percent, out of a possible  
17  
18 range from 0 to 100. So, this appears to be a case of our community hospital sample all having about the  
19  
20 same ratio of inpatients to outpatients, rather than a case of true technical nonsubstitutability. However,  
21  
22 from the viewpoint of modeling choice, the reason for the empirical lack of substitutability does not  
23  
24 matter and we have to honor the data. Therefore, a fixed proportion efficiency model was also applicable  
25  
26 for these outputs as well as the inputs. It may be worthwhile to note that substitutability or lack thereof  
27  
28 can be caused either by strict technical constraints, by other constraints such as regulations or norms, or  
29  
30 simply by the environment. But, whatever the reason, if outputs (inputs) are not substituted for each  
31  
32 other, then a *de facto* fixed proportion technology is present.  
33  
34  
35  
36  
37

### 38 39 Efficiencies reported by the ADD and FPA models 40

41 Using the FPA scores as the base, the ADD model reported efficiencies that were 42.4 percent greater  
42  
43 on the average, ranging from 3.6 percent greater to 100 percent greater. The two models measure  
44  
45 efficiency the same way and only differ on their identification of efficient points based on whether or not  
46  
47 the inputs were substitutable. We know that the FPA model was correct because the inputs are not  
48  
49 substitutable. Thus, if the ADD model were (inappropriately) applied to these data, it would greatly  
50  
51 overestimate mean efficiency. Moreover, the efficiency of some DMUs would be only slightly  
52  
53 overestimated and the overestimation would be substantial for others. In short, in the presence of  
54  
55 nonsubstitutable inputs, the conventional DEA additive model efficiency estimates were remarkably  
56  
57 biased and showed strikingly low precision.  
58  
59  
60

## 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 Figure 4, the CCR scores were much higher than the FPR scores, at all of the reported efficiency levels except for the highest 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.

1  
2  
3  
4           Although inpatients and outpatients were not substitutable for each other in our sample of community  
5  
6 hospitals, this lack of substitutability may not always be the case. However, our sample shows that it  
7  
8 should never be assumed that inpatient and outpatient substitutability exists without empirical testing to  
9  
10 justify the assumption. In the case of inputs, it seems unlikely that staffed beds and employment could be  
11  
12 substitutable under any circumstances.

13  
14  
15           Therefore, we suggest that conventional DEA hospital efficiency applications should never include  
16  
17 both employment and beds as separate inputs, and DEA should include both inpatients and outpatients as  
18  
19 separate outputs only if it has been shown that they are substitutes in the dataset being used.

20  
21  
22           If all inputs and all outputs are nonsubstitutable, then one alternative efficiency measure would be  
23  
24 the FPR indicator that we presented in this paper. Using this measure, employment and beds could be  
25  
26 included as separate inputs, as could inpatients and outpatients for cases where they are not substitutes.  
27  
28 However, use of the FPR measure would not be appropriate if some of the inputs (outputs) were  
29  
30 substitutable and others were not, or if all of the inputs and outputs were substitutable.

31  
32  
33           There are several methods that permit the use of conventional DEA models without suffering the bias  
34  
35 and precision problems illustrated in this paper. One method is to aggregate nonsubstitutable variables  
36  
37 using their prices (or some other logical choice) as weights. We believe that this solution is a good one  
38  
39 for inputs if prices are available and can be adjusted for price differences over time and among DMUs.  
40  
41 For hospitals, this solution might lead to using total operating costs (perhaps adding depreciation) as the  
42  
43 sole input variable, and using some reasonable weighting scheme to aggregate inpatients and outpatients  
44  
45 into one output variable if the two are not substituted in the sample at hand.

46  
47  
48           A second solution is to use conventional DEA models but utilize only one of the nonsubstitutable  
49  
50 inputs and one of the nonsubstitutable outputs. Because nonsubstitutable variables occur in a fixed  
51  
52 proportion for efficient DMUs, they will increase and decrease together. So, one can serve as a rough  
53  
54 proxy for all. The problem with this approach is that it does not account for differences in a DMU's  
55  
56 efficiency in producing different outputs or in using different inputs. But, in the absence of comparable  
57  
58

1  
2  
3  
4 prices or other acceptable weighting schemes, it may be the best choice available if one wishes to use  
5  
6 conventional DEA models. For hospitals, it would seem to us that the best input variable would be FTE  
7  
8 employment, and the best output variable would be number of inpatients, because employment seems to  
9  
10 be driven primarily by the inpatient load with other factor inputs seeming to have little effect.  
11  
12

13 A third way of using conventional DEA models is to combine nonsubstitutable variables with  
14  
15 methodologies such as Factor Analysis or Principle Components Analysis, [27], or a variation of two-  
16  
17 stage regression analysis [28, 29]. There undoubtedly are other applicable statistical methodologies.  
18  
19

## 20 21 **Conclusions**

22  
23 This paper identified the effects on efficiency estimates when conventional DEA models are applied  
24  
25 to hospitals that employ a fixed proportion technology. For our sample of community hospital data, the  
26  
27 inputs and outputs both occurred in fixed proportions. As a result, the DEA efficiency estimates were  
28  
29 substantially biased and provided little precision.  
30  
31

32 We suggest that when DEA models are to be used, all potential inputs (outputs) be empirically  
33  
34 tested to assure that substitutability exists. If any inputs (outputs) are not substitutable for each other,  
35  
36 then, before applying DEA, the nonsubstitutable variables should be combined using an appropriate  
37  
38 weighting scheme or statistical methodology, or only one of the nonsubstitutable inputs (outputs) should  
39  
40 be included. If the analyst wishes to include nonsubstitutable variables, then efficiency models allowing  
41  
42 nonsubstitutability should be used.  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

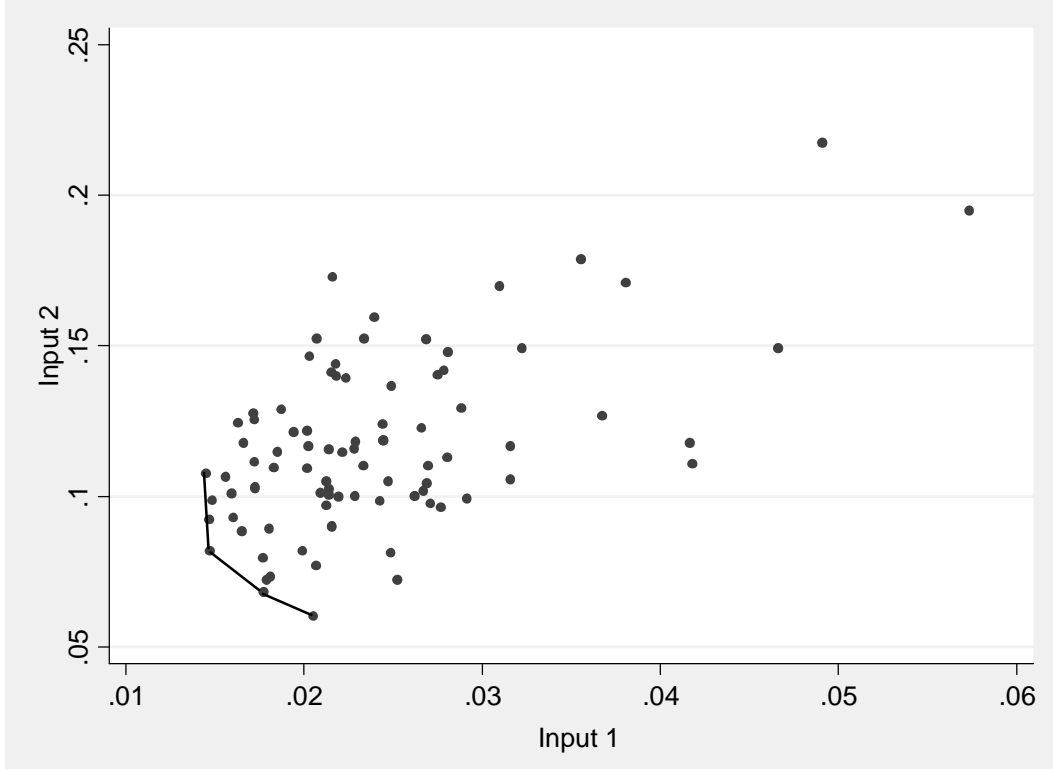


## References

1. Hollingsworth, B., The Measurement of Efficiency and Productivity of Health Care Delivery. *Health Econ.* 17(10): 1107-1128, 2008.
2. 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.
3. Sahin, I. and Ozcan, Y.A., Public Sector Hospital Efficiency for Provincial Markets in Turkey. *J. Med. Syst.* 24(6): 307-320, 2000.
4. 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.
5. 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.
6. 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.
7. 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.
8. 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.
9. Nayar, P. and Ozcan, Y., Data Envelopment Analysis Comparison of Hospital Efficiency and Quality. *J. Med. Syst.* 32(3): 193-199, 2008.
10. 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.* 2009 (In Press), <http://dx.doi.org/10.1007/s10916-008-9192-2>.
11. Hollingsworth, B. and Street, A., The market for efficiency analysis of health care organizations. *Health Econ.* 15: 1055-1059, 2006.
12. Beattie, B.R. and Taylor, C.R., *The Economics of Production*. Wiley, New York, 1985.
13. Farrell, M.J., The measurement of productive efficiency. *Journal of the Royal Statistical Society, Series A* 120(3): 253-290, 1957.
14. Charnes, A., Cooper, W.W., and Rhodes, E., Measuring the efficiency of decision making units. *Eur. J. Oper. Res* 2(6): 429-444, 1978.
15. 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.

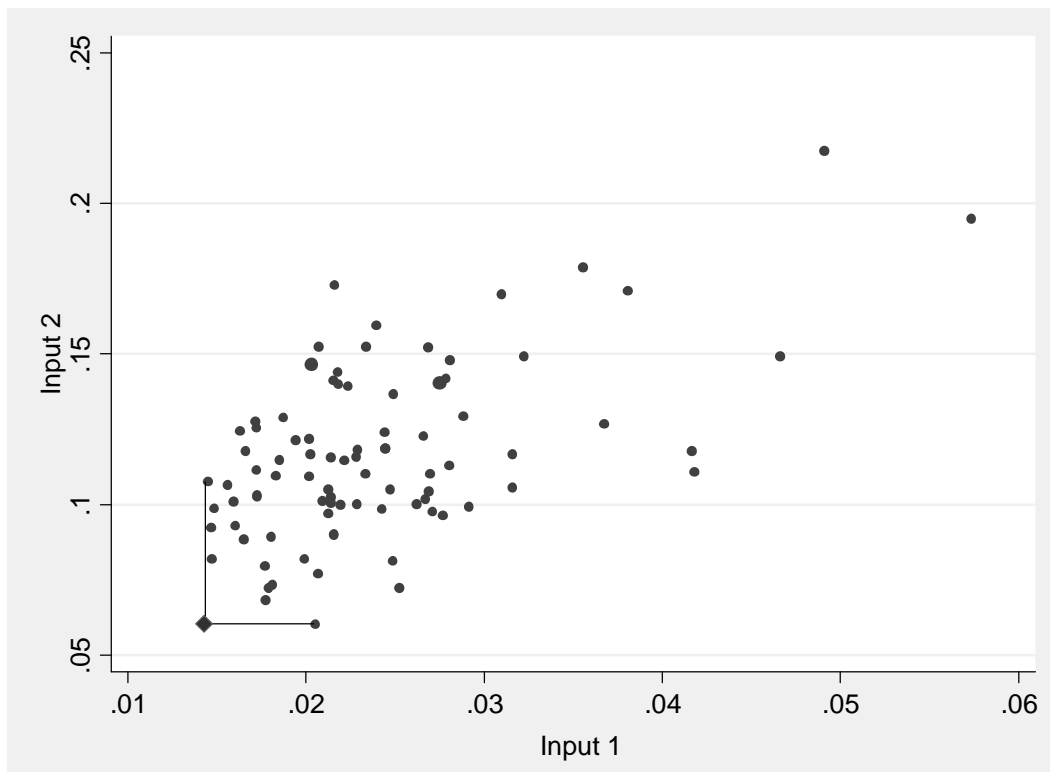
16. Färe, R., Grosskopf, S., and Lovell, C.A.K., *Production Frontiers*. Cambridge University Press, Cambridge, England, 1994.
17. Cooper, W.W., Seiford, L.M., Tone, K., and Zhu, J., Some models and measures for evaluating performances with DEA: past accomplishments and future prospects. *J Prod Anal* 28(3): 151, 2007.
18. Fried, H.O., Lovell, C.A.K., and Schmidt, S.S., *The Measurement of Productive Efficiency and Productivity Growth*. Oxford University Press, Oxford, 2008.
19. Casu, B. and Thanassoulis, E., Evaluating cost efficiency in central administrative services in UK universities. *Omega* 34(5): 417-426, 2006.
20. 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.
21. 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.
22. 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.
23. Bogetoft, P. and Hougaard, J.L., Efficiency Evaluations Based on Potential (Non-Proportional) Improvements. *J Prod Anal* 12(3): 233-247, 1999.
24. Charnes, A., Cooper, W.W., Lewin, A.Y., and Seiford, L.M., *Data Envelopment Analysis: Theory, Methodology and Application*. Kluwer Academic Publishers, Boston, 1994.
25. 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.
26. Sikka, V., Luke, R.D., and Ozcan, Y.A., The efficiency of hospital-based clusters: Evaluating sustem performance using data envelopment analysis. *Health Care Manage. Rev.* 34(3): 251-261, 2009.
27. Pedhazur, E.J. and Schmelkin, L.P., *Measurement, Design, and Analysis: An Integrated Approach*. Lawrence Erlbaum Associates, Hillsdale, NJ, 1991.
28. Wooldridge, J.M., *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge, MA, 2002.
29. Greene, W.H., *Econometric Analysis, 6th Ed*. Prentice Hall, Upper Saddle River, NJ, 2008.

Figure 1. Substitutable inputs best-practice (isoquant) frontier



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

Figure 2. Nonsubstitutable inputs best-practice (Point) frontier



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

Figure 3. Graphical representation of ADD and FPA measures

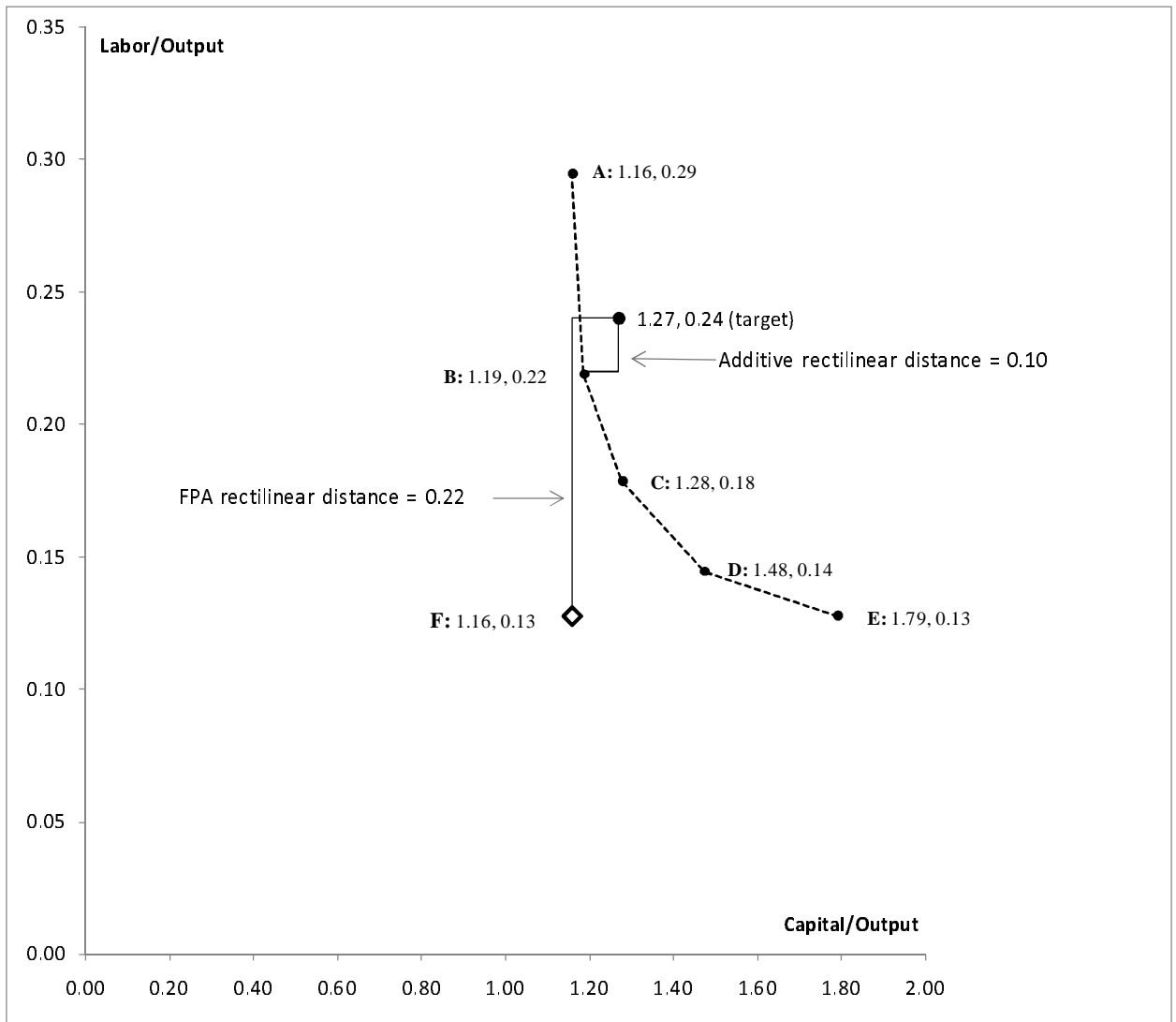


Figure 4. Comparison of CCR and FPR Efficiency Scores

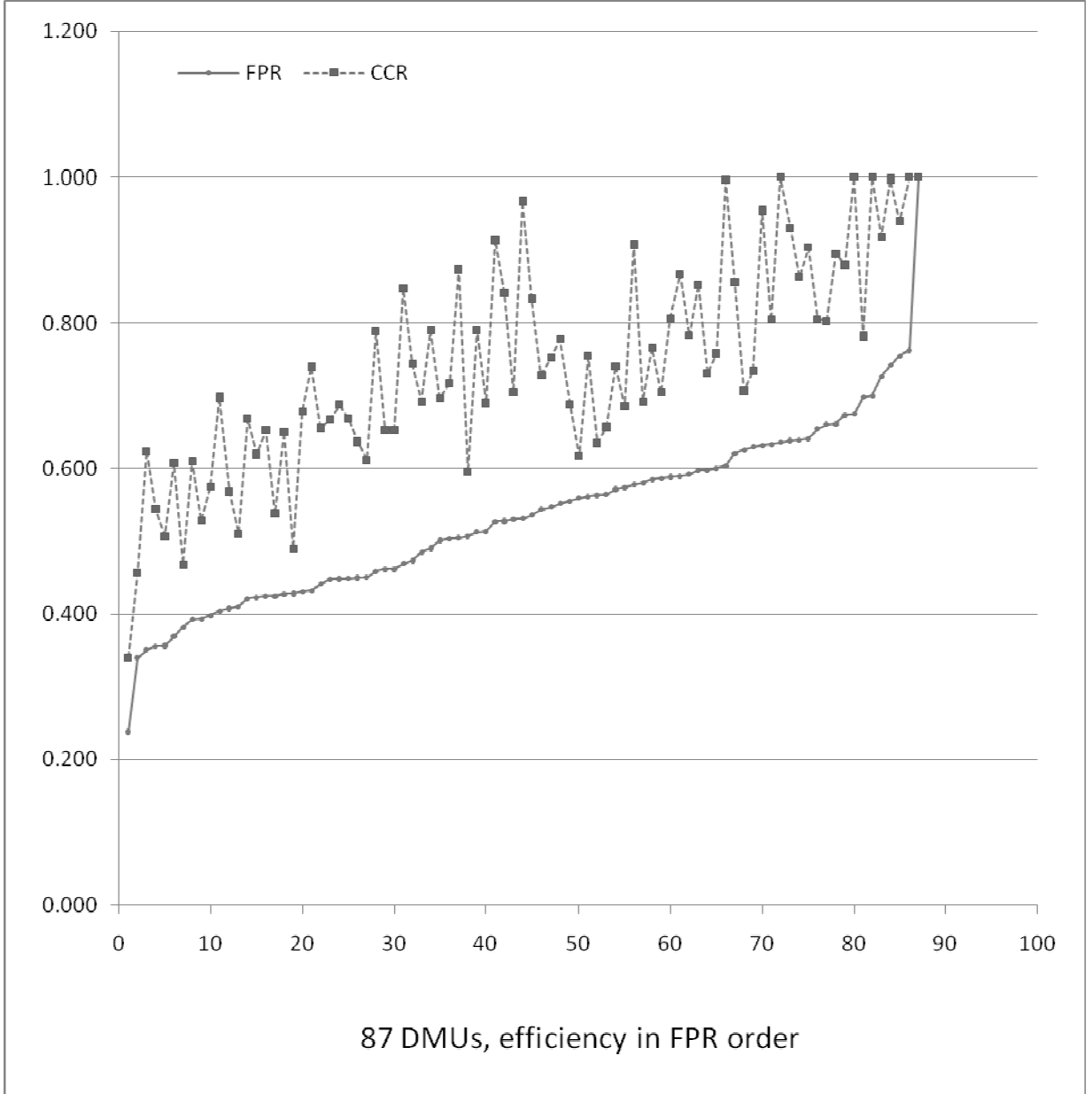


Table 1. Summary Statistics for 87 Community Hospitals

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

Table 2. Regression of Hospital Employees on Beds, Holding Inpatients and Outpatients Constant

Total FTE Employees	Coefficient	Robust Std. Err.	t	P>t (1-tail)
Staffed Beds	.8950035	.5668977	1.58	0.059
Annual Inpatients	.0817770	.0115093	7.11	0.000
Annual Outpatients	.0007314	.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.	t	P>t (2-tail)
Annual Outpatients	.0033999	.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

## Authors' Response to Reviewers

### Reviewer #2

This is a very well-written paper that deals with an issue that heretofore has not been properly recognized in the theoretical and (more importantly) the applied Data Envelopment Analysis literature. It raises serious questions about the validity of many published DEA analyses. The authors have made a good case regarding the substitutability/non-substitutability issue, and have developed procedures to deal appropriately with the case of non-substitutability. This paper definitely should be published.

Thank you!

1. Figure 3 is cited on page 7, but the heading of the figure refers to the ADD and FPA models, and these models are not discussed until page 9. The point the authors make on page 7 is supported by the graph in Figure 3 (that is, knowledge of the ADD and FPA models is not necessary in order to understand the point made on page 7), but the reader may be a bit confused because the title of the figure incorporates concepts not yet discussed.

We've moved the discussion related to Figure 3 to a point following where the ADD and FPA models are discussed.

2. In Tables 2 and 3, center the fourth column heading (the 't') over the column, rather than right-justifying it.

Done.

### Reviewer #3

The paper focuses on the conflict between the DEA theory and its application in hospitals. Measures are proposed to deal with the nonsubstitutability to measure the degree of efficiency / inefficiency of DEA application. The authors have done a good work in bringing out the necessity to propose such a measure.

Thank you!

The equation fonts are not appropriate and need to be edited to make the equation more clearer and readable. If required some equations can be split in to two equations for easier understanding.

We've used fonts from the MathType equation editor in Word, and increased the font size to make them clearer and more readable.



## Reviewer #4

. . . using regressions to address the substitution issue is not appealing. DEA is a non-parametric technique that has been used in over 2,000 published studies: precisely because of its flexibility in dealing with situations where the strict economic assumptions may not be met.

We appreciate the reviewer's perspective.

However, DEA does require strict adherence to a small set of key economic assumptions, one of them being the existence of substitutability. There can be no isoquants unless inputs are substitutable, and no production possibility frontiers unless outputs are substitutable. DEA estimation is based on isoquants and production possibility frontiers, so, if they are absent, then DEA estimates are erroneous. The only question is whether the errors are trivial or not. As we say in the paper,

DEA mathematical models and the economic theory underlying them require substitutability. Farrell (Farrell, 1957) and Charnes, Cooper and Rhodes (Charnes, Cooper and Rhodes, 1978) assume substitutability, as does Banker, Charnes and Cooper (Banker, Charnes and Cooper, 1984), Färe, Grosskopf and Lovell (Färe, Grosskopf and Lovell, 1994) and more recent work (Cooper *et al.*, 2007; Fried, Lovell and Schmidt, 2008).

The reviewer correctly states that DEA is a nonparametric technique. However, although DEA linear programs are nonparametric, DEA scores produced by the programs are stochastic, composed of the DMU's true efficiency and an error component (Banker and Natarajan, 2004). A decade ago, Seiford (1996, pp. 106-107) observed that stochastic DEA "appears on almost everyone's list of future research areas for DEA and presents a formidable challenge." He argued that stochastic DEA was the "most critical and difficult future issue in DEA," agreeing with Lovell, Grosskopf *et al.* (1994) that statisticians and others would remain skeptical of managerial and policy implications of DEA until there was a valid methodology for estimating true efficiencies from stochastic DEA scores.

Since 1996, statistical methodologies have often been used in DEA applications. It has been shown that DEA scores possess statistical characteristics that permit many types of hypothesis tests (Banker and Natarajan, 2004; Grosskopf, 1996), and the variations in DEA scores have been addressed with methodologies such as chance-constrained programming (Cooper *et al.*, 2002; Tser-yieth, 2002), bootstrapping (Simar and Wilson, 2000; Simar and Wilson, 2007), window analysis (Cooper, Seiford and Zhu, 2004), use of means (Ruggiero, 2004), and sensitivity-robustness-stability analysis (Cherchye *et al.*, 2008; Cooper *et al.*, 2004; Neralic and Wendell, 2004), and, most recently, Panel Data Regression (Barnum *et al.*, 2009a).

Indeed, a recent paper in the *Journal of Medical Systems* uses regression to estimate confidence intervals and trends for DEA efficiency in hospital pharmacies (Barnum *et al.*, 2009b).

In short, we must respectively disagree with the reviewer. Substitution requirements are one of the economic assumptions that strictly and absolutely must be met for DEA, as stated by the field's founders and many of its most respected researchers. Further, using regression with DEA has become a very common technique and, again, the field's top researchers have led the way in developing applications.

1. The authors criticize previous studies for using "beds" as a proxy for capital, and then use "beds" in their own example! The use of "beds" has been empirically validated as a proxy for capital (See Ozcan, 1993). If "beds" is not a valid input variable, then what other variables do the authors advise researchers to use?

We don't think that we criticized the use of beds as a proxy for capital. Here is all that we said in the paper:

Using different proxies for capital and labor *might* result in a different conclusion, but we have used the proxies common to most published hospital DEA articles.

We thank the reviewer for the Ozcan (1993) citation, and, as a result, were able to find a more recent citation by Ozcan *et al.* in which they update and re-justify the use of beds as a proxy for capital. We have replaced the sentence above with the sentence below:

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 (Sikka, Luke and Ozcan, 2009).

2. The authors applied their model to data from 87 hospitals and found that they are not substitutable. From this they conclude that all previous DEA studies using these variables are invalid.

3. The authors may consider framing "FPR" as an addition to the DEA practitioner's toolkit, rather than concluding that all previous studies using DEA were invalid.

We agree with the reviewer – our statements about other studies were too broad, and we have corrected the problem. Also, we have reframed the FPR as the reviewer suggested. Here is what we now say:

[ Last paragraph on page 5]

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.

[Pages 14-16]

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, (Pedhazur and Schmelkin, 1991), or a variation of two-stage regression analysis (Greene, 2008; Wooldridge, 2002). There undoubtedly are other applicable statistical methodologies.

## References

1. Banker R D, Charnes A and Cooper W W (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Manag Sci* 30 (9): 1078-1092.
2. Banker R D and Natarajan R (2004). Statistical tests based on DEA efficiency scores. In: Cooper W W, Seiford L M and Zhu J (eds). *Handbook on Data Envelopment Analysis*. Kluwer: Boston, pp 299-322.
3. Barnum D T, Gleason J M, Hemily B, Lin J and Wang P (2009a). Progressing from uncertainty to risk for DEA-based decisions. *J Oper Res Soc*. In Press doi:10.1057/jors.2009.120.
4. Barnum D T, Shields K L, Walton S M and Schumock G T (2009b). Improving the efficiency of distributive and clinical services in hospital pharmacy with DEA. *J Med Syst*. In Press <http://dx.doi.org/10.1007/s10916-009-9341-2>.
5. Charnes A, Cooper W W and Rhodes E (1978). Measuring the efficiency of decision making units. *Eur J Oper Res* 2 (6): 429-444.
6. Cherchye L, Moesen W, Rogge N, Puyenbroeck T V, Saisana M, Saltelli A, Liska R and Tarantola S (2008). Creating composite indicators with DEA and robustness analysis: The case of the Technology Achievement Index. *J Oper Res Soc* 59 (2): 239.
7. Cooper W W, Deng H, Huang Z and Li S X (2002). Chance constrained programming approaches to technical efficiencies and inefficiencies in stochastic data envelopment analysis. *J Oper Res Soc* 53 (12): 1347.
8. Cooper W W, Seiford L M, Tone K and Zhu J (2007). Some models and measures for evaluating performances with DEA: past accomplishments and future prospects. *J Prod Anal* 28 (3): 151.
9. Cooper W W, Li S, Seiford L M and Zhu J (2004). Sensitivity analysis in DEA. In: Cooper W W, Seiford L M and Zhu J (eds). *Handbook on Data Envelopment Analysis*. Kluwer Academic Publishers: Boston, pp 75-97.
10. Cooper W W, Seiford L M and Zhu J (2004). Data envelopment analysis: History, models and interpretations. In: Cooper W W, Seiford L M and Zhu J (eds). *Handbook on Data Envelopment Analysis*. Kluwer Academic Publishers: Boston, pp 1-40.
11. Färe R, Grosskopf S and Lovell C A K (1994). *Production Frontiers*. Cambridge University Press: Cambridge, England.
12. Farrell M J (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society, Series A* 120 (3): 253-290.
13. Fried H O, Lovell C A K and Schmidt S S (2008). *The Measurement of Productive Efficiency and Productivity Growth*. Oxford University Press: Oxford.
14. Greene W H (2008). *Econometric Analysis*, 6th Ed. Prentice Hall: Upper Saddle River, NJ.

15. Grosskopf S (1996). Statistical inference and nonparametric efficiency: A selective survey. *J Prod Anal* 7 (2-3): 161-176.
16. Lovell C A K, Grosskopf S, Ley E, Pastor J T, Prior D and Vanden Eeckaut P (1994). Linear programming approaches to the measurement and analysis of productive efficiency. *TOP* 2 (2): 175-248.
17. Neralic L and Wendell R E (2004). Sensitivity in data envelopment analysis using an approximate inverse matrix. *J Oper Res Soc* 55 (11): 1187.
18. Pedhazur E J and Schmelkin L P (1991). *Measurement, Design, and Analysis: An Integrated Approach*. Lawrence Erlbaum Associates: Hillsdale, NJ.
19. Ruggiero J (2004). Data envelopment analysis with stochastic data. *J Oper Res Soc* 55 (9): 1008-1012.
20. Seiford L M (1996). Data envelopment analysis: The evolution of the state of the art (1978-1995). *J Prod Anal* 7 (2-3): 99-137.
21. Sikka V, Luke R D and Ozcan Y A (2009). The efficiency of hospital-based clusters: Evaluating system performance using data envelopment analysis. *Health Care Manage Rev* 34 (3): 251-261.
22. Simar L and Wilson P W (2000). Statistical inference in nonparametric frontier models: The state of the art. *J Prod Anal* 13 (1): 49-78.
23. Simar L and Wilson P W (2007). Estimation and inference in two-stage, semi-parametric models of production processes. *J Econometrics* 136 (1): 31-64.
24. Tser-yieth C (2002). A comparison of chance-constrained DEA and stochastic frontier analysis: Bank efficiency in Taiwan. *J Oper Res Soc* 53 (5): 492.
25. Wooldridge J M (2002). *Econometric Analysis of Cross Section and Panel Data*. MIT Press: Cambridge, MA.