Journal of Productivity Analysis, 17, 157–174, 2002

Accounting for Environmental Effects and Statistical Noise in Data Envelopment Analysis

H. O. FRIED friedh@union.edu
Department of Economics, Union College, Schenectady, NY 12308, USA

C. A. K. LOVELL knox@terry.uga.edu
Department of Economics, University of Georgia, Athens, GA 30602, USA

S. S. SCHMIDT schmidts@union.edu
Department of Economics, Union College, Schenectady, NY 12308, USA

S. YAISAWARNG yaisawas@union.edu
Department of Economics, Union College, Schenectady, NY 12308, USA

Abstract

In this paper we propose a new technique for incorporating environmental effects and statistical noise into a producer performance evaluation based on data envelopment analysis (DEA). The technique involves a three-stage analysis. In the first stage, DEA is applied to outputs and inputs only, to obtain initial measures of producer performance. In the second stage, stochastic frontier analysis (SFA) is used to regress first stage performance measures against a set of environmental variables. This provides, for each input or output (depending on the orientation of the first stage DEA model), a three-way decomposition of the variation in performance into a part attributable to environmental effects, a part attributable to managerial inefficiency, and a part attributable to statistical noise. In the third stage, either inputs or outputs (again depending on the orientation of the first stage DEA model) are adjusted to account for the impact of the environmental effects and the statistical noise uncovered in the second stage, and DEA is used to re-evaluate producer performance. Throughout the analysis emphasis is placed on slacks, rather than on radial efficiency scores, as appropriate measures of producer performance. An application to nursing homes is provided to illustrate the power of the three-stage methodology.

JEL classification: D24, C44, C53

Keywords: environmental effects, managerial inefficiency, statistical noise, slacks

1. Introduction

It is commonplace to evaluate the performance of producers on the basis of their inputs and outputs. Producers are typically evaluated in terms of their ability to minimize input usage in the production of given outputs, or to maximize output production with given inputs, relative to the performance of other producers in some comparison set. Data envelopment analysis (DEA) (Charnes et al., 1978) is a widely used linear programming technique for conducting such an evaluation. However producer performance is influenced by three very different phenomena: the efficiency with which management organizes production activities, the characteristics of the environment in which production activities are carried out, and the impact of good and bad luck, omitted variables, and related phenomena which would be collected in a random error term in a regression-based evaluation of producer performance. The first phenomenon is endogenous, while the second and third are exogenous. It is of course desirable to disentangle the three influences on producer performance. This requires information concerning environmental characteristics, as well as data on outputs and inputs, and the development of a model that incorporates the environmental characteristics into the performance evaluation procedure. In addition, in order to capture the influence of luck on producer performance, the model must be stochastic. However most DEA models, and virtually all operational DEA models, are deterministic. The objective of this paper is to develop a DEA-based model of producer performance that contains a stochastic element designed to isolate the impact of luck from those of managerial performance and environmental impacts.[1]

Several models have been proposed to incorporate environmental effects into a DEA-based evaluation of producer performance. These models can be grouped somewhat loosely into one-stage models and two-stage models. One-stage models use data on outputs, inputs and observable environmental variables all at once, the objective being to control for observable environmental variables in the evaluation of producer performance. However these models are deterministic, and so fail to account for the effect of statistical noise. Their implementation also requires that the direction (if not the magnitude) of each included environmental effect be known in advance. Two-stage models use data on outputs and inputs in the first stage, and use data on observable environmental variables in the second stage, the objective being to determine the impact of the observable environmental variables on initial evaluations of producer performance. If the second stage is DEA-based, the resulting two-stage model is fully deterministic, and incapable of accounting for the effect of statistical noise on producer performance. However if the second stage is regression-based, this model is capable of attributing some portion of the variation in producer performance to the effect of statistical noise, although this important feature of a two-stage model is typically not exploited. The three-stage model developed in this paper is an extension of the regression-based two-stage model, the extension allowing for the first time a complete decomposition of the three sources of variation in producer performance.

The paper is organized as follows. In Section 2 we briefly review the extant approaches to incorporating environmental effects into a DEA-based producer performance evaluation. In Section 3 we introduce our new three-stage producer performance evaluation model. Our procedure begins with a first stage DEA-based initial evaluation of producer performance, based exclusively on output and input data. In the second stage, producer performance evaluations are decomposed into environmental effects, managerial inefficiencies, and statistical noise. We use stochastic frontier analysis (SFA) (Aigner et al., 1977; Meeusen and van den Broeck, 1977) to implement the decomposition. SFA is regression-based, and so has the virtue of being capable of isolating managerial inefficiency from both environmental effects and statistical noise, although it has the drawback of doing so within a parametric framework. In the third stage, producers' inputs or outputs (depending on the orientation of the performance evaluation) are adjusted to account for the environmental effects and the statistical noise identified in the second stage, and DEA is used again to re-evaluate producer performance. The re-evaluation provides improved measures of managerial efficiency, since environmental effects and statistical noise have been netted out of producers' input or output data in the second stage.

In Section 4 we provide an empirical example to illustrate the power of our new three-stage producer performance evaluation model. The example concerns the operations of a sample of 990 US hospital-affiliated nursing homes previously analyzed by Fried et al. (1999). When we implement our three-stage procedure for incorporating the separate impacts of environmental variation and statistical noise, we find no statistically significant correlation between our third-stage and first-stage performance evaluations. This finding provides compelling evidence of the importance of incorporating these two exogenous phenomena into a DEA-based evaluation of producer performance. Section 5 concludes with some summary observations.

2. Existing Approaches to Incorporating Environmental Effects in DEA

Single-stage approaches were developed by Banker and Morey (1986a, 1986b) for non-discretionary environmental variables (such as quasi-fixed inputs and/or outputs whose magnitudes are temporarily constrained by contractual arrangements), and also for categorical environmental variables (such as ownership form). The approach to non-discretionary variables is to include them along with the inputs and outputs, but to restrict the optimization to either inputs or outputs. An obvious requirement is that the direction of the impact on producer performance of each non-discretionary variable must be known in advance. The approach to categorical variables is to restrict the comparison set to other producers in the same or higher (or the same or lower) categories. This of course requires that the categories be nested, and reduces the size of the comparison set for most producers, thereby reducing the discriminatory power of the model. Both approaches are purely deterministic, and so are incapable of incorporating the effect of statistical noise on producer performance. More detailed commentary on the two single-stage approaches is provided by Adolphson et al. (1991).[2]

The typical two-stage approach follows a first stage DEA exercise based on inputs and outputs with a second stage regression analysis seeking to explain variation in first stage efficiency scores in terms of a vector of observable environmental variables. Timmer (1971) pioneered this approach, and several subsequent studies have improved upon Timmer's second stage by using limited dependent variable regression techniques (since efficiency scores are bounded, and frequently achieve their upper bound).[3] McCarty and Yaisawarng (1993) and Bhattacharyya et al. (1997) went a step further, by using the second stage regression residuals to adjust the first stage efficiency scores. Fried et al. (1993) expanded the focus of the second stage regression analysis by applying the second stage regression to first stage slacks (rather than to radial efficiency scores), using seemingly unrelated regression techniques.

Pastor (1995) proposed a novel variation on the two-stage approach by proposing a double DEA format. In the first stage he applied either input-oriented DEA to inputs and the environmental variables or output-oriented DEA to outputs and the environmental variables. He then replaced either the inputs or the outputs by their radial projections, in order to eliminate the effect of the environmental variables. In the second stage he again applied DEA to an expanded data set consisting of the originally efficient observations, the originally inefficient observations, and the radial projections of the originally inefficient observations. A comparison of the second stage efficiency scores of the originally inefficient observations with those of the radial projections of the originally inefficient observations reveals the impact of the environmental variables on producer performance.

It is also possible to extend the basic two-stage approach, as Fried et al. (1999) have done. In their approach an initial DEA evaluation is followed by a second stage tobit regression analysis to obtain predictions of the impacts of the environmental variables on the first stage performance evaluations. In the third stage, the original data are adjusted to account for these environmental impacts, and the DEA evaluation is repeated. The virtue of this approach is that the second stage is stochastic. The shortcoming of this approach is that the data adjustment accounts for environmental impacts, but not for the impact of statistical noise.

3. The Three-Stage DEA Model

Our approach to purging producer performance evaluation of environmental effects and statistical noise is a three-stage approach. In the first stage we apply DEA to input and output data to obtain an initial evaluation of producer performance. This evaluation does not account for the impacts of either the operating environment or statistical noise on producer performance. Therefore in the second stage we use SFA to attribute variation in first stage producer performance to environmental effects, managerial inefficiency, and statistical noise. In the third stage, we adjust producers' inputs or outputs in a manner that accounts for the environmental effects and the statistical noise uncovered in the second stage. We then repeat the first stage analysis by applying DEA to the adjusted data. This third stage re-evaluation of producer performance provides improved measures of managerial efficiency, since the data have been purged of both environmental effects and statistical noise in the second stage SFA regression.[4]

Stage 1: The Initial DEA Producer Performance Evaluation

The initial producer performance evaluation is conducted using a conventional oriented DEA analysis, using input quantity data and output quantity data only. Either orientation is appropriate, and we arbitrarily adopt an input orientation. For producer "o" the Banker et al. (1984) variable returns to scale envelopment problem can be expressed as the linear programming problem:

$$\begin{aligned}
\min_{\theta,\lambda} \quad & \theta \\
\text{subject to} \quad & \theta x^{o} \geq X\lambda \\
& Y\lambda \geq y^{o} \\
& \lambda \geq 0 \\
& e^{T}\lambda = 1,
\end{aligned} \tag{1}$$

where $x \geq 0$ is a producer's $N \times 1$ vector of inputs, $y \geq 0$ is a producer's $M \times 1$ vector of outputs, $X = [x_1, \ldots, x_I]$ is an $N \times I$ matrix of input vectors in the comparison set, $Y = [y_1, \ldots, y_I]$ is an $M \times I$ matrix of output vectors in the comparison set, $\lambda = [\lambda_1, \ldots, \lambda_I]$ is an $I \times 1$ vector of intensity variables, $e = [1, \ldots, 1]$ is an $I \times 1$ vector, and there are I producers in the comparison set. The data of the producer being evaluated are superscripted "^{o}", and the problem is solved I times, once for each producer in the comparison set.

The optimal solutions to the envelopment problem (1) provide initial performance evaluations for each producer, expressed in terms of the optimal values of $\theta \leq 1$ and the nonnegative slacks in the $M + N$ functional constraints. However actual performances are likely to be attributable to some combination of managerial inefficiencies, environmental effects, and statistical noise, and it is desirable to isolate the three effects. This cannot be accomplished within the framework of the above problem, from which the environmental variables and statistical noise are both missing.

Stage 2: Using SFA to Decompose Stage 1 Slacks

We focus on *total* (radial plus nonradial) Stage 1 slacks $[x - X\lambda] \geq 0$ and $[Y\lambda - y] \geq 0$, in the belief that both types of slack reflect initial managerial inefficiency. However we interpret these slacks more broadly, as being composed of three effects: environmental influences, managerial inefficiencies, and statistical noise arising from measurement errors in the input and output data used to generate the first stage slacks. The objective of the Stage 2 analysis is to decompose Stage 1 slacks into these three effects. We can meet this objective only by using SFA, in which Stage 1 slacks are regressed against observable environmental variables and a composed error term, which both captures and distinguishes the effects of managerial inefficiency and statistical noise. The overriding virtue of using SFA (rather than a limited dependent variable approach such as tobit) in Stage 2 is that its error term is asymmetric. Consequently it allows for the impact on Stage 1 slacks of environmental variables (the regressors), of managerial inefficiency[5](the one-sided error component) and of statistical noise (the symmetric error component).

When using SFA to regress Stage 1 slacks against the observable environmental variables, we have two pairs of options, or four options in all. The first pair of options concerns whether to attempt to explain variation in all $M + N$ Stage 1 slacks, or in just the N Stage 1 input slacks. A case could be made for either approach, but since our Stage 1 model is input-oriented, we prefer to focus on just the N Stage 1 input slacks. The second pair of options concerns whether to estimate N separate SFA regressions, one for each Stage 1 input slack, or to stack the N regressions and estimate a single SFA regression model. The advantage of the first approach is that it allows environmental variables to have different effects on each Stage 1 input slack. The advantage of the second approach is that it provides for greater degrees of freedom, and greater statistical efficiency in estimation. Again a case could be made for either approach, but we believe that the gain in flexibility offered by the former approach outweighs the sacrifice of degrees of freedom. Hence we opt for the estimation of N separate SFA regressions.

The dependent variables in the Stage 2 SFA regression models are the Stage 1 *total* input slacks $s_{ni} = x_{ni} - X_n\lambda \geq 0, n = 1, \ldots, N, i = 1, \ldots, I$, where s_{ni} is the Stage 1 slack in the usage of the nth input for the ith producer, X_n is the nth row of X, and $X_n\lambda$ is the optimal projection of x_{ni} onto the input efficient subset for output vector y_i. The independent variables in the Stage 2 SFA regression model are the elements of the K observable environmental variables $z_i = [z_{1i}, \ldots, z_{Ki}], i = 1, \ldots, I$. The N separate Stage 2 SFA regressions take the general form

$$s_{ni} = f^n(z_i; \beta^n) + v_{ni} + u_{ni}, \quad n = 1, \ldots, N, \;\; i = 1, \ldots, I, \tag{2}$$

where the $f^n(z_i; \beta^n)$ are deterministic feasible slack frontiers with parameter vectors β^n to be estimated and composed error structure $(v_{ni} + u_{ni})$. Consistent with a stochastic cost frontier formulation, we assume that the $v_{ni} \sim N(0, \sigma_{vn}^2)$ reflect statistical noise and that the $u_{ni} \geq 0$ reflect managerial inefficiency. If we make a distributional assumption on the u_{ni}, such as $u_{ni} \sim N^+(\mu^n, \sigma_{un}^2)$, and if we assume that the v_{ni} and the u_{ni} are distributed independently of each other, and of the z_i, each of the N regressions (2) may be estimated by maximum likelihood techniques. In each regression the parameters to be estimated are $(\beta^n, \mu^n, \sigma_{vn}^2, \sigma_{un}^2)$. All parameters are allowed to vary across the N input slack regressions, which allows the environmental variables, statistical noise and managerial inefficiency each to exert different impacts across inputs.[6]

The SFA regression models (2) are interpreted in the following way. The impact of the environment on Stage 1 slacks is captured by the deterministic feasible slack frontiers $f^n(z_i; \beta^n)$. However this relationship is noisy, and so the stochastic feasible slack frontiers (SFSF) are $[f^n(z_i; \beta^n) + v_{ni}]$. Since $u_{ni} \geq 0$, the SFSF represent the *minimum* slacks that can be achieved in a noisy environment characterized by variables (z_i, v_{ni}) and parameters (β^n, σ_{vn}^2). Any slacks in excess of the SFSF are attributable to managerial inefficiency, because the effects of both the environmental variables z_i and statistical noise v_{ni} are netted out, having been captured by the SFSF. The managerial inefficiency component of the slacks is captured by the nonnegative error components u_{ni}, with parameters (μ^n, σ_{un}^2) reflecting the variability of managerial inefficiency, both across producers and across inputs.

The Stage 2 SFA regression model has four virtues. First, it is not necessary to know the direction of the impact of any environmental variable on producer performance prior to the analysis. In fact, the directions and magnitudes of the impacts can vary across inputs, and are determined from the estimated regression parameters. Second, the statistical significance of the effects of the environmental variables, both individually and collectively, can be determined using conventional likelihood ratio tests. This enables a test of the hypothesis that environmental variation has no impact on producer performance as evaluated by the Stage 1 DEA analysis. Third, the hypothesis that managerial efficiency is invariant across producers can be evaluated by testing the hypothesis that $\sigma_{un}^2 = 0$, in which case producer performance variation is attributable exclusively to variation in the environment and to the effects of statistical noise. Fourth, the framework permits the environmental variables, statistical noise and managerial inefficiency each to exert different impacts across inputs.

We now consider how to utilize the results of the Stage 2 SFA analysis to adjust producers' inputs for the variable impacts of different operating environments and random statistical noise. The objective of the proposed adjustment is to level the playing field before repeating the DEA analysis. The essence of the proposed adjustment exploits the fact that producers operating in relatively unfavorable environments, and producers experiencing relatively bad luck, are disadvantaged in the Stage 1 DEA performance evaluation that does not take these factors into account. One way to level the playing field is to adjust downward the inputs of these producers in amounts determined by the extent to which they have been disadvantaged by their relatively unfavorable environments or by their relatively bad luck. The extent to which they have been disadvantaged by each source is revealed by the parameter estimates obtained in the Stage 2 SFA regressions. An alternative procedure is to adjust upward the inputs of producers who have been advantaged by their relatively favorable operating environments or by their relatively good luck. We adopt the second approach, to avoid the possibility that some extremely disadvantaged producers might have some inputs adjusted so far downward as to become negative.

Producers' adjusted inputs are constructed from the results of the Stage 2 SFA regressions by means of

$$x_{ni}^A = x_{ni} + [\max_i\{z_i\hat{\beta}^n\} - z_i\hat{\beta}^n] + [\max_i\{\hat{v}_{ni}\} - \hat{v}_{ni}], \quad n = 1, \ldots, N, \quad i = 1, \ldots, I, \quad (3)$$

where x_{ni}^A and x_{ni} are adjusted and observed input quantities, respectively.[7] The first adjustment on the right side of equation (3) puts all producers into a common operating environment, the least favorable environment observed in the sample. The second adjustment puts all producers into a common state of nature, the unluckiest situation encountered in the sample. Thus producers with relatively unfavorable operating environments and/or relatively bad luck have their inputs adjusted upward by a relatively small amount, while producers with relatively favorable operating environments and/or relatively good luck have their inputs adjusted upward by a relatively large amount. These adjustments vary both across producers and across inputs.

In order to implement equation (3) it is necessary to separate statistical noise from managerial inefficiency in the residuals of the SFA regression models (2) in order to obtain estimates of v_{ni} for each producer. This is accomplished by using the Jondrow et al. (1982) methodology that decomposes the composed error terms in equation (2). From the conditional estimators for managerial inefficiency given by $\hat{E}[u_{ni}|v_{ni}+u_{ni}]$, we derive estimators for statistical noise residually by means of

$$\hat{E}[v_{ni}|v_{ni} + u_{ni}] = s_{ni} - z_i\hat{\beta}^n - \hat{E}[u_{ni}|v_{ni} + u_{ni}], \quad n = 1, \ldots, N, \quad i = 1, \ldots, I, \quad (4)$$

which provide conditional (on $v_{ni} + u_{ni}$) estimators for the v_{ni} in equation (3). Since the $\hat{E}[u_{ni}|v_{ni} + u_{ni}]$ depend on $(\hat{\beta}^n, \hat{\sigma}_{vn}^2, \hat{\sigma}_{un}^2, \hat{\mu}^n)$, so do the $\hat{E}[v_{ni}|v_{ni} + u_{ni}]$. The elements of $\hat{\beta}^n$ provide estimates of the contributions of each observable environmental variable to slack in usage of the nth input, while the parameters $(\sigma_{un}^2, \sigma_{vn}^2, \mu^n)$ characterize the separate contributions of managerial inefficiency and statistical noise to slack in usage of the nth input. In particular, as $\gamma^n = \sigma_{un}^2/(\sigma_{vn}^2 + \sigma_{un}^2) \to 1$, the impact of managerial inefficiency dominates that of statistical noise in the determination of slack in usage of the nth input, while just the opposite occurs as $\sigma_{un}^2/(\sigma_{vn}^2 + \sigma_{un}^2) \to 0$.[8]

Stage 3: Adjusted DEA

Stage 3 is a repetition of Stage 1, with observed input data x_{ni} replaced with input data x_{ni}^A which have been adjusted for the impacts of both the observable environmental variables and statistical noise. The output of Stage 3 is a DEA-based evaluation of producer performance couched solely in terms of managerial efficiency, purged of the effects of the operating environment and statistical noise.[9]

4. An Empirical Application

We apply our three-stage methodology to a 1993 sample of 990 US hospital-affiliated nursing homes. The data are described in detail by Fried et al. (1999), and summary statistics are presented in Table 1. Two outputs are specified: inpatient days of skilled care (SKD), and inpatient days of intermediate care (ICD). Four inputs are specified: registered nurses (RN), licensed practical nurses (LPN), other personnel (OEMP), and non-payroll expenses (NEXP). Three types of environmental variable are specified: ownership form (for profit (F) and not-for-profit (NF)); location (attached to a hospital (A), on campus (ON), and off campus (OFF)); and facility capacity (four categories of bed size (BED1–BED4)). Including all possible ownership-location pairs generates six environmental dummy variables. Adding the four bed size dummy variables gives a total of ten environmental indicators. There are only two homes in the for-profit-off-campus category, and just four homes in the for-profit-on-campus category. These homes are removed from the data set, leaving eight environmental variables that are captured by two sets of dummy variables. The omitted dummies are not-for-profit-attached homes and the smallest bed size category. All three types of environmental variable are posited to influence nursing home performance, although without prior knowledge of the directions of their impacts, it is not possible to

Table 1. Descriptive statistics of hospital-affiliated nursing homes, 1993 (sample size = 990).

Variables	Mean	Std. Deviation	Minimum	Maximum
Outputs:				
SKD	13,162.1	15,356.7	0	166,517
ICD	2,876.2	8,184.9	0	64,970
Inputs:				
RN	6.3	5.5	0.5	43
LPN	7.4	6.1	0.5	47
OEMP	30.3	32.2	1.0	271
NEXP	912,513.9	1,046,032.1	4,774.0	8,395,873

use these variables in a single stage DEA performance evaluation. Hence our Stage 1 DEA analysis is conducted using only the two outputs and the four inputs.

The impacts of the environmental variables are investigated in our Stage 2 SFA analysis, after which we conduct our Stage 3 DEA analysis using adjusted input quantities.[10]An advantage of our approach is that it does not require that we impose priors on the directions of the impacts of environmental variables on managerial inefficiency. However, there are reasons to expect a for-profit environment to be more favorable if stockholders monitor managerial performance. On campus and attached locations are expected to provide more favorable environments than an off campus location, since these environments provide an opportunity to share resources with the associated hospital. Whether large bed size is related to the environment or not depends upon how well we controlled for scale effects in the VRS specification of the Stage 1 DEA model. Nonetheless it is important to include bed size as a control variable, since equally efficient small and large homes will have different levels of *absolute* slack.

Results of the Stage 1 DEA analysis are summarized in Table 6, in order to facilitate subsequent comparison with the Stage 3 DEA analysis. The Stage 1 DEA results suggest a relatively low mean efficiency relative to best practice, and a relatively large dispersion in performance. Although fully 86 homes form the best practice frontier, the mean radial efficiency score is just 0.52 with a standard deviation of 0.25. The natural inference to draw at this point is that there is considerable room for improvement in the sample, and that overall performance can be enhanced by bringing the laggards up to best practice standards. However managers of the laggard homes can be expected to protest that their initial performance evaluations are misleading, because of the unfavorable environments in which they operate, or because of otherwise unfavorable extenuating circumstances. It is the objective of the second and third stages of the analysis to investigate the validity of these claims, and to re-evaluate the performance of all homes in light of these phenomena.

Results of the Stage 2 SFA regressions are based on a half normal specification of the one-sided inefficiency error component, and are summarized in Table 2. These results suggest that the operating environment does indeed exert a statistically significant influence on nursing home performance. Between one and six of the environmental variables are significant, depending upon the input equation, and all six environmental variables have consistent signs across the four equations. Ownership form matters, with input slack being consistently smaller in for-profit homes than in not-for-profit homes, although the coefficient is significant only in the OEMP equation. Location matters also, with input slack being generally smaller in attached than on-campus and off-campus homes. Capacity appears to matter as well, with input slack being generally larger for the largest bed size category than for smaller bed size categories. There is clear evidence that excess input usage is related to the three sets of environmental variables. However the statistical significance of so many regressors does not imply that all managers of laggard homes have a legitimate complaint, since some of them may be operating in relatively favorable environments. We return to this issue below.[11]

The results summarized in Table 2 also suggest that a small but statistically significant portion of input slack is attributable to managerial inefficiency. A likelihood ratio test rejects the hypothesis that the one-sided error component makes no contribution to the composed

Table 2. Stochastic frontier estimation results (standard errors in parenthesis).

Independent Variable	Dependent Variable RN Slack	LPN Slack	OEMP Slack	NEXP Slack
Constant	3.067* (0.228)	3.160* (0.232)	9.570* (1.032)	353,513.04* (38,701.02)
For-profit attached, FA	−0.592 (0.485)	−0.490 (0.488)	−4.656* (2.159)	−50,021.49 (82,494.80)
Not-for-profit on-campus, NFON	0.711 (0.825)	1.740 (0.952)	12.293* (3.639)	585,049.38* (140,209.20)
Not-for-profit off-campus, NFOFF	0.353 (0.693)	0.934 (0.710)	11.896* (3.066)	166,993.41 (117,684.37)
Next to smallest, BED2	0.806 (0.320)	0.641* (0.322)	2.914* (1.435)	287.68 (54,536.76)
Next to largest, BED3	0.445 (0.430)	1.451* (0.435)	11.316* (1.918)	168,698.69* (73,683.33)
Largest, BED4	2.927* (0.431)	3.427* (0.439)	21.732* (1.965)	452,634.53* (73,398.17)
σ^2	16.90* (.761)	17.02* (0.786)	330.16* (14.865)	692,875.46* (16,542.57)
γ	0.591E-04* (0.207E-04)	0.2999E-05 (0.915E-05)	0.232E-02* (0.115E-03)	N/A
Log-likelihood function	−2,804.2	−2,807.7	−4,274.8	−13,525.7
LR test of the one-sided error	0.8345E-04	0.7142E-06	0.1857E-01	

*Significant at the 5% level or better.

error term, except for the NEXP equation, in which the stochastic frontier specification is rejected. This implies that variation in managerial inefficiency plays no role in the use of this input. Consequently the parameter estimates for NEXP reported in Table 2 are based upon a tobit specification.

Thus, after controlling for variation in the operating environment, there remains very little variation in managerial efficiency in the use of each input. This is apparent from Table 3, which reports mean input-specific efficiency scores by ownership-location-size category.[12] All categories manage non-payroll operating expense equally efficiently, since the frontier specification is rejected for this equation. For the remaining three inputs, inefficiency ranges from slightly greater than zero to around nine percent. There are no obvious patterns across ownership-location-bed categories.

The results summarized in Table 2 also shed light on the contribution of statistical noise to nursing home performance. The estimated values of the parameters γ^n are very small for the three labor inputs. This suggests that the environmental variables and statistical noise explain virtually all of the variation in slack in the three labor inputs, and that there is very little difference in the abilities of individual managers to adapt to the external environment. Variation in managerial efficiency matters in a statistical sense in the case of two inputs, but its economic impact is negligible.

Table 3. Mean input-specific efficiency scores (Stage 2) categorized by operating environment.

			Mean Efficiency Score by Category			
Ownership	Location	#	RN	LPN	OEMP	NEXP
Smallest size						
For profit	Attached	77	1.007	1.001	1.049	Na
Not for profit	Attached	314	1.007	1.001	1.056	Na
	On campus	2	1.008	1.001	1.056	Na
	Off campus	2	1.006	1.001	1.028	Na
This size class		395	1.007	1.001	1.054	Na
Next to smallest size						
For profit	Attached	6	1.008	1.002	1.088	Na
Not for profit	Attached	304	1.008	1.002	1.058	Na
	On campus	5	1.009	1.002	1.093	Na
	Off campus	7	1.009	1.002	1.088	Na
This size class		322	1.008	1.002	1.060	Na
Next to largest size						
For profit	Attached	1	1.008	1.002	1.077	Na
Not for profit	Attached	113	1.008	1.002	1.079	Na
	On campus	5	1.007	1.001	1.048	Na
	Off campus	7	1.007	1.001	1.049	Na
This size class		126	1.008	1.002	1.076	Na
Largest size						
For profit	Attached	3	1.007	1.001	1.044	Na
Not for profit	Attached	105	1.008	1.002	1.061	Na
	On campus	15	1.007	1.001	1.052	Na
	Off campus	24	1.007	1.001	1.054	Na
This size class		147	1.008	1.002	1.059	Na
Overall		990	1.008	1.002	1.060	Na

The parameter estimates obtained from the Stage 2 SFA regressions are used to predict slacks attributable to the operating environment and to noise for all size-location-ownership categories. Table 4 contains the means of predicted slacks attributable to the operating environment. All of the means are positive. This corresponds to an unfavorable environment on average. The most favorable operating environment for all four inputs appears to be for-profit attached, for all bed size categories. The least favorable environment appears to be not-for-profit, on campus. The smallest bed size category for all location-ownership combinations appears to be a more favorable environment than larger bed-size categories.

Table 5 contains the means of predicted slacks due to noise for all size-location-ownership combinations. For the three labor inputs the estimated value of γ is very small, and variation in predicted slack reflects primarily the effect of noise. Since the frontier specification is rejected for non-payroll operating expense, there is no evidence that variation in efficiency matters, and variation in predicted slack is due exclusively to noise. For the RN and OEMP

Table 4. Mean and maximum estimated slacks (Stage 2) categorized by operating environment.

			Mean Estimated Slack by Category			
Ownership	Location	#	RN	LPN	OEMP	NEXP
Smallest size						
For profit	Attached	77	2.475	2.670	4.915	303491.55
Not for profit	Attached	314	3.067	3.160	9.570	353513.04
	On campus	2	3.777	4.900	21.863	938562.42
	Off campus	2	3.420	4.094	21.467	520506.45
Maximum			3.777	4.900	21.863	938562.42
Next to smallest size						
For profit	Attached	6	2.556	3.311	7.829	303779.23
Not for profit	Attached	304	3.147	3.801	12.484	353800.72
	On campus	5	3.858	5.541	24.777	938850.10
	Off campus	7	3.500	4.735	24.381	520794.13
Maximum			3.858	5.541	24.777	938850.10
Next to largest size						
For profit	Attached	1	2.920	4.121	16.231	472190.24
Not for profit	Attached	113	3.511	4.611	20.886	522211.73
	On campus	5	4.222	6.351	33.179	1107261.11
	Off campus	7	3.864	5.545	32.782	689205.14
Maximum			4.222	6.351	33.179	1107261.11
Largest size						
For profit	Attached	3	5.402	6.097	26.646	756126.08
Not for profit	Attached	105	5.993	6.587	31.302	806147.57
	On campus	15	6.704	8.327	43.595	1391196.95
	Off campus	24	6.346	7.521	43.198	973140.98
Maximum			6.704	8.327	43.595	1391196.95
Overall maximum			6.704	8.327	43.595	1391196.95

equations, the maximum mean values for noise occur in the largest bed size and for-profit-attached category. For the NEXP equation, the maximum mean value of noise occurs in the smallest bed size and not-for-profit on-campus category. These are unlucky environments. For the LPN input, the unluckiest environment is the next to smallest bed size, not-for-profit-off-campus. The luckiest environment is different for each input.

Observed input usage is adjusted for the influences of the environment and noise, by inserting the parameter estimates obtained from the Stage 2 regressions into equation (3).[13] This procedure reflects the impact of variation in the operating environment and statistical noise. It does not reflect variation in efficiency in managing the external environment, which is negligible or zero for each input. The Stage 3 DEA analysis is based on observed outputs and adjusted inputs. Results are summarized in Tables 6 and 7.

The Stage 3 DEA efficiency scores summarized in Table 6 suggest that, after adjusting for variation in the operating environment and for the influence of statistical noise, mean

Table 5. Mean and maximum estimated noise (Stage 2) categorized by operating environment.

			Mean Estimated Noise by Category			
Ownership	Location	#	RN	LPN	OEMP	NEXP
Smallest size						
For profit	Attached	77	−0.149	0.076	0.389	34094.70
Not for profit	Attached	314	0.034	−0.010	0.348	21321.66
	On campus	2	1.562	−1.032	−11.626	692299.15
	Off campus	2	−0.439	−0.073	−15.417	−17639.36
Maximum			17.779	17.103	97.534	2190676.59
Next to smallest size						
For profit	Attached	6	0.395	−0.013	−3.613	−17183.13
Not for profit	Attached	304	−0.024	−0.039	−0.083	44745.95
	On campus	5	0.098	−0.659	−6.638	−616612.48
	Off campus	7	0.684	2.172	7.859	12515.80
Maximum			23.920	27.274	136.742	6153855.64
Next to largest size						
For profit	Attached	1	−2.084	0.062	−8.774	−292075.96
Not for profit	Attached	113	−0.068	−0.002	−1.349	19610.11
	On campus	5	2.362	1.784	10.123	599889.38
	Off campus	7	−0.370	−1.289	−1.276	20011.59
Maximum			22.531	21.762	130.617	4489541.06
Largest size						
For profit	Attached	3	3.916	−1.901	11.299	16494.93
Not for profit	Attached	105	0.027	0.142	−1.182	72903.81
	On campus	15	−1.055	−0.246	−2.109	−11425.00
	Off campus	24	−0.077	−0.258	−2.935	33349.85
Maximum			29.307	23.911	87.669	4028179.53

Table 6. Comparison of DEA initial and final results.

	Initial Results	Final Results
Mean efficiency score	0.522	0.905
Standard deviation	0.247	0.096
Minimum	0.054	0.707
Maximum	1.000	1.000
Number of efficient nursing homes	86	222

Kendall rank correlation coefficient between initial and final results = −0.00339.
The 95% confidence interval is (−0.08437, 0.08391).

efficiency scores improve dramatically and dispersion declines. This is consistent with the hypothesis that at least some nursing homes that received relatively high initial performance evaluations did so in part due to their relatively favorable operating environments or their relatively favorable extenuating circumstances. Not all of them were as well managed as

Table 7. Mean rank for initial and final results categorized by operating environment (ranked from 1 = most efficient to 990 = least efficient).

Ownership	Location	#	Mean Rank by Category	
			Initial Stage	Final Stage
Smallest size				
For profit	Attached	77	553	163
Not for profit	Attached	314	545	223
	On campus	2	847	409
	Off campus	2	734	350
Next to smallest size				
For profit	Attached	6	550	930
Not for profit	Attached	304	476	825
	On campus	5	470	671
	Off campus	7	546	777
Next to largest size				
For profit	Attached	1	227	521
Not for profit	Attached	113	447	563
	On campus	5	691	688
	Off campus	7	567	519
Largest size				
For profit	Attached	3	232	524
Not for profit	Attached	105	390	465
	On campus	15	542	496
	Off campus	24	487	499

their initial evaluations indicated. It also lends support to the hypothesis that at least some nursing homes that received relatively low initial performance evaluations did indeed have a valid complaint, due to their relatively unfavorable operating environments or their relatively unfavorable extenuating circumstances. Not all of them were as poorly managed as their initial evaluations indicated. However if this was generally the case, the rank correlation between Stage 1 and Stage 3 DEA efficiency scores would be significantly negative. But the Kendall rank correlation coefficient[14] is not significantly different from zero. Thus nursing homes that received relatively high (low) initial performance evaluations did so in relatively favorable (unfavorable) operating environments and extenuating circumstances. Although adjusting performance evaluations for variation in the operating environment and for extenuating circumstances levels the playing field, it does not eliminate all variation in managerial performance.

The impact of the adjustments is examined in detail in Table 7, which reports arithmetic means of initial and final performance rankings of nursing homes in each category. In the initial stage the category receiving the highest mean efficiency rank is the next-to-largest bed size, having a for-profit status and an attached location. After adjusting for its relatively favorable operating environment and for its benign extenuating circumstances, the mean rank deteriorates from 227 to 521. At the other extreme, the mean rank of the category with the smallest bed size, having a not-for-profit status and an on-campus location, improves from 847 to 409 when account is taken of its relatively unfavorable operating environment and its relatively difficult extenuating circumstances. But not all categories experience such a dramatic change in their performance evaluations when adjustments are made. The relatively poor performance of some categories remains largely unaffected by the adjustments, suggesting that whatever relatively unfavorable circumstances they suffer are offset by other relatively favorable circumstances. Generally speaking, the rankings of the smallest nursing homes improve, while the rankings of the three larger size categories decline, when adjusted data are used in the performance evaluation. Small nursing homes have a relatively favorable operating environment, and this association outweighs (and therefore conceals) any association between either ownership form or location and the operating environment.

5. Summary and Conclusions

As it is conventionally employed, DEA has two drawbacks. First and foremost, it is deterministic, and so is plagued by measurement errors in included variables and by the omission of unobserved but potentially relevant variables, the impacts of which would be captured by a disturbance term in a stochastic model. Second, chief among the omitted variables are what we have referred to as environmental variables, those that capture features of the operating environment which are posited to have an impact on the efficiency with which conventional inputs are used to produce conventional outputs. These variables are typically omitted not because they are unobserved, but because the lack of prior knowledge of the direction of their impacts precludes their introduction in a single stage DEA analysis.

In this paper we have endeavored to address both of these drawbacks, by developing and implementing a three-stage DEA model. The first stage is a standard deterministic DEA model, based on conventional inputs and outputs. The second stage is regression-based, and so is stochastic. The objective of the second stage is to explain variation in first stage performance in terms of three phenomena: observable characteristics of the operating environment, statistical noise, and managerial inefficiency. The structure of the second stage is a set of SFA regressions, one for each input, in which Stage 1 input slacks are regressed against the environmental variables to construct stochastic feasible slack frontiers. The structure of these frontiers reflects the direction and the intensity of the impact of each environmental variable on slack in the use of each input. The structure of the disturbance terms associated with these frontiers apportions excess input slacks to statistical noise and managerial inefficiency. After the SFA regressions have been estimated, the original inputs are adjusted to account for the effects of variation in the operating environment and variation in statistical noise. These adjustments penalize producers operating in relatively favorable environments or enjoying otherwise benign extenuating circumstances. The third stage of the analysis is another standard deterministic DEA model, but this model is based on conventional outputs and adjusted inputs. This final performance evaluation is conducted on a level playing field, since variation in operating environments and the vagaries of luck have been accounted for in the second stage of the analysis.[15]

We have applied our three-stage methodology to the evaluation of the performance of a 1993 sample of 990 US nursing homes. These homes operate in very different environments, which we characterize by their ownership form, their location, and their capacity. Our second stage SFA analysis demonstrates that these environmental variables do indeed influence performance, as measured by input slacks obtained in the first stage DEA analysis, in a statistically significant way. Our second stage analysis also demonstrates that, even after controlling for the impacts of the environmental variables, managerial efficiency in the use of each input varies across homes. Finally our second stage analysis demonstrates the role played by statistical noise in the use of one of the inputs. When we adjust input use for the separate impacts of the operating environment and statistical noise, we find considerable differences between our Stage 1 DEA results and our Stage 3 DEA results. Our evaluation of the performance of various categories of nursing home is dramatically altered when we incorporate phenomena that are ignored in a conventional single stage DEA analysis.

Acknowledgments

We dedicate this paper to WWC, a longtime friend and influential scholar. Previous versions of this paper were presented at the Second Biennial Georgia Productivity Workshop and at the Oviedo Workshop on Efficiency and Productivity. We are grateful to a referee for posing several questions that led us to sharpen the econometric analysis.

Notes

1. We refer to DEA as a deterministic model because it has no explicit error term. Simar and Wilson (2000) define a statistical model from which they derive the asymptotic sampling distribution of the DEA efficiency scores in the bivariate case. They also discuss the use of bootstrapping to approximate sampling distributions of DEA efficiency scores in the multivariate case. They do not discuss the statistical properties of DEA slacks, which are the focus of our analysis.
2. Charnes et al. (1981) developed an extreme approach to the categorical variable problem, which generated a two-stage DEA model. In the first stage the data were divided into two mutually exclusive categories, DEA was performed separately on each category, and the data were projected to the separate category frontiers to eliminate managerial inefficiency within each category. In the second stage the managerially efficient data sets were merged, and DEA was performed a second time, the objective being to compare the performance of the two categories.
3. Note the date of Timmer's contribution. It appeared seven years prior to the formal introduction of DEA.
4. We originally conjectured that DEA could be used as an alternative to SFA in the second stage, to adjust Stage 1 producer performance evaluations for environmental effects (if not for the effects of statistical noise). We proposed a radial input-oriented second stage DEA model in which Stage 1 slacks were treated as "inputs" to be minimized, and environmental variables were treated as "outputs" which constrained the ability of producers to minimize slacks. However, this formulation failed to distinguish managerial inefficiency from environmental effects (much less from the effects of statistical noise), because producers who were Koopmans-efficient in the first stage could not be included in the second stage DEA model, since they had null slack vectors. For these producers it was not possible to determine whether their first stage efficiency was due to managerial efficiency or a favorable environment.
5. Some producers are able to cope with the external environment better than other producers, and some inputs are easier to adjust than others. The managerial inefficiency terms capture these differences, both across producers and across inputs.

6. The error components are probably not iid, since the slacks are derived from the first stage DEA model. Thus it might be preferable to stack the regressions and estimate via SUR, allowing error terms to be correlated across inputs. This would be feasible within a conventional error structure, but it would be difficult within a composed error framework. The idea suggests a new project, the merger of SFA with SUR, allowing either or both error components to be correlated across equations. We leave this project to others.
7. Note that we have replaced $f^n(z_i; \hat{\beta}^n)$ with $z_i\hat{\beta}^n$. We have imposed linearity on the feasible slack frontiers because a logarithmic specification would require $s_{ni} > 0$, which in turn would eliminate many producer/input observations having $s_{ni} = 0$ from the second stage SFA regressions.
8. The separate estimates of inefficiency and noise obtained from equation (2) are both *conditional* (on the residuals, which only estimate their sum, $v + u$). In a cross-section context, large N does not suffice for consistency, which is one reason Schmidt and Sickles (1984) favor the use of panel data when estimating SFA models. However our cross section is sufficiently large to give us confidence that β is estimated consistently, so that we have consistent estimates of the impacts of the environmental variables on input slacks.
9. Our second and third stages generate two distinct indicators of managerial performance. The first is a set of N predictors $\hat{E}[u \mid v + u]$ obtained in the second stage. The second is a single third stage DEA radial efficiency score. The former provide input-specific measures of management's ability to adapt to the vagaries of the external environment, and it may be easier to manage some inputs than others. The latter provides a comprehensive assessment of managerial performance, in terms of adapting to the environment and in terms of conserving primary input use. To obtain this measure requires that we adjust the data *and* re-run DEA, in order to evaluate all producers relative to environmentally adjusted best practice.
10. The Stage 1 DEA analysis allows for variable returns to scale in the four variable inputs. The inclusion of a fixed input (facility capacity, as proxied by bed size categories) in the Stage 2 SFA analysis permits us to test the hypothesis that facility capacity influences the efficiency of variable input usage. The four bed size categories are [3,30), [30,60), [60,90), [90,260].
11. Regressions (2) are linear regressions. The regressors are categorical variables, with not-for-profit attached homes and the smallest bed size serving as the benchmark categories. The parameter estimates reported in Table 2 show the impact on input slacks of moving from the benchmark category to any other category. For example, the impact of moving from an attached category to on-campus location while retaining not-for-profit status and holding bed-size constant would be to increase expected slack in non-payroll expense by \$585,049 or by about 64% on average.
12. In particular, note that there is very little variation in LPN-efficiency scores across nursing homes after controlling for the operating environment. This supports the insignificance of γ for LPN equation in Table 2.
13. We implement equation (3) by using maximum values that are specific to a bed-size category.
14. Stage 1 and Stage 3 efficiency scores contain a large number of ones that create ties in calculating the ranks. Ties are broken by calculating the number of times an efficient unit appears in the efficient reference set of an inefficient unit. For example, an efficient unit that is a peer more often for inefficient units is ranked higher than an efficient unit that is a peer less often.
15. The main advantages of our three-stage approach to the introduction of environmental variables and statistical noise into DEA are its computational simplicity and its modest data requirements relative to other approaches such as chance-constrained DEA and bootstrapped DEA. These alternative approaches are surveyed in Olesen, Petersen and Lovell (1996).

References

Adolphson, D. L., G. C. Cornia and L. C. Walters. (1991). "A Unified Framework for Classifying DEA Models." *Operational Research 90*, 647–657.

Aigner, D., C. A. K. Lovell and P. Schmidt. (1977). "Formulation and Estimation of Stochastic Frontier Production Function Models." *Journal of Econometrics* 6(1), 21–37.

Banker, R. D., A. Charnes and W. W. Cooper. (1984). "Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis." *Management Science* 30(9), 1078–1092.

Banker, R. D. and R. C. Morey. (1986a). "Efficiency Analysis for Exogenously Fixed Inputs and Outputs." *Operations Research* 34(4), 513–521.

Banker, R. D. and R. C. Morey. (1986b). “The Use of Categorical Variables in Data Envelopment Analysis.” *Management Science* 32(12), 1613–1627.

Bhattacharyya, A., C. A. K. Lovell and P. Sahay. (1997). “The Impact of Liberalization on the Productive Efficiency of Indian Commercial Banks.” *European Journal of Operational Research* 98(2), 332–347.

Charnes, A., W. W. Cooper and E. Rhodes. (1978). “Measuring the Efficiency of Decision Making Units.” *European Journal of Operational Research* 2(6), 429–444.

Charnes, A., W. W. Cooper and E. Rhodes. (1981). “Evaluating Program and Managerial Efficiency: An Application of Data Envelopment Analysis to Program Follow Through.” *Management Science* 27(6), 668–697.

Fried, H. O., C. A. K. Lovell and P. Vanden Eeckaut. (1993). “Evaluating the Performance of U.S. Credit Unions.” *Journal of Banking and Finance* 17(2/3), 251–265.

Fried, H. O., S. S. Schmidt and S. Yaisawarng. (1999). “Incorporating the Operating Environment into a Nonparametric Measure of Technical Efficiency.” *Journal of Productivity Analysis* 12(3), 249–267.

Jondrow, J., I. Materov, C. A. K. Lovell and P. Schmidt. (1982). “On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Model.” *Journal of Econometrics* 19(2/3), 233–238.

McCarty, T. and S. Yaisawarng. (1993). “Technical Efficiency in New Jersey School Districts.” In H. O. Fried, C. A. K. Lovell and S. S. Schmidt (eds.), *The Measurement of Productive Efficiency: Techniques and Applications*. New York: Oxford University Press.

Meeusen, W. and J. van den Broeck. (1977). “Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error.” *International Economic Review* 18(2), 435–444.

Olesen, O. B., N. C. Petersen and C. A. K. Lovell (eds.), (1996). “Efficiency and Frontier Analysis.” *Journal of Productivity Analysis* 7(2/3).

Pastor, J. T. (1995). “How to Account for Environmental Effects in DEA: An Application to Bank Branches.” Working Paper, Departamento de Estadistica e Investigacion Operativa. 03071 Alicante, SPAIN: Universidad de Alicante.

Schmidt, P. and R. C. Sickles. (1984). “Production Frontiers and Panel Data.” *Journal of Business and Economic Statistics* 2(4), 367–374.

Simar, L. and P. W. Wilson. (2000). “Statistical Inference in Nonparametric Frontier Models: The State of the Art.” *Journal of Productivity Analysis* 13(1), 49–78.

Timmer, C. P. (1971). “Using a Probabilistic Frontier Production Function to Measure Technical Efficiency.” *Journal of Political Economy* 79(4), 776–794.