ELSEVIER

European Journal of Operational Research 99 (1997) 72–88

EUROPEAN JOURNAL OF OPERATIONAL RESEARCH

Measures of inefficiency in data envelopment analysis and stochastic frontier estimation

W.W. Cooper [a,*], K. Tone [b]

[a] Graduate School of Business, The University of Texas at Austin, Austin, TX 78712-1174, USA
[b] Graduate School of Policy Science, Saitama University, Urawa, Saitama 338, Japan

Received 1 September 1995; revised 1 June 1996

Abstract

This paper discusses recent work in developing scalar measures of inefficiency which (a) comprehend *all* inefficiencies, including non-zero slacks, and (b) are readily interpretable and easily used in a wide variety of contexts. The opening section of the paper discusses some of the varied contexts in which uses of DEA are now being reported. This provides background for some of these measures. The closing section turns to simulation studies of DEA–regression combinations and possible inefficiency measures. Serious problems of bias in SF (Stochastic Frontier) regression approaches are identified. Extensions and modifications are suggested which can make a development of other inefficiency measures worthwhile for SF extensions to input-specific and multiple output evaluations. © 1997 Elsevier Science B.V.

Keywords: Data envelopment analysis; Regression analysis; Measures of inefficiency

1. Introduction

DEA can best be described as 'data-oriented' in that it effects its evaluations and inferences directly from observed data. A flood of reports on uses and extensions of DEA is now available, and increasing at an increasing rate. For instance, the bibliography published by Seiford [43] contains some 500 references to published articles, and many more have appeared since. Many (but not all) of these references use DEA to evaluate the performances of not-for-profit and governmental entities which are involved in activities that have proved resistant to other methods of inference and evaluation. The entities evaluated are referred to as DMUs (Decision Making Units) which are engaged in activities that use multiple inputs to produce multiple outputs with no easily identified 'bottom line'. Examples include schools and universities, military services, hospitals, court systems, prisons [23,31], and, more recently, whole economic and social systems [33,39].

Developments in DEA involve OR tools, such as mathematical programming, as well as economic and managerial concepts of efficiency and effectiveness. Concepts had to be modified, however, and tools had to be reoriented so that they could be used to effect inferences from already generated data (ex post) for purposes of

* Corresponding author.

PII S0377-2217(96)00384-0

evaluation and control – e.g., as distinguished from their more traditional ex ante planning uses in OR and economics. This is accomplished in the following manner: Application is by means of a series of mathematical programming optimizations to observed data generated from past behavior in order to effect inferences and evaluate the performances of each DMU.

This method of effecting inferences from observed data can be regarded as replacing or supplementing customary approaches in statistics. The two approaches may be contrasted as follows: In typical uses of statistics the inferences are obtained from optimizations over all observations. In DEA, the inferences are obtained from solutions which are optimal for each observation. Opportunities are thus opened for complementing as well as for competing uses of DEA and statistical methods. This includes uses such as 'cross-checking' by comparing results from statistical inferences with results from DEA (and vice versa). See Mendoza [41] or Ferrier and Lovell [30]. It also includes joint uses of DEA and statistical approaches that promise to complement, extend and sharpen the capabilities of both in addressing problems in ways that would otherwise not be available (see Bardhan [11]; see also Arnold et al. [7]).

Common to all uses of DEA are the choices of (a) the *inputs* and *outputs* from which evaluations are to be effected, and (b) the choices of DMUs as the entities to be evaluated. Different choices of DMUs can lead to different results (and yield different insights) and the same is true of the input and output choices. These choices also offer possibilities that can be exploited in ways that include expanding or contracting the number of inputs and outputs or the number of DMUs. This provides additional routes for conducting sensitivity analyses (see Thompson et al. [44] and Banker et al. [8]).

Weights can be used, of course, or suitable aggregates can be assembled from initially designated inputs, outputs and DMUs. A priori choices of weights, however, are *not* required by DEA and, as emphasized in Cook and Kress [24], the developments in DEA have opened new ways of approaching preference structures and rankings (see also Green et al. [36] as well as Golany [32]). Finally, a use of *exact* weights may be replaced with upper or lower bounds while allowing DEA to determine a best set of values directly from records of past performance as in, for instance, the 'assurance region' approaches of Thompson et al. [44] (see also Cooper et al. [27] for an alternate approach called 'cone-ratio envelopments').

It is not possible to cover all of these topics here, so we proceed selectively as follows. The next three sections deal with some of the new approaches to inefficiency measurement and their uses in DEA. For instance, the Section 2 discusses the TDT measure which provides fundamental insights into the nature of DEA evaluations and measurements. Section 3 introduces measures referred to as MEP and MIP which comprehend non-zero slacks as well as the customarily used radial (weak efficiency) measures with values that are invariant to the units in which different inputs and outputs might be measured. Section 4 introduces another measure called RAM which extends this invariance to a choice of origins so that negative as well as positive inputs may also be treated without losing contact with the body of DEA theory. Attention is then turned to recent research dealing with ways to combine DEA and statistical methods for efficiency evaluation, where (a) deficiencies in the usual stochastic frontier (= econometric) approaches are identified, and (b) extensions to ordinary least squares are found to provide very satisfactory results. Suggestions for uses and additional problems for research are introduced at various points.

2. Measures of efficiency

The materials in this section are adapted from Banker and Cooper [10] (see also Cooper et al. [26]). Our discussions start with a measure of efficiency that has recently been introduced into the literature of DEA by Thompson et al. [44], viz.,

$$\max_{u,v} \frac{\sum_{r=1}^{s} u_r y_{ro}}{\sum_{i=1}^{m} v_i x_{io}} \bigg/ \frac{\sum_{r=1}^{s} u_r y_{rk}}{\sum_{i=1}^{m} v_i x_{ik}}, \tag{1}$$

where, for any choice of (u,v),

$$\frac{\sum_{r=1}^{s} u_r y_{rk}}{\sum_{i=1}^{m} v_i x_{ik}} = \max\left\{ \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \middle| j = 1,\ldots,n \right\}.$$

We refer to this as the TDT measure of relative efficiency. Here (u,v) are vectors, with components $u_r, v_i \geq 0$, determined from the observed values of the $i = 1,\ldots,m$ inputs used and $r = 1,\ldots,s$ outputs produced for each of $j = 1,\ldots,n$ DMUs. DMU_o, as represented in the numerator for the objective in (1), is the DMU to be evaluated by choosing (u,v) to maximize its value *relative* to the *highest* score that this *same* (u,v) choice accords to the similarly formed ratios for the entire collection of DMU_j, $j = 1,\ldots,n$, with DMU_o included in this collection. Hence the objective is to choose a best set of relative weights, with these weights changing in value as different DMUs are designated as DMU_o.

We can formally represent the 'ratio of ratios' in the objective of (1) by

$$0 \leq \frac{y_o}{x_o} \bigg/ \frac{y_k}{x_k} \leq 1, \tag{2}$$

where y_o, y_k represent 'virtual outputs' and x_o, x_k represent 'virtual inputs'. With the data all non-negative, this ratio has a lower bound of zero because only non-negative values are admitted for the u_r and v_i. Because y_k/x_k is maximal over the set $k = 1,\ldots,n$, which includes $k = o$, we have $y_o/x_o \leq y_k/x_k$. The above ratios therefore have a maximum value of unity – which is achievable if and only if DMU_o's performance is *not* bettered by some *other* DMU using the weights which give DMU_o the highest *relative* score.

This TDT measure may be interpreted in various ways. It might be regarded as an extreme value statistic, for example, and treated by suitably extended versions of the statistical theory of extreme values. The formulation in (1) may also be approached deterministically as a mathematical programming problem to be modeled and solved in ways suited to choosing the u and v vectors, and this is the way in which we now proceed.

The following DEA model, known as the CCR ratio model, as given Charnes et al. [20] – see also Rhodes [42] – will help to clarify what is involved,

$$\begin{aligned}
\max_{u,v} \quad & \frac{\sum_{r=1}^{s} u_r y_{ro}}{\sum_{i=1}^{m} v_i x_{io}} \\
\text{s.t.:} \quad & \\
& \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \leq 1; \quad j = 1,\ldots,n, \\
& \frac{u_r}{\sum_{i=1}^{m} v_i x_{io}} \geq \varepsilon; \quad r = 1,\ldots,s, \\
& \frac{v_i}{\sum_{i=1}^{m} v_i x_{io}} \geq \varepsilon; \quad i = 1,\ldots,m.
\end{aligned} \tag{3}$$

Here the only new element is ε, a positive non-Archimedean infinitesimal. We elaborate on its mathematical properties later after noting that its use ensures that *all* inputs and outputs are accorded 'some' positive value in these u_r and v_i choices. These ε values need not be specified explicitly but can be dealt with by computational processes like the ones described in [6].

A restriction to positive values for all variables is not present in (1). Hence, one referee remarked that setting $v = 0$ produces an infinite solution so that *any* DMU_o will thereby be characterized as efficient. We pursue this remark because it allows us to use TDT to frame what is involved in (1) by noting that the same $v = 0$ applies to the numerator ratio as well as the denominator ratio. Hence the choice suggested by the referee will result in an answer of $0/0$, which is not admissible because it is not a well-defined mathematical expression. The same result occurs for a choice of $u = 0$ which results in the numerator of each ratio being zero. Restricting (1) to well-defined mathematical expressions is therefore found to be equivalent to the treatment in Charnes et al. [22] (also intended to be fundamental), where it is required that at least one term in each of the numerators and in each of these denominators must be non-zero in (1). (This is a weaker condition than the requirement that all u_r, v_i must be positive as specified in the last s + m constraints in (3).)

The solution set for (1) is also unbounded from above. A question therefore arises with respect to possibilities in which some $u_r \to \infty$ in (1). The following two cases can then occur.

Case 1. When $y_o(u)/x_o(v)$ becomes and remains as large or larger than the largest of the other ratio values over the set $k = 1, \ldots, n$ as these $u_r \to \infty$ then

$$\lim_{u_r \to \infty} \frac{y_o}{x_o} \bigg/ \frac{y_k}{x_k} = \frac{y_o}{x_o} \bigg/ \frac{y_o}{x_o} = 1.$$

Case 2. When $y_k(u)/x_k(v)$ for some $k \neq o$ becomes and remains as large or larger than the largest of the ratio values which exceed $y_o(u)/x_o(v)$ as these $u_r \to \infty$, then

$$\lim_{u_r \to \infty} \frac{y_o}{x_o} \bigg/ \frac{y_k}{x_k} = 0.$$

Hence, the previously noted bounds continue to apply for (1) even in these extreme cases.

To interpret these two cases we note that the numerator and denominator fractions in (2) are both stated in terms of rates – viz., virtual output per unit virtual input. Hence, the rate of increase in the denominator ratio for Case 2, above, is such that even a small increase in input will produce a vastly greater (= infinite) relative increase in virtual output for DMU_k than for DMU_o. Case 1 is interpreted along similar lines by comparing DMU_o with any DMU_k which does not reach the limit at a rate exceeding the rate obtainable by DMU_o for the outputs associated with these $u_r \to \infty$.

The formulation in (3) is less general than (1), but it provides a route for implementation and further interpretation. It, too, is also very general for, as shown in [20], the formulation in (3), greatly generalizes the usual single-output-to-single-input ratio definitions of efficiency that are used in engineering and science. We can also relate these engineering-science definitions and usages to definitions in economics – e.g., the Pareto–Koopmans–Farrell definition of efficiency given in Charnes et al. [19] – which can be accorded operationally implementable form as follows:

Efficiency: The performance of DMU_o is to be considered fully (100%) DEA efficient if and only if the performance of other DMUs does not provide evidence that some of its inputs or outputs could be improved without worsening some of its other inputs or outputs.

We will shortly provide a transformation of (3) that makes it possible to identify the sources and estimate the relative amounts of inefficiency in *each* input and output for *every* DMU from evidence supplied by the data. Here we note that a relation to (1) is established by observing that a necessary condition for optimality in (3) is that at least one of the $j = 1, \ldots, n$ output-to-input ratios in the constraints must be at its upper bound of unity. The denominator in (1) then has a value of unity and the efficiency evaluation for DMU_o reduces to whether the numerator in (1) is unity or less.

3. Linear programming equivalents

Reference to (3) shows that it is a non-linear, non-convex programming problem, and hence is best used for conceptual clarification. To give these concepts computationally implementable form, we introduce new variables defined as follows,

$$\begin{aligned} &\mu_r = tu_r, \quad r = 1,\ldots s, \\ &\nu_i = tv_i, \quad i = 1,\ldots,m, \\ &1 = \sum_{i=1}^{m} \nu_i x_{io}. \end{aligned} \tag{4}$$

These are the so-called 'Charnes–Cooper transformations' from Charnes and Cooper [18] which initiated the field of 'fractional programming'. Here we use them to transform the problem in (3) to the problem on the right in the following dual pair of linear programming problems with assurance (from fractional programming) that their optimal values will also be optimal for (3).

$$\begin{aligned} \min \quad & \theta - \varepsilon\left(\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+\right) & \max \quad & \sum_{r=1}^{s} \mu_r y_{ro} \\ \text{s.t.:} \quad & & \text{s.t.:} \quad & \\ & 0 = \theta x_{io} - \sum_{j=1}^{n} x_{ij}\lambda_j - s_i^-, & & -\sum_{i=1}^{m} \nu_i x_{ij} + \sum_{r=1}^{s} \mu_r y_{rj} \le 0, \\ & y_{ro} = \sum_{j=1}^{n} y_{rj}\lambda_j - s_r^+, & & \sum_{i=1}^{m} \nu_i x_o = 1, \\ & 0 \le \lambda_j,\ s_i^-,\ s_r^+ & & -\nu_i \le -\varepsilon, \\ & & & -\mu_r \le -\varepsilon. \end{aligned} \tag{5}$$

As before, $i = 1,\ldots,m$ indexes the inputs, $r = 1,\ldots,s$ indexes the outputs, and $j = 1,\ldots,n$ indexes the DMUs. Also, $j = o$ is used to identify the DMU to be evaluated by (a) placing it in the objective while also (b) leaving it in the constraints. Leaving the data for DMU_o in the constraints guarantees that solutions exist for both problems in (5) and, by the dual theorem of linear programming, it follows that they will have finite and equal optimal values.

We now assume that the x_{ij}, y_{rj} are all positive. (This condition can be relaxed; see Charnes et al. [22].) This means that we can move back and forth between (5) and (3) since the constraint $\sum_{i=1}^{m} \nu_i x_{io} = 1$, $\nu_i \ge 0$, for all i, means that we have $t > 0$ in (5) and hence in (4). We then have the full power of available linear programming algorithms and computer codes to solve (5) or (3), as we wish. We also have its interpretative power available (after suitable adaptations) for use in DEA efficiency analyses and inferences.

Using * to denote an optimal value, the condition for full (100%) DEA efficiency, as defined above, becomes

$$\sum_{r=1}^{s} \mu_r^* y_{ro} = 1 \tag{6}$$

for the problem on the right in (5) which is referred to as being in 'multiplier form'. Interest usually attaches to identifying sources and amounts of inefficiency in each input and output. This is most easily done from the problem on the left in (5) – which is referred to as the 'envelopment form'. For this problem the conditions for efficiency become;

$$\begin{aligned} &\text{(i)} \quad \theta^* = 1, \\ &\text{(ii)} \quad \text{All optimum slack values are zero.} \end{aligned} \tag{7}$$

It is to be noted that the presence of non-zero slacks means that the measure of inefficiency resulting from (7) involves a two-component number of the form $\theta^* - k^*\varepsilon < 1$, where k^* = sum of slacks. Both θ^* and k^* are real numbers and hence are Archimedean, whereas ε is a non-Archimedean infinitesimal so that $\theta^* - k^*\varepsilon$ is not a real number unless $k^* = 0$. From the duality theorem of linear programming it follows that Eq. (6) must also involve such non-Archimedean elements in its expression at an optimum for such cases. (For detailed discussions, including the new linear programming theorems this usage leads to, see Arnold et al. [6].)

Use of the envelopment model makes it possible to distinguish between 'mix' and 'technical' inefficiencies in the inputs. This is accomplished by noting that 'mix' refers to the proportions in which inputs are used. Minimizing θ preemptively, as is implied in the objective for the envelopment model of (5), provides a measure of inefficiencies in the proportion $(1 - \theta^*)$. This reduction applies to all inputs and hence does not alter their mix proportions. We refer to this proportionate reduction in all inputs as a measure of 'pure technical inefficiency' with value $0 < (1 - \theta^*) \leq 1$. Proceeding to maximization of the slacks – as is done in a 2^d stage without altering the value of θ^* – may produce non-zero slacks. Consistent with our definition of efficiency, as formalized mathematically in (7), we are then bound to recognize all such non-zero slacks as evidence of inefficiency obtained from the data. However, eliminating such non-zero slack will alter the observed input proportions and hence change the mix. We therefore refer to non-zero slacks as 'mix inefficiencies'. These mix inefficiencies can then be combined with pure technical inefficiencies to obtain $x_{io} - (\theta^* x_{io} - s_i^{-*})$ as the amounts needed to eliminate both types of inefficiencies in each of DMU$_o$'s $i = 1, \dots, m$ inputs.

4. MIP and MEP measures of inefficiency and efficiency proportions

We might note here that both problems in (5) provide scalar measures of efficiency when (6) is fulfilled. As has just been noted, however, a presence of non-zero slacks in a solution for the envelopment model in (5) implies that $\sum_{r=1}^{s} \mu_r^* y_{ro} < 1$ with non-Archimedean elements possibly involved in some of these μ_r^* values. A single real-number measure of inefficiency is then not available.

If desired, an alternative real-number measure may be developed as follows from the MID and MED measures developed by Bardhan et al. [12] for measuring 'efficiency dominance'. First, we observe that the non-Archimedean element $\varepsilon > 0$ is not present in the constraints for the envelopment problem on the left in (5). Hence, the values in the constraints involve only real numbers. Thus, when an optimal solution is available we have,

$$\theta^* x_{io} - s_i^{-*} = \sum_{j=1}^{n} x_{ij}\lambda_j^* = x_{io}^*,$$

where x_{io}, $i = 1, \dots, m$, represents the thus adjusted 'efficiency' value of the ith input for DMU$_o$. We therefore have

$$x_{io} - x_{io}^* \leq x_{io} \quad \text{or} \quad \frac{x_{io} - x_{io}^*}{x_{io}} \leq 1, \tag{8}$$

for each of the $i = 1, \dots, m$ inputs. Similarly, we can write

$$y_{ro} + s_r^{+*} = y_{ro}^* \quad \text{or} \quad \frac{y_{ro}^* - y_{ro}}{y_{ro}^*} \leq 1, \tag{9}$$

since y_{ro}^*, the efficient value, satisfies $y_{ro}^* \geq y_{ro}$ for each of the $r = 1, \dots, m$ outputs of DMU$_o$.

We now note that the numerators and denominators in (8) and (9) are expressed in the same units and hence are 'dimensionless' – i.e., multiplying numerator and denominator elements by any positive constant to change the unit of measure of the corresponding input or output will not change the value of any of these measures. We also note that the expressions in (8) and (9) are all non-negative. Hence we can use them to derive the following measure of inefficiency,

$$0 \leq \left(\sum_{i=1}^{m} \frac{x_{io} - x_i^*}{x_{io}} + \sum_{r=1}^{s} \frac{y_r^* - y_{ro}}{y_r^*} \right) / (s+m) \leq 1, \tag{10}$$

where s is the number of outputs and m is the number of inputs. As can be seen, this is a real number between zero and one with a value that represents an average of the inefficiency proportions due to (i) excessive inputs in the first term, and (ii) output shortfalls in the second term. To convert this to a measure of efficiency we replace the above with

$$0 \leq 1 - \left[\left(\sum_{i=1}^{m} \frac{x_{io} - x_i^*}{x_{io}} + \sum_{r=1}^{s} \frac{y_r^* - y_{ro}}{y_r^*} \right) / (s+m) \right] \leq 1. \tag{11}$$

Referring to the measure in (10) as MIP (Measure of Inefficiency Proportions) and the measure in (11) as MEP (Measure of Efficiency Proportions) we see that full efficiency is attained only with MEP = 1 or MIP = O.

This provides a way of reducing the conditions represented in (7) to a single real number without losing the ability to identify the inefficiencies that may be present in inputs, in outputs, or in subsets thereof. Extensions to weighted measures and the treatment of zeros which may appear to be sources of trouble in some of the denominators are discussed in Bardhan et al. [12]. This treatment requires only very natural extensions of the MID and MED measures discussed in Bardhan et al. [12] as well as in Banker and Cooper [10] so we do not cover them here. We do need to note, however, that these measures lend themselves to rankings of DMU performances whereas this is not the case for the θ^* values obtained from (5) because (a) the latter measure is incomplete, and (b) these θ^* values will generally be determined from different facets – which means that these values are being derived from comparisons involving performances of different sets of DMUs.

5. RAM – A range adjusted measure of inefficiency

The just described MIP and MEP measures were developed on the assumption that all of the observed x_{ij} and y_{rj} are positive. On this assumption these measures are always well-defined and have the desirable property that their values do not depend on the units in which any input or output is measured. Undesirable properties include the appearance of y_{ro}^* in the denominator for the output terms (10). This causes no trouble if (9) or (10) is used as a measure *after* the model results have been obtained (see the discussion in Bardhan et al. [12]). It can, however, cause trouble when it is desired to use (10) as an objective for the envelopment problem in (5). This is due to the non-linearities associated with such an objective. Another trouble can be encountered in relaxing the requirement that all data must be positive. This relaxation may be important in dealing with certain outputs or inputs which are 'unwanted' or 'undesirable' – e.g., 'net losses' as contrasted with 'positive profits' as outputs, or degree days below zero when weather is used as an input, etc.

Recent work by Cooper and Pastor [25] deals with both of these problems through the following class of 'additive models',

$$\begin{aligned} \max \quad & \sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+ \\ \text{s.t.:} \quad & \\ & x_{io} = \sum_{j=1}^{n} x_{ij}\lambda_j - s_i^-, \quad i = 1,\ldots,m \\ & y_{ro} = \sum_{j=1}^{n} y_{rj}\lambda_j + s_r^+, \quad r = 1,\ldots,s \\ & 1 = \sum_{j=1}^{n} \lambda_j \\ & 0 \le \lambda_j,\ s_i^-,\ s_r^+,\ j = 1,\ldots,n. \end{aligned} \tag{12}$$

No non-Archimedean elements are involved in this class of models so the following single condition replaces the two conditions in (7),

Definition. DMU_o, the DMU_j being evaluated in (12), will be fully DEA efficient if and only if all slacks are zero at an optimum.

We next note that the convexity condition $\Sigma_{j=1}^{n}\lambda_j = 1$ may be adjoined to (5). This will convert (5) from its present form, called the 'CCR model' (as first given in Charnes et al. [20]) into what is called the 'BCC model' (as first given in Banker et al. [9]). We can use the following theorem and remark to relate this BCC version of (5) to (12).

Theorem. (Ahn et al. [1]) *A DMU_o is efficient when the BCC version of* (5) *is employed if and only if it is efficient when the additive model represented in* (12) *is used.*

Remark. When a DMU_o is inefficient, its sources and inefficiency amounts may differ because of differences in metrics employed for (12) and the BCC version of (5).

(See also Yu et al. [46] for a single model which can be used to represent these, and other, DEA models by varying one parameter.)

We amplify by noting that (12) uses a simple linear function represented in the objective of (12). This is expressed in the so-called ℓ_1 metric. (See discussion in Appendix A of Charnes and Cooper [17].) When full efficiency is achieved this sum is independent of the units of measure used, but not when any inefficiency is present. Indeed the solution choices may then depend on the units of measure used and the resulting sum may exceed unity.

These shortcomings may all be eliminated by replacing the objective in (12) by the following measure, called RAM (Range Adjusted Measure), as given in Cooper and Pastor [25]:

$$\mathrm{Max} \sum_{i=1}^{m} \frac{s_i^-}{R_i^-} + \sum_{r=1}^{s} \frac{s_r^+}{R_r^+}, \tag{13.1}$$

where R_i^- and R_i^+ are the Ranges defined by

$$R_i^- = \bar{x}_i - \underline{x}_i, \quad i = 1, \ldots, m,$$
$$R_r^+ = \bar{y}_r - \underline{y}_r, \quad r = 1, \ldots, s, \tag{13.2}$$

with $\bar{x}_i$, $\underline{x}_i$ and $\bar{y}_r$, $\underline{y}_r$ representing maximal and minimal observed values taken over the $j = 1, \ldots, n$ DMUs for *each* of the inputs and outputs in question.

Because the numerators and denominators in (13.1) are expressed in the same units, it follows that changes in the units of measure will not affect the value of any of these terms. Thus, as was true for MIP and MEP, the ratio in each of the terms in (13.1) is 'dimensionless'. Moreover, the numerator and denominator have the property that additions of an arbitrary constant to any input of output will not alter the value of (13.1) or the choice of λ_j values in the constraints. This is true because

$$(\bar{x}_i + d_i) - (\underline{x}_i + d_i) = \bar{x}_i - \underline{x}_i = R_i^-$$

and

$$(\bar{y}_r + c_r) - (\underline{y}_r + c_r) = \bar{y}_r - \underline{y}_r = R_r^+ ,$$

for each $i = 1, \ldots, m$ or $r = 1, \ldots, s$ as just defined. Similarly, as first shown in Ali and Seiford [4],

$$(x_{io} + d_i) - \sum_{j=1}^{n} (x_{ij} + d_i)\lambda_j = x_{io} - \sum_{j=1}^{n} x_{ij}\lambda_j = s_i^-$$

and

$$\sum_{j=1}^{n} (y_{rj} + c_r)\lambda_j - (y_{ro} + c_r) = \sum_{j=1}^{n} y_{rj}\lambda_j - y_{ro} = s_r^+$$

because $\Sigma_{j=1}^{n}\lambda_j = 1$, It follows that the choice of an origin is arbitrary. Hence, no problem need be encountered in dealing with negative inputs or outputs because these may be eliminated by adding positive constants which are large enough to make all of the thus adjusted data positive without changing the optimal solution sets or the value of the objective. Indeed, this property carries over into the outputs, but not the inputs, of (5) when it is accorded the BCC form. Hence, as suggested by one referee, we can use this property as a guide in choosing between these different models. For instance, if negative terms are to be dealt with in both inputs and outputs, then the choice should be in favor of (12) even though this involves extra labor to distinguish between the purely technical and mix inefficiencies that (5) automatically provides.

The expression (13.1) is non-negative, of course, but it need not be bounded by unity, as in (5), (3) and (1). However, this 'unity condition' can be satisfied, if desired, by dividing by the number of inputs plus outputs to obtain

$$0 \le \left[\sum_{i=1}^{m} \left(\frac{s_i^-}{R_i^-} \right) + \sum_{r=1}^{s} \left(\frac{s_r^+}{R_r^+} \right) \right] / (m+s) \le 1. \tag{14}$$

The result is evidently a measure of the average of both the input and output inefficiencies.

This RAM formulation has other desirable properties besides those we have just noted. For instance, it exhibited 'scaling' properties that made it possible to deal with large disparities in the inputs and outputs associated with the entities providing water services to different population centers in Japan (see Aida et al. [2]). Moreover, subdivisions of (13.1) may be used so that, say, input and output inefficiencies may be separately measured, if desired. Some of these properties may be carried over to (5) by dividing each slack variable by its possible range of values in the objective of the-envelopment model in (5), and this renders the expression dimensionless because θ^* is also dimensionless. A further modification which converts (5) into BCC form also makes it possible to deal with negative outputs (but not inputs) in a convenient manner.

There are possible problems and further developments with respect to such measures which we do not deal with here. Instead we refer to Thrall [45] where these topics are treated in general and in detail to provide methods for dealing with the slack components in any DEA model. With this reference available, we therefore simply close our discussion of RAM by noting that each (s^-/R_i^-) and (s_r^+/R_r^+) in (14) measures the amount of inefficiency (in the numerator) relative to the range of possible inefficiencies (in the denominator) for each input and output. That is, R_i^- and R_r^+ represent the maximum possible inefficiency over $j = 1, \ldots, n$ for *each i* and r. Hence, (14) represents an average of the maximum possible inefficiency proportions exhibited by DMU_o in each input and output. The result will be unity if and only if equality is attained for every term in (14). The result will be zero, and hence full efficiency will be achieved, if and only if all slacks are zero – in accordance with the definition given immediately after (12), above.

We can also orient our measure toward efficiency, rather than inefficiency, by replacing (14) with the following expression,

$$0 \le 1 - \left\{ \left[\sum_{i=1}^{m} \left(\frac{s_i^-}{R_i^-} \right) + \sum_{r=1}^{s} \left(\frac{s_r^+}{R_r^+} \right) \right] / (m+s) \right\} \le 1 \tag{15}$$

Hence, we have bounded our measure of efficiency to lie between zero and one with assurance that the resulting real-valued scalar comprehends all of both the purely technical and mix inefficiencies identified in the solutions for any DMU_o.

Classical literatures dealing with measures of inefficiency have devoted little, if any, attention to the dual (= multiplier) problem in (5). Indeed, the FDH (Free Disposal Hall) approach described in de Borger and Kerstens [28] corresponds to an integer programming problem which has no dual. See Bardhan et al. [12]. Confinement to the envelopment form in (5) has resulted in troubles for attempts to comprehend both of the conditions in (7) in a real-valued scalar measure of inefficiency. These problems along with attempts to resolve them are described in de Borger and Kerstens [28] and hence need not be discussed here. The MIP and RAM measure described above constitute new approaches and, of course, other approaches are also now being essayed by exploiting new models (like the additive model) which were not available in the classical literature (see, Färe et al. [29] for these classical approaches and Lovell and Pastor [40] for new modeling alternatives).

6. Stochastic frontiers and DEA-regression combinations

Green [37] provides a relatively up-to-date survey of statistical regression approaches to efficiency evaluations which he separates into (i) regressions which are deterministic, as in the formulations by Aigner and Chu [3], and (ii) regressions which involve statistical errors. The latter are referred to as 'stochastic' and further subdivided into (1) OLS (Ordinary Least Squares), and (2) SF (Stochastic Frontier models), with the former being directed to 'central tendency' and the latter being directed to 'frontier' estimates for effecting evaluations of observed performances.

The studies we now turn to represent yet another approach in which deterministic DEA approaches are combined with stochastic regressions to open additional avenues for development. Such combinations can be effected in a variety of ways, but the studies we examine here involve a two-stage approach which proceeds as follows: In stage one, DEA is applied to the data in order to distinguish which observations are associated with efficiently and which are associated with inefficiently performing DMUs. In stage two, the results of stage one are incorporated as 'dummy variables' in the regressions to be estimated. These may take the form of either OLS or SF regressions.

The research we now describe had its origins in empirical work directed to developing improved methods for evaluating the performances of public schools in Texas. This was followed by a simulation study which utilized statistically controlled experimental design approaches to better understand the results from this in-the-field study and it is the latter, as reported in Bardhan et al. [13], that we turn to in our discussion.

The following very simple Cobb–Douglas production function was used in one part of this simulation study

$$y = a x_1^{\alpha_1} x_2^{\alpha_2} e^{\varepsilon}. \tag{16.1}$$

This function was chosen because it represented a simplification of the Cobb–Douglas forms used in the field study and so its use could facilitate the task of understanding what might have happened in both the unsuccessful and successful uses of these functional forms in the previous field applications (see Bardhan [11]).

In classical economics a 'production function' maximizes the value of the single output, y, for any of the inputs that might be used in the amounts x_1 and x_2. Such a function therefore defines the technologically (absolutely) efficient frontier to which the observations must conform. We need to allow for statistical errors, however, so the frontier will be known only stochastically when the parameters a, α_1, and α_2 are estimated from observational data. These statistical errors are represented by ε in Eq. (16.1).

For the simulations reported in Bardhan et al. [11], the following parameters were used:

$$a = 0.75, \quad \alpha_1 = 0.65, \quad \alpha_2 = 0.55. \tag{16.2}$$

One thousand x_1 and x_2 values were then determined by random draws from uniform probability distributions and substituted in Eq. (16.1) to obtain the corresponding 1,000 values of y. Statistical error terms were then added to each of these y values by sampling from normal distributions $N(0,\sigma_\varepsilon^2)$.

After this was done the 'efficient' x_1 and x_2 input values were replaced with new inputs

$$\hat{x}_1 = x_1 e^{\tau_1} \text{ and } \hat{x}_2 = x_2 e^{\tau_2}, \quad \text{with } \tau_1, \tau_2 \geq 0. \tag{17}$$

The literatures dealing with SF regressions commonly assume that inefficiencies conform to either the exponential or half-normal distributions. Hence these distributions were used to replace Eq. (16.1) with

$$y = \beta \hat{x}_1^{\beta_1} \hat{x}_2^{\beta_2} e^{\varepsilon}, \tag{18}$$

where $\hat{x}_1 \geq x_2$ and $\hat{x}_2 \geq x_2$ simulated the actually observed inputs. Then, for estimating purposes, Eq. (18) was transformed into the following log-linear form

$$\ln y = \beta_0 + \beta_1 \ln \hat{x}_1 + \beta_2 \ln \hat{x}_2 + \varepsilon, \tag{19}$$

where the β's serve as estimators for the true (production function) parameters in Eq. (16.2).

The expression in Eq. (19) is the OLS form, which we want to compare with

$$\ln y = \beta_0 + \beta_1 \ln \hat{x}_1 + \beta_2 \ln \hat{x}_2 + \delta D + \delta_1 D \ln \hat{x}_1 + \delta_2 D \ln \hat{x}_2 \varepsilon, \tag{20}$$

where

$$D = \begin{cases} 1 & \text{if a DMU is 100\% DEA efficient,} \\ 0 & \text{if a DMU is not 100\% DEA efficient.} \end{cases} \tag{21}$$

Thus, (20) reflects our two-stage process in which DEA is first used to determine whether an observation reflects the performance of an efficient or inefficient DMU. (This condition was relaxed to allow values of $\theta^* \geq 0.98$ to serve in place of only $\theta^* = 1$ in (7) but, as noted in Bardhan et al. [13], no appreciable difference in results occurred.)

As emphasized in (1) ff., DEA deals only with *relative* efficiency. Here, however, we want to gauge our results against the known levels of absolute efficiency associated with Eq. (16.1) and Eq. (16.2). The study reported in Bardhan et al. [13] therefore arranged to have some of the τ values set equal to zero in proportions ranging from 0.10 to 0.25 which were assigned randomly to the different observations so it was possible to be sure the resulting parameter estimates could be meaningfully compared with the true values for absolutely efficient performance as given in Eq. (16.2).

Table 1
OLS regression estimates without dummy variables – Case 1: *Exponential* distribution of input inefficiencies

Parameter estimates [a]	Case A $\sigma_\varepsilon^2 = 0.04$ (1)	Case B $\sigma_\varepsilon^2 = 0.0225$ (2)	Case C $\sigma_\varepsilon^2 = 0.01$ (3)	Case D $\sigma_\varepsilon^2 = 0.005$ (4)
β_0	1.30 *	1.58 *	1.40 *	1.43 *
	(0.19)	(0.15)	(0.13)	(0.10)
β_1	0.46 *	0.43 *	0.45 *	0.46 *
	(0.024)	(0.02)	(0.016)	(0.013)
β_2	0.48 *	0.47 *	0.47 *	0.46 *
	(0.02)	(0.013)	(0.01)	(0.01)

The asterisk '*' denotes statistical significance at the 0.05 significance level or better.
[a] The values for σ_ε^2 shown in the top row of the table represent the true variances for the statistical error distributions. Standard errors are shown in parentheses.

Results from these simulation experiments, using an exponential distribution for selecting the τ_1 and τ_2 values, are reproduced in Tables 1 and 2 for the 4 levels of normally distributed statistical errors associated with $N(0,\sigma_\varepsilon^2)$ that are recorded at the top of each column. As is clear from Table 1, which reports results for Eq. (19), the estimates differ significantly from the true parameter values in every case. The converse situation occurs for Eq. (20) since, as shown at the bottom of Table 2, the t values are all so low that it is not possible to reject H_o without experiencing unacceptably high risks of rejecting the true parameter values – viz., $\alpha_1 = 0.65$ and $\alpha_2 = 0.55$ as specified in Eq. (16.2).

The same results occurred when half-normal rather than exponentially distributed values of τ were used. That is, H_o could not be rejected in *any* of the two-stage dummy variable regressions represented by Eq. (20)

Table 2
OLS regression estimates with dummy variables on DEA-efficient DMUs – Case 1: *Exponential* distribution of input inefficiencies

Parameter estimates [a]	Case A $\sigma_\varepsilon^2 = 0.04$ (1)	Case B $\sigma_\varepsilon^2 = 0.0225$ (2)	Case C $\sigma_\varepsilon^2 = 0.01$ (3)	Case D $\sigma_\varepsilon^2 = 0.005$ (4)
β_0	1.07 *	1.47 *	1.28 *	1.34 *
	(0.21)	(0.17)	(0.14)	(0.11)
β_1	0.49 *	0.43 *	0.46 *	0.47 *
	(0.03)	(0.02)	(0.02)	(0.01)
β_2	0.48 *	0.48 *	0.48 *	0.46 *
	(0.02)	(0.015)	(0.01)	(0.01)
δ	−1.57 *	−2.30 *	−1.50 *	−1.50 *
	(0.64)	(0.43)	(0.35)	(0.21)
δ_1	0.155	0.26 *	0.16 *	0.16 *
	(0.075)	(0.05)	(0.04)	(0.03)
δ_2	0.12 *	0.12 *	0.10 *	0.09 *
	(0.05)	(0.04)	(0.03)	(0.02)
Combining parameters with *dummy variables*				
H_o: $\beta_1 + \delta_1 = 0.65$	$t_1 = 0.07$	$t_1 = 0.87$	$t_1 = -0.72$	$t_1 = 0.82$
H_a: $\beta_1 + \delta_1 \neq 0.65$				
H_o: $\beta_2 + \delta_2 = 0.55$	$t_2 = 1.09$	$t_2 = 1.76$	$t_2 = 1.02$	$t_2 \cong 0$
H_a: $\beta_2 + \delta_2 \neq 0.55$				

The asterisk '*' denotes statistical significance at the 0.05 significance level or better. Standard errors are shown in parentheses
[a] The values for σ_ε^2 shown in the top row of the table represent the true variances for the statistical error distributions. Standard errors are shown in parentheses.

but H_o was *always* rejected when the one-stage approach represented in Eq. (19) was used. These same results were also secured in other parts of this simulation experiment where a CES (Constant Elasticity of Substitution) production function was used in place of Eq. (16.1) as well as in sensitivity analyses undertaken with (a) different σ_ε^2 values, and (b) different proportions assigned to the $\tau = 0$ values.

There seems to be little point in pursuing these OLS studies in more detail here. We therefore turn to results from the SF (Stochastic Frontier) regressions that were also covered in these studies. For this purpose, we replace Eq. (19) with

$$\ln y = \beta_0 + \beta_1 \ln \hat{x}_1 + \beta_2 \ln \hat{x}_2 + \nu - \tau. \tag{22.1}$$

This is the 'composed-error' version of a SF regression function in which ν and τ are both statistically determined, with $\tau \geq 0$ representing inefficiencies and ν, which is unconstrained, representing statistical noise – generally assumed to be normally distributed $N(0,\sigma_\nu^2)$. We can gain some insight into the way this SF regression deals with input inefficiencies by rewriting Eq. (22.1) in the following form

$$\ln y + \tau = \beta_0 + \beta_1 \ln \hat{x}_1 + \beta_2 \ln \hat{x}_2 + \nu. \tag{22.2}$$

As this representation makes clear, the output in this SF regression is to be augmented by $\tau \geq 0$ to reflect the extra input amounts in $\hat{x}_1$ and $\hat{x}_2$. In the terminology of Gong and Sickles, [34,35], these $\tau \geq 0$ values are referred to as 'foregone outputs' and interpreted to mean that this extra output amount is *estimated* to be attainable in place of the actually observed y values in response to the inputs $\hat{x}_1 \geq x_1$ and $\hat{x}_2 \geq x_2$. Here, however, we emphasize that this approach does not try to adjust the input inefficiencies in an 'input-specific' manner.

Simulation studies have to date reported fairly good behavior for these composed error models. Gong and Sickles [34], for instance, report that this model performs better than DEA when the simulations are based on sufficiently complex production functions. However such studies have all used procedures to generate inefficiencies in conformance with Eq. (22.1) or Eq. (22.2). No attempt was made to assign inefficiencies to inputs in the manner we used to generate $\hat{x}_1$ and $\hat{x}_2$. When this is done, however, the results are far from satisfactory. In fact, as Table 3 shows, every one of the *estimated* parameter values differs significantly from their true values. In short, this composed error model give erroneous results in every one of these simulated experiments.

Table 3
Stochastic frontier regression estimates without dummy variables – Case 1: *Exponential* distribution of input inefficiencies

Parameter estimates [a]	Case A $\sigma_\varepsilon^2 = 0.04$ (1)	Case B $\sigma_\varepsilon^2 = 0.0225$ (2)	Case C $\sigma_\varepsilon^2 = 0.01$ (3)	Case D $\sigma_\varepsilon^2 = 0.005$ (4)
β_0	1.42 *	1.62 *	1.25 *	1.28
	(0.19)	(0.14)	(0.14)	(0.11)
β_1	0.46 *	0.43 *	0.48 *	0.46 *
	(0.024)	(0.02)	(0.017)	(0.01)
β_2	0.48 *	0.47 *	0.48 *	0.47 *
	(0.017)	(0.013)	(0.01)	(0.01)
σ_τ	0.15 *	0.11 *	0.15 *	0.15 *
	(0.035)	(0.01)	(0.01)	(0.01)
σ_ν	0.15 *	0.13 *	0.08 *	0.04
	(0.01)	(0.02)	(0.01)	(0.025)

The asterisk '*' denotes statistical significance at the 0.05 significance level or better.
[a] The values for σ_ε^2 shown in the top row of the table represent the true variances for the statistical error distributions. Standard errors are shown in parentheses.

We can get some insight into these failures by returning to Eq. (22.1). Then because our τ_1, $\tau_2 \geq 0$ values were generated from the same distribution we have $E\tau_1 = E\tau_2 = E\tau$ so, utilizing $E\nu = 0$, we can write

$$E(\ln y \mid \hat{x}_1 \hat{x}_2) = \beta_0 + \beta_1 \ln x_1 + \beta_2 \ln x_2 + (\beta_1 + \beta_2 - 1)E'\tau, \tag{23}$$

where $E'\tau = (1-\gamma)E\tau$ and $(1-\gamma)$ reflects the probability mixtures used to generate $\tau = 0$ and $\tau \geq 0$, respectively. We evidently have $E\tau > 0$ when either exponential or half-normal distributions are used to generate the inefficiencies associated with τ_1 and τ_2 in (17). Hence to have relatively unbiased estimates when $\hat{x}_1 = x_1$ and $\hat{x}_2 = x_2$ we must have $\beta_1 + \beta_2 \approx 1$, at least approximately, if we are to obtain the correct frontier, or anything even close to it. However, this approximation will produce a downward bias in the parameter estimates since $\alpha_1 + \alpha_2 = 1.20$. Thus the bias has its source in the orientation toward correct estimates of y on the frontier at the expense of correct estimates of the parameters associated with each of the inputs. Moreover, this bias is further accentuated from the observations associated with inefficient DMUs which, except for statistical errors, must always lie below the efficient frontier defined by Eq. (16.1) and Eq. (16.2).

Maximum likelihood methods are used to obtain the β values in Table 3 and hence these estimates may be expected to have some bias in exchange for improved efficiency. The small variances for the statistical error terms recorded at the top of Table 3 make it clear, however, that the error terms (and their associated variances) will be close to parameter values which are *not* correct. Hence the efficiency properties of these maximum likelihood estimates do not compensate for this bias in a satisfactory manner. Finally, with samples of size $n = 1{,}000$, properties of statistical consistency – or the other desirable properties of such maximum likelihood estimates – also cannot be relied upon.

The serious nature of these underestimates from an economics standpoint is made clear from the fact that

Table 4
Stochastic frontier regression estimates with dummy variables on DEA-efficient DMUs – Case 1: *Exponential* distribution of input inefficiencies

Parameter estimates [a]	Case A $\sigma_\varepsilon^2 = 0.04$ (1)	Case B $\sigma_\varepsilon^2 = 0.0225$ (2)	Case C $\sigma_\varepsilon^2 = 0.01$ (3)	Case D $\sigma_\varepsilon^2 = 0.005$ (4)
β_0	1.18 *	1.50 *	0.80 *	1.40 *
	(0.23)	(0.16)	(0.16)	(0.13)
β_1	0.50 *	0.44 *	0.53 *	0.49 *
	(0.03)	(0.02)	(0.02)	(0.01)
β_2	0.48 *	0.49 *	0.50 *	0.47 *
	(0.02)	(0.02)	(0.01)	(0.02)
δ	−1.60 *	−2.4 *	−1.25 *	−1.55 *
	(0.57)	(0.56)	(0.38)	(0.23)
δ_1	0.16 *	0.26 *	0.13 *	0.15 *
	(0.07)	(0.06)	(0.04)	(0.03)
δ_2	0.11 *	0.13 *	0.086	0.09 *
	(0.05)	(0.04)	(0.04)	(0.03)
σ_1	0.23 (0.03) *	0.16 (0.01) *	0.17(0.007) *	0.11 (0.04) *
σ_2	0.15 (0.02) *	0.10 (0.02) *	0.08 (0.02) *	0.05 (0.03)
σ_ν	0.13 (0.01) *	0.09 (0.01) *	0.05 (0.01) *	0.04 (0.01)
Combining parameters with *dummy variable*				
H_o: $\beta_1 + \delta_1 = 0.65$	$t_1 = 0.20$	$t_1 = 0.93$	$t_1 = 0.28$	$t_1 = -0.4$
H_a: $\beta_1 + \delta_1 \neq 0.65$				
H_o: $\beta_2 + \delta_2 = 0.55$	$t_2 = 1.03$	$t_2 = 1.90$	$t_2 = 1.16$	$t_2 = 0.45$
H_a: $\beta_2 + \delta_2 \neq 0.55$				

The asterisk '*' denotes statistical significance at the 0.05 significance level or better.

[a] The values for σ_ε^2 shown in the top row of the table represent the true variances for the statistical error distributions. Standard errors are shown in parentheses.

$\beta_1 + \beta_2 < 1$ in every case whereas $\alpha_1 + \alpha_2 = 1.2$. Hence the returns-to-scale properties of Eq. (16.1) and Eq. (16.2) are reversed by these estimates. Increasing returns are replaced with decreasing returns to scale in every case.

The same situation holds for important statistical properties. For instance, as shown in Charnes et al. [21], the β_0 terms may be adapted so they can be used as MDI (Minimum Discrimination Information) statistics – which have the unusual property that the best alternative hypothesis is automatically supplied when H_o is rejected. (See Brockett et al. [14] for details and interpretations.) In the present case, however, $\log a < 1$ whereas $\beta_0 = \ln \beta > 1$, so rejection of H_o points in the *wrong* direction.

The column headings in Tables 1–4 show that the statistical error terms were generated from distributions which tend to be leptokurtic and hence tend to have relatively small error values. The objective of the study was to focus on the behavior of the inefficiency components which were therefore allowed much wider ranges of variation. It was believed that this would make it easier to identify inefficiencies with their input sources as a guide to further research which we comment on as follows: A lot has been invested in these composed error approaches, so we are not suggesting that they be abandoned. We believe rather that what has already been done can be built upon by moving further toward developing input-specific formulations and methods of estimation like those given in Kumbhakar [38] and Chaffai [16]. It would be even better if this could be extended to formulations that could handle multiple-output as well as multiple-input specific inefficiencies. Recourse to more complex procedures will undoubtedly be required to accomplish this and some time will be needed for their development.

A use of our two-stage procedure for the composed error model represented in Eq. (22.1) and Eq. (22.2) also produces a modification of these SF approaches which yield the very satisfactory results shown in Table 4. However, these estimates, as obtained from these two-stage uses of SF regressions are no better (or worse) than those obtained from our two-stage procedure with OLS as reported in Table 2. Thus, the extra effort in securing these SF estimates may not be worthwhile unless special interest attaches to the behavior of σ_ν and the other standard deviation values shown near the bottom of Table 4.

We do not discuss the latter topics here. Instead we simply refer to the discussions in Bardhan et al. [13]. We can then close our present paper by returning to issues related to choices of measure like those we discussed for DEA.

To obtain a measure of the indicated kind, we can return to Eq. (20) and utilize $\mathrm{E}\varepsilon = 0$ to obtain

$$\begin{aligned} &\mathrm{E}(\ln y \mid D=1) - \mathrm{E}(\ln y \mid D=1) \\ &\quad = \beta_0 + \beta_1 \ln \hat{x}_1 + \beta_2 \ln \hat{x}_2 + \delta + \delta_1 \ln \hat{x}_1 + \delta_2 \ln \hat{x}_2 - \beta_0 + \beta_1 \ln \hat{x}_1 + \beta_2 \ln \hat{x}_2 \\ &\quad = \delta + \delta_1 \ln \hat{x}_1 + \delta_2 \hat{x}_2 . \end{aligned} \tag{24}$$

Bardhan et al. [13] suggest using the last of these expressions as a measure of the inefficiency associated with the $\hat{x}_1$, $\hat{x}_2$ choices for each DMU. Indeed, if the data satisfy

$$\delta_1 \ln \hat{x}_1 + \delta_2 \ln \hat{x}_2 > -\delta, \tag{25}$$

we can then use these expressions to develop a measure analogous to the one we considered for DEA in the following form:

$$0 \le \frac{\hat{x}_1^{\beta_1} \hat{x}_2^{\beta_2}}{\mathrm{e}^{\delta} \hat{x}_1^{\beta_1+\delta_1} \hat{x}_2^{\beta_2+\delta_1}} = \frac{1}{\mathrm{e}^{\delta} \hat{x}_1^{\delta_1} \hat{x}_2^{\delta_1}} \le 1. \tag{26}$$

This is a measure of output inefficiency. Input inefficiency measures are readily available from the first-stage use of DEA which can be aggregated, if desired, in the form of the MIP, MEP and RAM measures discussed earlier in this paper.

In keeping with other parts of this paper we now close by noting another problem for further research as follows. It is possible that Eq. (25) may hold for some choices of $\hat{x}_1$, $\hat{x}_2$ and fail for other choices. This means that the functional forms connected with $\mathrm{E}(\ln y \mid D=1)$ and $\mathrm{E}(\ln y \mid D=0)$ have crossed over. (See Bardhan et al. [13] for a discussion of Eq. (20) as two separate equations.) This can result from statistical error. It can also occur because DEA misclassified some of the DMUs in stage 1. Development of statistical tests to discriminate between these two possibilities is therefore a subject which can be added to our topics for further research and use. See Brockett and Golany for a discussion of non-parametric (= distribution-free) statistical approaches.

Acknowledgements

This paper has benefitted from comments supplied by R.M. Thrall. Support for this research by the IC^2 Institute of the University of Texas at Austin is gratefully acknowledged.

References

[1] Ahn, T., Charnes, A., and Cooper, W.W., "A note on the efficiency characterizations obtained in different DEA models", *Soci-Economic Planning Sciences* 22 (1988) 253–257.
[2] Aida, K., Cooper, W.W., and Pastor, J.T., "Evaluating water supply services in Japan with RAM-A range-adjusted measure of inefficiency", *Omega* (1997), to appear.
[3] Aigner, D., and Chu, S., "On estimating the industry production function", *American Economic Review* 58 (1968) 826–839.
[4] Ali, I., and Seiford, L., "Translation invariance in data envelopment analysis", *Operations Research Letters* 9 (1990) 403–4.
[5] Arnold, V., Bardhan, I., and Cooper, W.W., "A two-stage DEA approach for identifying and rewarding efficiency in Texas secondary schools", in: W.W. Cooper F.Y. Phillips, A. Whinston and S. Thore (eds.), *IMPACT: How IC^2 Research Affects Public Policy and Business Markets*, Kluwer Academic Publishers, New York, 1997.
[6] Arnold, V., Bardhan, I., Cooper, W.W. and Gallegos, A., "Primal and dual optimality in computer codes using two-stage solution procedures in DEA", in: J. Aronson and S. Zionts (eds.), *Operations Research: Methods, Models and Applications, A Conference in Honor of G.L. Thompson on His 70^{th} Birthday*, Kluwer Academic Press, forthcoming, 1996.
[7] Arnold, V., Bardhan, I., Cooper, W.W., and Kumbhakar, S., "New uses of DEA and statistical regressions for efficiency evaluation and estimation – With an illustrative application to public secondary schools in Texas", *Annals of Operations Research* (1996) forthcoming.
[8] Banker, R.D., Chang, H., and Cooper, W.W., "Simulation studies of efficiency and returns to scale and misspecification with nonlinear functions in DEA", *The Annals of Operations Research* (1996) to appear.
[9] Banker, R.D., Charnes, A., and Cooper, W.W., "Models for estimating technical and scale efficiencies", *Management Science* 30 (1984) 1078–1092.
[10] Banker, R.D., and Cooper, W.W., "Validation and generalization of DEA and its uses", TOP 2, Sociedad Española de Estadistica e Investigación Operativa, Madrid, Spain, 1994, 249–297..
[11] Bardhan, I., "DEA and stochastic frontier regression approaches for evaluating the efficiency of public sector activities applications of public school performances in Texas", Ph.D. Thesis, The University of Texas at Austin, Graduate School of Business, 1995; Also available from Ann Arbor, Mich., University Microfilms, Inc.
[12] Bardhan, I., Bowlin, W.F., Cooper, W.W., and Sueyoshi, T., "Models for efficiency dominance in DEA. Part I: Additive models and MED measures; Part II: Free disposal hull (FDH) and Russell measure (RM) approaches", *Journal of the Operations Research Society of Japan* 39 (1996) 322–332; 333–344.
[13] Bardhan, I., Cooper, W.W., and Kumbhakar, S.C., "A simulation study of joint uses of data envelopment analysis and statistical regressions for production function estimation and efficiency evaluation", *Journal of Productivity Analysis* (1997) forthcoming.
[14] Brockett, P.L., Charnes, A., Cooper, W.W., Leamer, D.B., and Phillips, F.Y., "Information theory as a unifying statistical approach in marketing research", *European Journal of Operational Research* 83 (1995) 310–329.
[15] Brockett, P.L., and Golany, B., "Determining programmatic differences in DEA", *Management Science* 42 (1996) 466–472.
[16] Chaffai, M.E., "Estimating input-specific technical inefficiency: The case of the Tunisian banking industry", *European Journal of Operational Research* 98 (1997) 314–331.
[17] Charnes, A., and Cooper, W.W., *Management Models and Industrial Applications of Linear Programming*, John Wiley and Sons, Inc., New York, 1961.
[18] Charnes, A., and Cooper, W.W., "Programming with linear fractional functionals", *Naval Research Logistics Quarterly* 8 (1962) 181–186.
[19] Charnes, A., Cooper, W.W., Golany, B., Seiford, L., and Stutz, J., "Foundations of data envelopment analysis and Pareto–Koopmans empirical production functions", *Journal of Econometrics* 30 (1975) 91–107.
[20] Charnes, A., Cooper, W.W., and Rhodes, E., "Measuring efficiency of decision making units", *European Journal of Operational Research* 1 (1977) 429.

[21] Charnes, A., Cooper, W.W., and Schinnar, A., "An extended Cobb–Douglas form for use in production economics", in: T.J. Kaul and J.K. Sengupta (eds.), *Essays in Honor of Karl A. Fox*, North-Holland, Amsterdam, 1991.

[22] Charnes, A., Cooper, W.W., and Thrall, R.M., "A structure for classifying and characterizing inefficiencies in data envelopment analysis", *Journal of Productivity Analysis* 2 (1991) 197–237.

[23] Chilingerian, J.A., "Exploring why some physicians are more efficient: Taking DEA inside the hospital", in: A. Charnes, W.W. Cooper, A. Lewin and L. Seiford (eds.), *Data Envelopment Analysis: Theory, Methodology and Applications*, Kluwer Academic Publishers, Norwell, MA, 1994.

[24] Cook, W.D., and Kress, M., *Ordinal Information and Preference Structures* Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1992.

[25] Cooper, W.W., and Pastor, J.T., "Generalized efficiency measures (GEMs) and model relations for use in DEA", *Journal of Productivity Analysis* (submitted).

[26] Cooper, W.W., Thompson, R.G., and R.M. Thrall, "Introduction: Extensions and new development in DEA" *Annals of Operation Research* 66 (1996) 3–45.

[27] Cooper, W.W., Tone, K., Takamori, H., and Sueyoshi, T., "Data envelopment analysis: Survey and interpretations", *Communications of the Operations Research Society of Japan* (1994); in Japanese, English translations can be secured from the authors.

[28] de Borger, B., and Kerstens, K., "Radial and nonradial measures of technical efficiency: An empirical illustration for Belgian local government", *Journal of Productivity Analysis* 7 (1996) 41–62.

[29] Färe, R., Grosskopf, S., and Lovell, C.A.K., *The Measurement of Efficiency of Production*, Kluwer–Nijhoff Publishing, 1985.

[30] Ferrier, G.D., and Lovell, C.A.K., "Measuring cost efficiency in banking: Econometric and linear programming evidence", *Journal of Econometrics* 46 (1990) 229–244.

[31] Ganley, J.A., and Cubbin, J.S., *Public Sector Efficiency Measurement*, North-Holland, Amsterdam, 1992.

[32] Golany, B., "A note on including ordinal relations among multipliers to be used in data envelopment analysis", *Management Science* 34 (1988) 1029–1032

[33] Golany, B., and Thore, S., "The competitiveness of nations", in: W.W. Cooper, D. Gibson, F. Phillips, S. Thore and A. Whinston (eds.), *IMPACT: Essays in Honor of George Kozmetsky*, Kluwer Academic Press, Norwall, MA, 1997, forthcoming.

[34] Gong, B.H., and Sickles, R., "Finite sample evidence on the performance of stochastic frontiers and data envelopment analysis using panel data", *Journal of Econometrics* 51 (1990) 259–284.

[35] Gong, B.H., and Sickles, R., "Finite sample evidence on the performance of stochastic frontier models using panel data", *The Journal of Productivity Analysis* 1 (1989) 229–261.

[36] Green, R.H., Doyle, J.R., and Cook, W.D., "Preference voting and project ranking using DEA and cross-evaluation", *European Journal of Operational Research* 90 (1996) 461–472.

[37] Green, W.H., "The econometric approach to efficiency analysis", in: H.O. Fried, C.A.K. Lovell and S.S. Schmidt (eds.), *The Measurement of Productive Efficiency*, Oxford University Press, New York, 1993.

[38] Kumbhakar, S.C., "Estimation of input-specific technical and allocative inefficiency relative to stochastic production and cost estimations", *Oxford Economic Papers* 40 (1988) 535–549.

[39] Lovell, C.A.K., and Pastor, J.T., "Macroeconomic performance of sixteen Ibero-American countries over the period 1980–1981", Working paper, University of Georgia, Department of Economics, Athens, GA, 1994.

[40] Lovell, C.A.K., and Pastor, J.T., "Units invariant and translation invariant DEA models", *Operations Research Letters* (1995).

[41] Mendoza, M.N.F., "Essays in production theory: Efficiency measurement and comparative statistics." Ph.D. Thesis, University of British Columbia, Vancouver, British Columbia, Canada, 1989.

[42] Rhodes, E., "Data envelopment analysis and related approaches for measuring the efficiency of decision making units with an application to program follow through in U.S. public school education, Ph.D. Thesis Pittsburgh, PA, Carnegie Mellon University, School of Urban and Public Affairs, 1978; Also available from Ann Arbor, Mich., University Microfilms, Inc.

[43] Seiford, L.M., "A DEA bibliography (1978–1992)", in: A. Charnes, W.W. Cooper, A. Lewin and L. Seiford (eds.), *Data Envelopment Analysis: Theory, Methodology and Applications*, Kluwer Academic Publishers, Boston, 1994.

[44] Thompson, R.G., Dharmapala, P.S., and Thrall, R.M., "DEA sensitivity analysis of efficiency with applications to Kansas farming and illinois coal mining", in: A. Charnes, W.W. Cooper, A. Lewin and L. Seiford (eds.), *Data Envelopment Analysis: Theory, Methodology and Applications*, Kluwer Academic Publishers, New York, 1994.

[45] Thrall, R.M., "Duality, classification and slacks in DEA", *Annals of Operations Research* 66 (1996).

[46] Yu, G., Wei, Q., and Brockett, P.L., "A generalized data envelopment analysis model – A unification and extension of existing methods for efficiency analysis of decision making units", *Annals of Operations Research* 66 (1996).