NONPARAMETRIC EFFICIENCY ANALYSIS UNDER PRICE UNCERTAINTY

A FIRST-ORDER STOCHASTIC DOMINANCE APPROACH

TIMO KUOSMANEN
Helsinki School of Economics and Business Administration
P.O. Box 1210, FIN-00101 Helsinki, Finland

THIERRY POST
Erasmus University Rotterdam
P.O.B. 1738 3000 DR Rotterdam, The Netherlands

ABSTRACT

This paper extends the nonparametric approach to efficiency analysis to deal with uncertainty of input-output prices. We generalize the notion of economic efficiency to derive necessary and sufficient first-order stochastic dominance (FSD) efficiency conditions. Interestingly, the FSD conditions include as limiting cases the traditional conditions for economic efficiency and technical efficiency. Furthermore, we propose empirical tests for these FSD conditions, which require minimal assumptions concerning the preferences of the decision-maker and the statistical distribution of the prices. From operational point of view, the FSD conditions can be tested empirically using standard mathematical programming techniques. An empirical application to the Dutch electricity distribution sector illustrates the approach.

Key words: *nonparametric efficiency analysis, Data Envelopment Analysis, performance evaluation under uncertainty, stochastic dominance, electricity distribution sector.*

1. INTRODUCTION

The nonparametric approach to analyzing efficiency of production (Farrell, 1957; Afriat, 1972), sometimes dubbed as *Data Envelopment Analysis* (DEA) due to Charnes *et al.* (1978), has been credited for its ability to deal with multiple input and output variables while using only minimal assumptions about the relationship between the variables. A well-recognized limitation of the methodology is its deterministic nature. Since nonparametric analysis relies on comparison with extreme observations, the results can be extremely sensitive to sampling error and errors-in-variables. In the last decade, much research effort has been directed towards this problem. Approaches to account for the sampling errors (e.g. Simar and Wilson, 2000) have been well received in the literature. Some approaches also exist for dealing with the errors-in-variables (Varian, 1985; Olesen and Petersen, 1995; Cooper *et al.* 1998; and Post, 2001a) and new approaches are being developed (e.g. Cherchye *et al.*, 2000a; Kuosmanen *et al.*, 2000; Kuosmanen and Post, 2001a).

Sampling error and errors-in-variables are sources of *data uncertainty* faced by the analyst about the production process of the firms. A problem not duly recognized so far is *economic uncertainty*, i.e. uncertainty faced by the *firms* about the outcomes of the production process. In many industries, production represents a problem of optimization under uncertainty. For example, bank performance generally depends on the uncertain influence of uncontrollable external risk factors such as interest rates, foreign exchange rates and the business cycle. Still, the methodology invariably represents production as a problem of optimization under certainty.

Those few papers available on the subject address *technological uncertainty* i.e. uncertainty about the physical outcomes of the production process (e.g. Post and Spronk, 2000; Moschini, 2001). Another aspect of economic uncertainty is *price uncertainty.* In reality, input-output prices are rarely known with full certainty when production decisions are made. Prices are typically determined by demand and supply forces that are (partly) beyond the control of the firms, and in addition that are (partly) unpredictable for the firms. For example, in the classic Cournot (1838) oligopoly model, firms control their production plans, but prices are determined on the market. Under such circumstances, firms require perfect information both on demand responses, and production plans of the competitors, to perfectly anticipate (the Nash equilibrium) prices. Needless to say, firms rarely have all this information available for their decision making. Moreover, even if the nominal prices would be rigid e.g. due to long-term contracts or regulatory acts, unpredictable fluctuations in the uncontrollable environmental factors such as the 'business cycle', the interest rates, and the inflation guarantee that the real prices in terms of the opportunity cost typically involve uncertainties beyond hedging. Thus, evaluating ex ante efficiency of firms at the realized prices observed ex post by the analyst may be unfair to firms who actually did not have that information when they fixed their production plans.

This paper extends the notion of economic efficiency to include price uncertainty into efficiency analysis. Using the theory of first-order stochastic dominance (Hadar and Russell 1969), we derive necessary and sufficient conditions for efficiency under uncertainty. To preserve the nonparametric orientation, we develop tests that require minimal assumptions about the statistical distribution of the prices and the organizational objectives of the firms. Interestingly, these conditions include the traditional technical efficiency criteria as the limiting case where prices are unrestricted random variables. Similarly, the condition for profit efficiency (Debreu, 1951; Nerlove, 1965) is the limiting case where prices are known with complete certainty. In addition, we introduce empirical tests that can be implemented using standard mathematical programming techniques.

The remainder of the text is organized as follows. Section 2 generalizes the notion of efficiency and discusses its relationship to optimality. Section 3 discusses production analysis under price uncertainty. To reduce the information requirements, Section 4 introduces the first-order stochastic dominance conditions. Section 5 illustrates how the FSD conditions can be tested empirically. Section 6 illustrates the proposed approach by an application to Dutch electricity distribution sector. Finally, Section 7 draws conclusions and gives suggestions for future research.

2. EFFICIENCY AND OPTIMIZING BEHAVIOR

Our purpose is to evaluate ex ante efficiency of firms $j \in J \equiv \{1,...,n\}$ that employ a common technology, characterized by the production possibilities set $T \in \Re^q$, which involves q netputs indexed by $i \in Q \equiv \{1,...,q\}$ (positive netputs represent outputs, negative netputs represent inputs). Throughout the text, the observed netputs of the firms are represented by the matrix $Y \equiv (y_1 ... y_n)$ with $y_j \equiv (y_{1j} ... y_{qj})^T \in T$.

We define economic efficiency as the extent to which the firm has succeeded in maximizing its objective function $g : \Re^q \to \Re$ subject to the constraints imposed by the production set. In the spirit of Debreu (1951), Farrell (1957), and Nerlove (1965), we define firm $k \in J$ as overall economic efficient if and only if

(1) $$\underset{y \in T}{Max}\{g(y) - g(y_k)\} = 0.$$

This generalized definition conveniently combines all known alternative concepts of economic efficiency. For example, if the firm operates under conditions of perfect competition and full certainty about the netput prices $p_k \in \Re_+^q$, the objective function g boils down to profit at given netput prices, i.e. $g(y_k) = y_k p_k$. According to the definition of Nerlove (1965), firm k is said to be *profit efficient* if and only if it is not possible to achieve higher profit, given the technology and the prices, i.e.:

(2) $$\underset{y \in T}{Max}\{yp_k - y_k p_k\} = 0.$$

Note that the 'test statistic' (2) is equivalent to Debreu's (1951) "coefficient of resource utilization", expressed in difference rather than ratio form.[1] In the nonparametric literature, Varian (1984), Banker and Maindiratta (1988), and Färe and Grosskopf (1995), among others, have presented empirical tests and efficiency measures related to profit efficiency.

It is worth emphasizing that the general efficiency definition (1) is not limited to the unconstrained profit maximization in the case of the competitive firm. Indeed, the definition is equally well compatible to the cost minimization (i.e. output constrained profit maximization) and revenue maximization (i.e. input constrained profit maximization) objectives, which motivate the commonly used input and output oriented technical efficiency measures. Moreover, the objective function g need not be restricted to the class of linear functions, but may be non-linear like e.g. in the value efficiency approach of Halme *et al.* (1999).

[1] It is worth noting that Debreu proposed his ratio "coefficient" as a measure of Paretian (in-) efficiency in an entire economy, placing emphasis on welfare aspects, while Nerlove's analogous measure was intended for gauging profitability of a business enterprise. Unfortunately, it appears that the equivalence of the efficiency notions by Debreu and Nerlove has been ignored thus far. Debreu's notion is often (questionably) associated with Farrell's input and output oriented measures.

In practice, the objective function generally is not fully known. Already Debreu (1951) and Farrell (1957) expressed their concern about the ability to measure prices accurately enough to measure profit efficiency. For example, accounting data can give a poor approximation for economic prices (i.e. marginal opportunity costs), because of debatable valuation and depreciation schemes. Still, necessary and sufficient conditions for optimizing behavior can be obtained even if the objective function is not fully known. Suppose that the objective function g is known to belong to a particular family of functions G. In the nonparametric spirit (inspired in particular by Debreu and Farrell), a necessary condition is given by:

$$(3) \qquad \underset{y \in T}{Max}\left(\underset{g \in G}{Min} \ \{ g(y) - g(y_k) \} \right) = 0 .$$

That is, we apply the overall efficiency criterion to the 'most favorable' objective function included in the family G. In similar spirit, a sufficient condition for optimizing behavior could be obtained by evaluation at the 'least favorable' objective function:

$$(4) \qquad \underset{y \in T}{Max}\left(\underset{g \in G}{Max} \ \{ g(y) - g(y_k) \} \right) = 0 .$$

For illustration, consider the classic case of profit maximization as discussed above. Suppose that the exact netput prices are unknown to the efficiency analyst, but the prices are known to be constant and non-negative, with a strictly positive price for at least for one netput. Without loss of generality, we can phrase in terms of normalized prices. Specifically, we will use weighting vector $w = (w_1 \ \cdots w_q) \in \Re^q_+$ and normalize the prices such that the weighted sum of prices sums to unity, i.e. $wp_k = 1$. The family G thus reads

$$(5) \qquad G = \{ g : \Re^q \to \Re \mid g(y) = yr \, ; wr = 1; \, r \in \Re^q_+ \}.$$

Using this normalization, the necessary condition (3) becomes

$$(6) \qquad \underset{y \in T}{Max}\left(\underset{r \in \Re^q_+}{Min} \{ yr - y_k \, r \mid wr = 1 \} \right) = 0 .$$

The max-min 'test statistic' (6) effectively measures 'technical efficiency'. It can be interpreted as a reformulation of the min-max interpretation of Debreu's (1951) "coefficient", expressed as a normalized difference. Interestingly, as shown by Färe and Grosskopf (1997) and Chambers *et al.* (1998), the dual formulation of this statistic is the directional distance function:

$$(7) \qquad d(y_k, w, T) = \max_d \{ d \mid (y_k + dw) \in T \}.$$

As discussed by Chambers *et al.* (1998), this directional distance function contains all known technical efficiency gauges as its special cases, including the Farrell input and output oriented measures.

The technical efficiency measure (profit efficiency at 'most favorable' prices) gives a *necessary* condition for profit efficiency. In similar vein, Kuosmanen and Post (2001b) derived *sufficient* conditions for economic efficiency using the 'least favorable' prices, focusing on the notion of cost efficiency. Moreover, Kuosmanen and Post demonstrated how incorporating additional price information, which may come in incomplete form, can narrow the gap between the necessary and sufficient conditions. Some of the developments in Section 4 build on these insights.

3. EX ANTE EFFICIENCY ANALYSIS UNDER UNCERTAINTY

In contrast to the conventional approach, we assume that the firm under evaluation does not know the netput prices until *after* the production plan is fixed. More specifically, we assume that the prices are random variables with domain $D \subseteq \Re_+^q$ and joint distribution function $F : D \to [0,1]$. Both D and F may harmlessly be assumed as firm specific. The price distribution effectively reflects exogenous market conditions. Different firms can face different market conditions, and comparing the profit of firms with different conditions is not fair to firms that face unfavorable conditions. Therefore, the tests below compare the profit of the evaluated firm with the hypothetical profits of other firms at the price distribution of the evaluated firm. Still, for notational convenience, we will not use a firm specific index for D and F.

For convenience, we continue the use of normalized prices (as above). Specifically, we assume the following structure for the price domain:

(8) $$D \equiv \left\{ p \in \Re_+^q \middle| Ap \geq 0; wp = 1 \right\}.$$

The domain D contains all non-negative price vectors that satisfy l linear inequalities characterized by the $l \times q$ matrix $A \equiv \left(A_1 \cdots A_q\right)$ and the normalization constraint. The assumed form of the price domain may turn out impractical or irrelevant in some empirical cases. However, in many cases, price information does take the form of linear inequalities, or linear inequalities can give a good approximation for more complicated price structures. In addition, the limiting cases of complete price information on the one hand and no price information whatsoever on the other, are special cases of this more general price domain, as shown in more detail below. Finally, the assumed structure of the price domain is computationally convenient, and can easily be included in Linear Programming models (see Section 5 below).[2]

To include price uncertainty in efficiency analysis, we formulate firm objectives in terms of the Von Neumann - Morgenstern (1944) Expected Utility

[2] Note that apart from the normalization constraint, our price domain is reminiscent of so-called assurance region (AR) or weight-restricted (WR) DEA models (first introduced by Thompson *et al*., 1986, 1990; see Allen *et al*., 1997, for a detailed survey). Those models enrich original DEA models with managerial judgement by restricting the dual multipliers to convex polyhedral cones. For a systematic framework for using polyhedral convex cones to represent *price* information for measuring economic efficiency, we refer to Kuosmanen and Post (2001b).

Theory.[3] More specifically, assume that the evaluated firm selects its production plan to maximize the expected value of a Von Neumann - Morgenstern utility function $U:\Re^1 \rightarrow \Re^1$ defined over profit realizations, subject to the technological constraints and subject to the distribution of prices. Hence, we effectively assume

$$(9) \qquad g(y) = E(U(yp)) = \int_D U(yp)\partial F(p) .$$

In microeconomic theory, McCall (1969) and Sandmo (1974), among others, have considered (9) as an appropriate firm objective under price uncertainty.

The purpose of our ex ante efficiency analysis is to judge whether or not the observed choice of the production variables was rational, maximizing the expected utility function of the firm. Thus, analogous to the traditional economic efficiency condition (4), the firm k maximizes expected utility if and only if:

$$(10) \qquad \max_{y \in T} \left\{ \int_D U(yp)\partial F(p) - \int_D U(y_k p)\partial F(p) \right\} = 0 .$$

If the preference structure (U) and the price distribution (F) are known to the analyst, then (10) provides a ready test statistic. Unfortunately, such detailed information is often very difficult to obtain for the purposes of external efficiency assessment. Hence, we intend to resort to minimal distribution and preference assumptions, so as to preserve the nonparametric orientation. We will next derive necessary and sufficient conditions in the spirit of (3) and (4) that apply for large families of utility functions and distribution functions.

4. FIRST-ORDER STOCHASTIC DOMINANCE APPROACH

To allow for a wide class of utility functions, we propose to employ decision criteria from the theory of stochastic dominance (henceforth SD), which is a theory of choice under uncertainty that has seen considerable theoretical development and empirical application in economics in the last decades (see e.g. Levy, 1992). SD makes no assumptions with respect to the specific functional form for decision-maker

[3] The expected utility framework has a number of well-known limitations. For example, many firm decisions are taken by a group of individuals, and group preferences may not always satisfy the transitivity axiom required for the existence of a Von Neumann-Morgenstern utility function. Nevertheless, the expected utility framework remains a standard analytical tool, mainly because of its mathematical tractability. Furthermore, there are many firms in which essentially one person makes the decisions, and there are presumably many firms in which preferences are sufficiently similar within the group of decision-makers to guarantee the existence of a group preference function. Finally, we emphasize that the below tests also apply for a whole range of non-expected utility theories of choice behavior under uncertainty; FSD is widely accepted as a choice criterion in non-expected utility theories. Moreover, it is well supported by empirical evidence. As Starmer (2000) summarizes, first order stochastically dominated options are usually not selected when stochastic dominance is transparent.

preferences. Rather, stochastic dominance relies on assumptions about the general characteristics of decision-maker preferences. There are three progressively stronger preference assumptions that are employed in the SD literature. They lead directly to the first-, second-, and third-order SD criteria. First-order SD (henceforth FSD; Hadar and Russell, 1969) assumes decision-makers are non-satiable preferring more over less, i.e. $U'(\cdot) \geq 0$. Second-order SD assumes, in addition to non-satiation, that decision-makers are risk aversive, that is, they dislike uncertainty, i.e. $U''(\cdot) \leq 0$. Finally, third-order SD adds to non-satiation and risk aversion the assumption that decision-makers have decreasing absolute risk aversion, i.e. their dislike for absolute uncertainties decreases as the levels of the outcomes increase, $U'''(\cdot) \geq 0$. Higher order criteria involve more discriminating power than lower order ones, because they induce a larger reduction of the set of not-dominated choice alternatives. However, that power has to be balanced against the stringency of the preference assumptions imposed. In general, striking that balance requires a careful consideration of the structure and the context of the decision problem considered.

In this paper, we focus on FSD conditions for firm efficiency, and we leave generalizations towards higher order conditions for future research. In general, a distribution function $G_1(x)$ stochastically dominates another distribution function $G_2(x)$ by first-order, if and only if the former involves a probability of exceeding any value for x greater than or equal to that of the latter (with inequality for at least one value for x), i.e.:

$$(11) \qquad G_2(x) \succ G_1(x) \;\; \forall x \in \Re^1 .$$

Here '$\succ$' is used to denote the condition that '$\geq$' holds for the entire domain with '$>$' holding for at least some values of the domain.

In the production problem discussed above, prices are considered random variables. Therefore, profit is a random variable with the following distribution function:

$$(12) \qquad P\left(y_k p \leq p\right) = \int_{\{p \in D | y_k p \leq p\}} \partial F(p) .$$

Applying the FSD rule (11) to this distribution function, we find that the evaluated firm is stochastically dominated by first-order by some feasible production plan, if and only if:

$$(13) \qquad \exists y \in T : \int_{\{p \in D | y_k p \leq p\}} \partial F(p) \succ \int_{\{p \in D | yp \leq p\}} \partial F(p) \quad \forall p \in \Re^1 .$$

Consequently, a *necessary* condition for optimal production is:

$$(14) \qquad \underset{y \in T}{Max} \left[\underset{p}{Min} \left\{ \int_{\{p \in D | y_k p \leq p\}} \partial F(p) - \int_{\{p \in D | yp \leq p\}} \partial F(p) \right\} \right] = 0 .$$

Similarly, the evaluated firm stochastically dominates all feasible production plans if and only if:

(15) $$\int_{\{p\in D | yp \le p\}} \partial F(p) \succ \int_{\{p\in D | y_k p \le p\}} \partial F(p) \quad \forall p \in \Re^1; y \in T.$$

Consequently, a *sufficient* condition for optimal production is:

(16) $$\underset{\substack{y\in T\\ p}}{Max}\left\{ \int_{\{p\in D | y_k p \le p\}} \partial F(p) - \int_{\{p\in D | yp \le p\}} \partial F(p) \right\} = 0.$$

These rules do not require preference assumptions in addition to non-satiability. However, they do require information on the distribution function *F*. Next, we consider relaxing this condition further to allow for an arbitrary distribution.

In the spirit of (3), the following *necessary* condition for (14) can be obtained for an arbitrary distribution:

(17) $$\underset{y\in T}{Max}\left[\underset{p\in D}{Min}\{yp - y_k p\} \right] = 0.$$

In other words, optimal behavior requires that netput vectors that yield higher profit than the evaluated firm at all possible price vectors are not attainable.

Similarly, the following *sufficient* condition for (16) can be obtained for an arbitrary distribution:

(18) $$\underset{\substack{y\in T\\ p\in D}}{Max}\{yp - y_k p\} = 0.$$

In other words, if the evaluated firm yields the maximum profit at all possible price vectors, it necessarily maximizes expected utility.

Interestingly, duality theory can express conditions (17) and (18) in terms of the directional distance function (3):

THEOREM 1 *The necessary FSD condition* (17) *can equivalently be defined as*

(19) $$d(y_k, w, W(T, D)) = 0,$$

where

(20) $$W(T, D) \equiv \{ y | y \le y' - zA; z \in \Re^l_+; y' \in T \}$$

is a 'price-augmented' production set that represents profit levels that are feasible at *<u>all</u> prices from D.*

PROOF Problem (17) embeds the following Linear Programming problem, using y^* for the optimal solution to the maximization problem $\underset{p\in D}{Min}\{y^* p - y_k p)\}$. The dual formulation of this problem is $\underset{\substack{d\in\Re_+ \\ z\in\Re_+^l}}{Max}\{d \,|y_k + dw \leq y^* - zA\}$. Substituting this problem in (17), we find $\underset{\substack{y\in T \\ d\in\Re_+}}{Max}[d \,|y_k + dw \leq y - zA] = 0$, or alternatively $d\,(y_k, w, W(T,D)) = 0$.

Q.E.D.

THEOREM 2 *The sufficient FSD condition* (18) *can equivalently be defined as*

(21) $\quad d\,(y_k, w, V(T,D)) = 0$,

where

(22) $\quad V(T,D) \equiv \{y | y_i \leq y'_i + zA_i \;\; i\in Q, y'\in T, \forall z\in \Re^l_+,\}$

is a 'price-augmented' production set that represents profit levels that are feasible at <u>*some*</u> *prices from D.*

PROOF Using y^* for the optimal solution, problem (18) embeds $\underset{p\in D}{Max}\{\, y^* p - y_k p\}$. The dual formulation of this problem is $\underset{\substack{d\in\Re_+ \\ z\in\Re_+^l}}{Min}\{d \,|y_k + d\, w \geq y^* + zA\}$, or equivalently $\underset{d\in\Re_+}{Max}\{d \,|y_{ik} + d\, w_i \;\leq y_i^* + zA_i \;\; i\in Q, \forall z\in \Re^l_+\}$. Substituting this problem in (18), we find $\underset{\substack{y\in T \\ d\in\Re_+}}{Max}\left[d \,|y_{ik} + d\, w_i \;\leq y_i + zA_i \;\; i\in Q, \forall z\in \Re^l_+\right] = 0$, or alternatively
$d\,(y_k, w, V(T,D)) = 0$. *Q.E.D.*

The FSD conditions involve some interesting limiting cases. Specifically, if the price domain is restricted to a unique price vector, i.e. $D = p_k$, both the necessary condition (17) (or (19)) and the sufficient condition (18) (or (21)) are equivalent to the condition for Nerlovian profit efficiency (2). By contrast, if the relative prices are completely unrestricted, i.e. $A = 0$, we find $W(T,D) = T$, and the necessary FSD condition (17) (or (19)) is equivalent to the technical efficiency criterion in terms of the directional distance function (7).

Figures 1, 2 and 3 graphically illustrate the FSD conditions for a single-input single-output technology. Figure 1 illustrates the price domain used in this example as represented by the polyhedral convex cone *D*. It represents the information that the output price is at least 0.5 times the input price, and at most 2 times the input price, i.e. we can set $A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$.

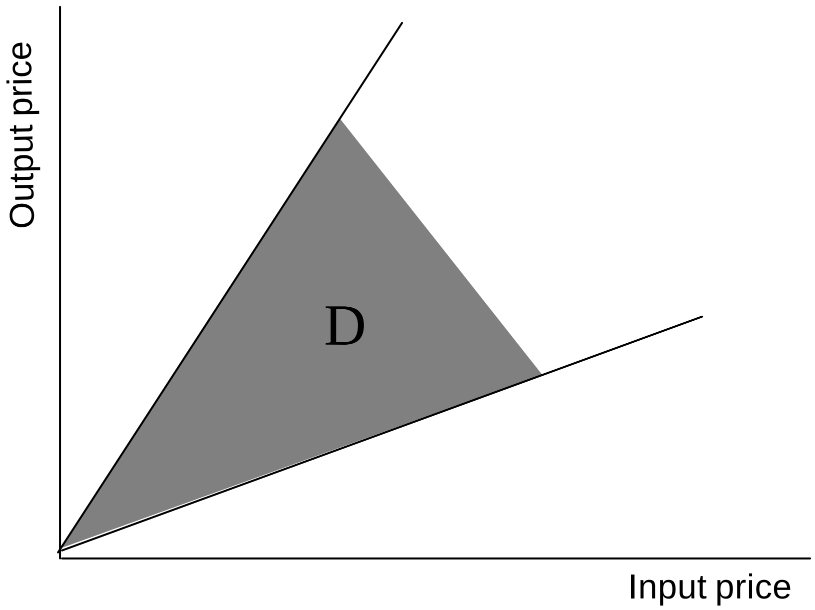

Figure 1: Example of price domain as represented by polyhedral convex cone *D*.

Figure 2 displays the production set T (the hypograph of the curved line) and the augmented production set $W(D,T)$ (the shaded area). All production vectors in the interior of the latter set *necessarily* are stochastically dominated by first-order, as T contains production vectors that yield higher profit at *all* prices in *D*.

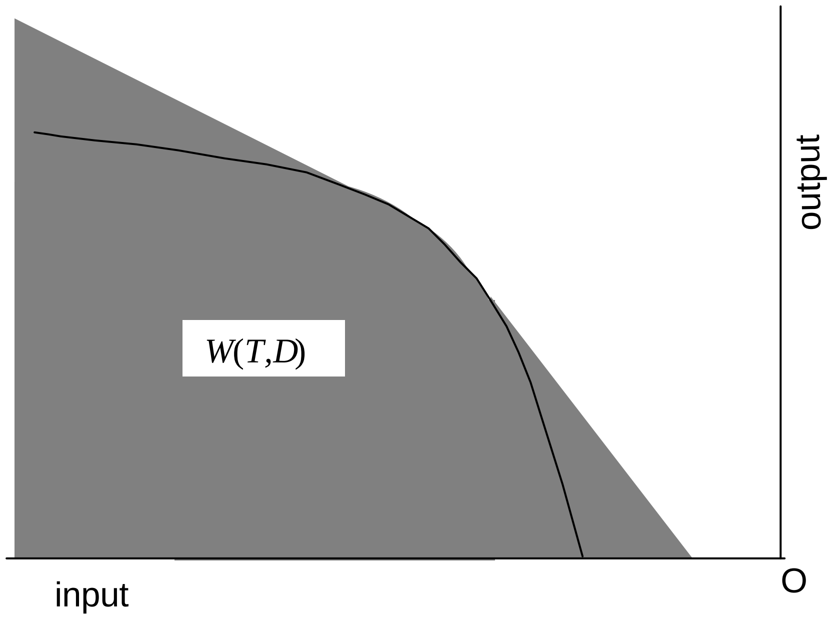

Figure 2: Example of the 'price-augmented' production set *W*(*D*,*T*) containing all production vectors that necessarily are dominated by FSD.

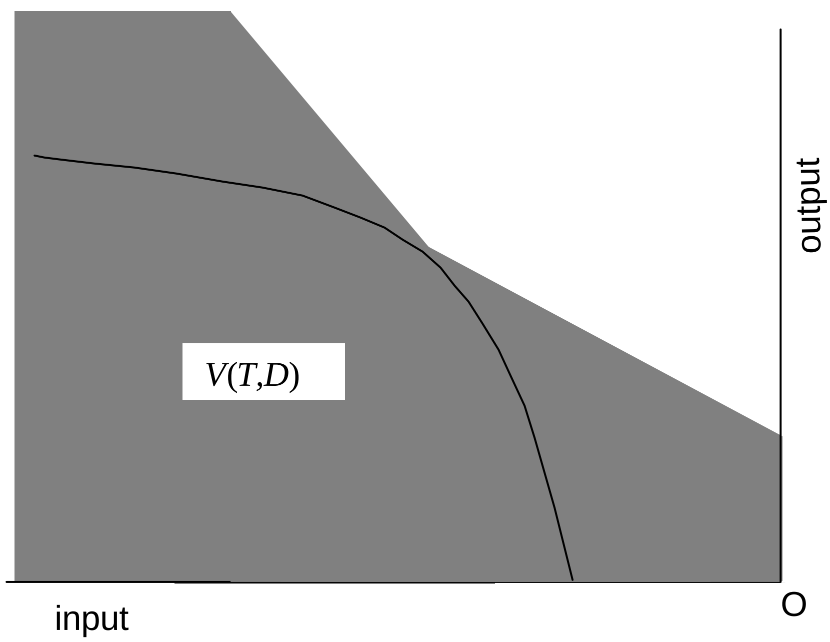

Figure 3: Example of the 'price-augmented' production set *V*(*D*,*T*) containing all production vectors that possibly are dominated by FSD.

Similarly, Figure 3 displays the augmented production set $V(D,T)$ (the shaded area). All production vectors in the interior of this set *possibly* are stochastically dominated by first-order, as T contains production vectors that yield higher profit at *some* prices in D. In this example, *no* production vector can be classified as stochastically not-dominated by first-order; the entire production set is enveloped by $V(D,T)$.

5. EMPIRICAL TESTS

Resorting to the above FSD conditions can reduce the information requirement for the preference structure and the price distribution. However, information on the production set is still required. Unfortunately, one typically faces imperfect information on the production possibilities in the empirical analysis of efficiency. To undertake empirical tests, however, we can approximate the true production set by an empirical production set constructed from empirical data.

Various approximations can be employed, depending on the assumptions imposed on the production set and the data used for approximating it. In the most standard approach due to Afriat (1972) and Banker *et al.* (1984), three key assumptions are imposed. Firstly, the production set is assumed to contain all observations, i.e. $y_j \in T \;\; \forall j \in J$. Furthermore, netputs are assumed free disposable, i.e. $T = T - \Re^q_+$. Finally, the production set is assumed convex, i.e. $T = co(T)$.

Following the so-called *minimum extrapolation* principle (Banker *et al.*, 1984), the smallest set in netput space that satisfies the above assumptions is the *convex monotone hull* (henceforth CMH) of observations, formally defined as:

$$(23) \qquad \hat{T}_{CMH} = \left\{ y \middle| y \le l^T Y;\; l^T e = 1;\; l \in \Re^n_+ \right\},$$

where $e \equiv (1 \ldots 1)^T$ denotes the nx1 unit vector. This empirical production set has become the standard approximation for T in the nonparametric literature. For example, it was used by Afriat (1972) and Banker *et al.* (1984), among others, for testing and measuring technical efficiency. In addition, Varian (1984), and Banker and Maindiratta (1988), among others, used it for testing and measuring profit efficiency.

Attractively, the CMH is that it can be used for assessing profit efficiency (2) even if the true production set does not satisfy free disposability and convexity. If the true production set violates free disposability or convexity, the CMH can contain netputs vectors outside the production set. However, profit is linear and increasing netputs. Therefore, for each production vector in the CMH and each possible price vector, the data set necessarily contains an observation that yields at least as much profit. Consequently, free disposability and convexity are 'harmless' assumptions; they do not interfere with the analysis of profit efficiency:

$$(24) \qquad \underset{y \in T}{Max}(yp) = \underset{j \in J}{Max}(y_j p) \;\; \forall p \in D.$$

This insight was employed by e.g. Varian (1984) and Banker and Maindiratta (1988).

However, under uncertainty, the CMH can approximate the production set only if the latter truly is convex. This is because the data set generally does not contain observations that yield at least as much profit, at *all* prices from the price domain, for each for each production vector in the CMH. In terms of the necessary condition (17): $\underset{y\in \hat{T}_{CMH}}{Max}\left[\underset{p\in D}{Min}(yp)\right]$ need not equal $\underset{j\in J}{Max}\left[\underset{p\in D}{Min}(y_j p)\right]$.

Therefore, the goodness of the CMH depends on whether the production set truly is convex or not. Convexity requires the production technology to exhibit particular theoretical characteristics such as non-increasing marginal products and non-increasing marginal rates of substitution and transformation (Madden, 1986). Unfortunately, there is little empirical evidence on these theoretical characteristics (see e.g. Cherchye *et al.*, 2000b,c, for further discussion). In case of violations, an alternative empirical production set is required. A number of alternative sets that do not require convexity have been proposed in the nonparametric literature, including Deprins *et al.* (1984), Tulkens (1993), Petersen (1990), Bogetoft (1996), Post (2001b,c), and Kuosmanen (2001).

Relaxing convexity, the smallest subset in netput space consistent with the assumptions of containment of observations and free disposability is the so-called *Free Disposable Hull* (henceforth FDH; Afriat, 1972; Deprins *et al.*, 1984; Tulkens, 1993). More formally, the FDH can be represented by:

(25) $\hat{T}_{FDH} = \{y \mid y \leq y_j ;\, j \in J\}.$

In contrast to convexity, free disposability can be imposed without harm. This is because observed netput vectors y_j yields at least as much profit, at all possible prices, as an arbitrary feasible netput vector $y : y \leq y_j$. Therefore, the following equality applies:

(26) $\underset{y\in \hat{T}_{FDH}}{Max}\left[\underset{p\in D}{Min}(yp)\right] = \underset{j\in J}{Max}\left[\underset{p\in D}{Min}(y_j p)\right].$

Substituting the FDH set for the true production set in the necessary FSD condition
(19) gives the following empirical condition:

(27) $d\,(y_k, w, W(\hat{T}_{FDH}, D)) = 0 \Leftrightarrow$

$$\underset{j\in J}{Max}\left[\underset{\substack{d\in \Re_+ \\ z\in \Re_+^l}}{Max}\left\{d \,\middle|\, y_k + d\,w \leq y_j - zA\right\}\right] = 0 .$$

If the observations are contained within the production set, this measure gives a necessary condition for optimal production under uncertainty.

Similarly, substituting the FDH set for the true production set in the sufficient FSD condition (18) gives the following empirical condition:

(28) $$\underset{\substack{y \in \hat{T}_{FDH} \\ p \in D}}{Max} \{yp - y_k p\} = 0 \Leftrightarrow$$

$$\underset{j \in J}{Max}\left[\underset{p}{Max}\left\{y_j p - y_k p \,\middle|\, Ap - Bw \geq 0;\, p^T w = 1\right\}\right] = 0$$

Since the FDH set typically is smaller than the true production set, this condition does not give a sufficient condition for optimal production. However, it does give a sufficient condition for dominance over all observations in the sample.

Checking the conditions (27) and (28) involves little computational burden; one can enumerate the solutions to the n embedded Linear Programming problems ($j = 1, \cdots, n$) and subsequently compute the maximum of those solutions. The following two observations can further reduce the computational burden in efficiency tests: 1) We can already reject the null hypothesis of full efficiency if the optimal solution to the embedded LP problem exceeds 0 for any $j \in J$. Thus computation can terminate immediately if such a solution is found. 2) We can focus attention to the non-dominated (i.e. FDH efficient) $j \in J$ because a dominated firm cannot solve the second maximization problem in (27) or (28). As a preliminary step, it is straightforward to eliminate the dominated firms from the reference set by enumeration (see e.g. Tulkens, 1993).

6. EMPIRICAL APPLICATION

Nonparametric efficiency analysis has seen numerous applications in regulating the electricity sector in various countries. For example, the regulator of the Dutch electricity sector recently applied the approach for a system of price cap regulation. For the purpose of illustration, we performed an application of our approach to 1999 data for the electricity distribution sector in the Netherlands, which included 18 electricity-distributing units or EDUs at that point in time. For the sake of simplicity, we used a simplified representation of the production technology that involves a single input (=negative netput), operating expenses in thousands of Dutch guilders (1 guilder $\approx$ 0.45 Euro), and two outputs, (1) number of small customers and (2) number large customers.[4] Table 1 lists the full data set.

---

[4] For a full description of the variables and the EDUs in our model, we refer to the homepage of the regulatory office for the Dutch electricity sector (www.dte.nl). Kuosmanen *et al*. (2000) used the same data set, but a slightly different set of variables. Specifically, to correct for differences in electricity consumption between consumers, Kuosmanen *et al*. included 'gigawatt hours distributed' as an output variable in addition to the number of customers. By contrast, we distinguish between small and large customers. This distinction allows for demonstrating the effect of incomplete price information, as is elaborated below.

Table 1 Data set

EDU	Operating expenses	Small customers	Large customers
1	9122	44625	690
2	196685	1048076	4020
3	8857	84236	497
4	97953	487479	3242
5	513819	2589620	19125
6	47707	187160	1190
7	12403	45599	936
8	10874	97405	1104
9	2313	11598	2
10	6832	48455	168
11	2765	29010	150
12	16667	115335	862
13	5311	37841	169
14	2416	19799	236
15	183441	921326	7753
16	185323	881196	6174
17	97912	401213	3906
18	7710	44974	273

We stress that this application is for illustrative purposes only. A more realistic representation of the production technology would account for differences across EDUs in e.g. quality of service, geography, and network architecture. In addition, a sound study would improve the power of the tests by using additional data e.g. from time series or from international comparisons.

In 1999, EDUs in the Netherlands were regional monopolists, and price uncertainty was not very important. By contrast, within a couple of years, the market for electricity distribution will be fully liberalized, and EDUs will have to compete for customers. Apart from decreasing prices, competition can be expected to increase uncertainty about prices. Therefore, it is interesting to 'simulate' the effect of uncertainty on efficiency (even if the distributors did not face substantial uncertainty during the sample period).

We tested the empirical necessary FSD conditions for efficiency relative to the CMH and relative to the FDH. We first considered the case where no information on the price domain is available (i.e. A=0). The test results are displayed in Table 2; 10 EDUs are classified as inefficient relative to the CMH, while only 2 EDUs are classified as inefficient relative to the FDH. This difference may signal a violation of the convexity assumption e.g. because of economies of scale (recall that convexity is not harmless under uncertainty!). Alternatively, the difference may signal a lack of discriminating power in small samples associated with the FDH.

As discussed above, including information on the price domain can increase discriminating power. One possibility is to use the insight that large customers yield more revenues than small customers do. We do not have detailed price and volume data for the two groups of customers. Still, based on our knowledge of this sector, we believe the average revenue per large customer equals at least 100 times the average revenue per small customer. We therefore considered a third model that includes this information ($A = (0 \quad -100 \quad 1)$) in the FDH model. Interestingly, based on the additional price information, two additional EDUs were classified as inefficient. For example, EDU 2 is efficient relative to the FDH set; no other EDU serves more small customers and more large customers at less cost. Still, if one large customer represents the revenue of at least 100 small customers, then EDU 15 necessarily involves *higher profit* than EDU 2. In other words, EDU 2 is in the interior of the 'price-augmented' set $W(D,\hat{T}_{FDH})$.

Table 2 Test results

EDU	**X = Classified as inefficient**		
	CMH A=0	**FDH A=0**	**FDH A=(0,-100,1)**
1	X		
2			X
3			
4	X		
5			
6	X		
7	X	X	X
8			
9			
10	X		
11			
12	X		X
13	X		
14			
15			
16	X	X	X
17	X		
18	X		

While various EDUs fail to pass the necessary FSD condition, no EDU satisfies the sufficient FSD condition for efficiency (even for the CMH and with the price information). This suggests that the sufficient condition can discriminate between EDUs only if additional price information is introduced.

7. CONCLUDING REMARKS

We have presented an approach based on first-order stochastic dominance that extends the traditional framework for nonparametric efficiency analysis, so as to deal with uncertainty related to input-output prices. We have derived necessary and sufficient first-order stochastic dominance conditions for ex ante optimizing behavior. Interestingly, the necessary FSD efficiency condition contains as limiting cases the standard economic efficiency and technical efficiency conditions. We find this a quite encouraging result, since this finding suggests that the conventional efficiency measures remain meaningful under uncertainty.

In addition, we have discussed empirical conditions that can test whether observed behavior is consistent with optimal behavior under uncertainty. These conditions can be checked using standard Linear Programming. A complication for testing FSD empirically is that convexity for the production set cannot be assumed without harm. Therefore, if violations of convexity are anticipated, the FDH is preferred to the CMH, which is typically used for testing profit efficiency in case of full certainty. From a methodological point of view, the approach adopted in this paper provides a powerful economic justification for the FDH approximation of the technology set; the production possibilities simply can not harmlessly be assumed convex in a case of price uncertainty. Nevertheless, it should be understood that the FSD approach is not limited to particular empirical production sets. Alternative empirical approximations or estimations could be used, depending on the purposes of the application and the available information on the production relationships.

We introduced *tests* for efficiency that can classify firms as efficient or inefficient. We have not proposed *degree measures* for efficiency. Most of the current efficiency measures are directly or indirectly related to a specific economic objective function (e.g. a cost function or profit function). There are no generally accepted objective functions in case of uncertainty (apart from general notions like shareholder value or risk-adjusted return on capital which are too general to operationalize), and we have circumvented the specification of an objective function by resorting to general stochastic dominance conditions. Therefore, the test results should not be interpreted as degree measures.

We see this paper as a first step towards developing a framework for empirical production analysis that applies (even) if production can not be described by the traditional theory of the competitive firm. In this respect, the many exiting routes for future research are still open:

1. *Higher-order stochastic dominance rules.* For the sake of generality, we deliberately avoided assumptions that could be considered restrictive in any sense. Resorting to first order stochastic dominance allowed us to effectively reduce the information requirement on the utility function and price distributions. However, in principle nothing prevents one to apply higher-order stochastic dominance rules. Higher order criteria involve more discriminating power than lower order ones, because they induce a larger reduction of the set of not-dominated choice alternatives. Nevertheless, that power has to be balanced against the stringency of the preference imposed assumptions (e.g. risk aversion for second-order stochastic dominance and decreasing absolute risk aversion for third-order stochastic dominance). Still, extending the efficiency criteria towards higher-order stochastic dominance criteria is an interesting challenge for future research.

2. *Recovering the technology*. We have discussed testing efficiency of individual firms, given a specific theoretical or empirical production set. However, the problem could be easily reversed to test properties of the production set, given the set of efficient firms, as e.g. in Hanoch and Rotschild (1972) and Varian (1984). A follow-up paper (Cherchye *et al.*, 2000c) provides a first step in this direction.

3. *Production uncertainty.* We relaxed the standard assumption of full certainty for the input-output *prices*. However, we adhered to the standard assumption that input-output *quantities* are fully controllable and certain. However, it would also be interesting to consider a reversed setting á la Bertrand (1883) oligopoly model, where prices are the controlled decision variables, while quantities are determined

on the market and may involve uncertainty. Investigating this reversed setting is left for future research.

4. *Data uncertainty.* We have focused on including *economic uncertainty* for the firms. For simplicity, we abstracted from *data uncertainty* faced by analyst due to errors-in-variables and sampling error. As discussed in the introduction, a number of approaches currently are available for accounting for data uncertainty. Future research could focus on integrating these approaches with our approach to economic uncertainty, so as to account for both uncertainty for the firm and for uncertainty for the analyst.

5. *Endogenous prices*. The traditional framework does not account for uncertainty, and we have tried to include it in the analysis. Another phenomenon not accounted for in the traditional framework is *imperfect competition*. In many cases, the production plans of individual firms affect the market prices. Note that price endogeneity and price uncertainty often occur simultaneously. For example, under endogenous prices a particular production plan can be associated with different price equilibria, which immediately implies ex ante price uncertainty. Cherchye *et al.* (2000c) provides a first step towards including imperfect competition in the analysis.

REFERENCES

Afriat, S. (1972): Efficiency Estimation of Production Functions, *International Economic Review* 13(3), 568-598.

Allen, R., A. D. Athanassopoulos, R.G. Dyson and E. Thanassoulis (1997): Weight Restrictions and Value Judgements in DEA: Evolution, Development and Future Directions, *Annals of Operations Research* 73, 13-34.

Banker, R.D., Charnes, A., 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 A. Maindiratta (1988): Nonparametric analysis of technical and allocative efficiencies in production, *Econometrica* 56 (6), 1315-1332.

Bertrand, J. (1883): Théorie Mathématique de la Richesse Sociale, *Journal des Savants*, 499 - 508

Bogetoft, P. (1996): DEA on Relaxed Convexity Assumptions, *Management Science* 42, 457-465.

Chambers, R. G., Y. Chung, and R. Färe (1998): Profit, Directional Distance Functions, and Nerlovian Efficiency, *Journal of Optimization Theory and Applications* 98(2), 351-364

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

Cherchye, L., T. Kuosmanen and T. Post (2000a): New Tools for Dealing with Errors-in-Variables in DEA, Discussion Paper D/2000/2020/6, Centre for Economic Studies, Catholic University Leuven.

Cherchye, L., T. Kuosmanen and T. Post (2000b): Why Convexify? An assessment of convexity assumptions in DEA, Helsinki School of Economics and Business Administration, Working Paper W-270.

Cherchye, L., T. Kuosmanen, and G. T. Post (2000c): Nonparametric Production Analysis under Alternative Price Conditions, Working Paper

Cooper, W.W., Z.M. Huang, V. Lelas, S. Xi and O.B. Olesen (1998): Chance constrained programming formulations for stochastic characterizations of efficiency and dominance in DEA, *Journal of Productivity Analysis* 9(1): 53-79.

Cournot, A. (1838): *Recherches sur les Principes Mathématiques de la Théorie des Richesses.* English edition: *Researches into the Mathematical Principles of the Theory of Wealth*, Edited by N. Bacon, New York, Macmillan

Debreu, G. (1951): The Coefficient of the Resource Utilization, *Econometrica* 19(3), 273-292.

Deprins, D., D. Simar, H. Tulkens (1984): Measuring Labor Efficiency in Post Offices, p. 243 - 267 in M. Marchand, P. Pestieau and H. Tulkens (Eds.), *The Performance of Public Enterprises: Concepts and Measurements*, North Holland.

Färe, R., and S. Grosskopf (1995): Non-Parametric Tests of Regularity, Farrell Efficiency, and Goodness of Fit, *Journal of Econometrics* 69, 415-425.

Färe, R., and S. Grosskopf (1997): Profit efficiency, Farrell decompositions and the Mahler inequality, *Economics Letters* 57, 283-287.

Farrell, M. J. (1957): The Measurement of Productive Efficiency, *Journal of the Royal Statistical Society* Series A 120 (3), 253-281.

Halme, M,.T.Joro, P.Korhonen, S.Salo and J.Wallenius (1999):A Value Efficiency Approach to Incorporating Preference Information in Data Envelopment Analysis, *Management Science* 45(1), 103-115

Hadar, J. and W. R. Russell (1969): Rules for Ordering Uncertain Prospects, *American Economic Review* 59, 25-34.

Levy, H. (1992): Stochastic dominance and expected utility: survey and analysis, *Management Science* 38(4), 555-593.

Kuosmanen, T. (2001): DEA with Efficiency Classification Preserving Conditional Convexity, *European Journal of Operational Research* 132(2), 83-99

Kuosmanen, T. and G.T. Post (2001a): Non-parametric tests of firm efficiency in case of errors-in-variables: Efficiency depth approach, Erasmus Research Institute of Management (ERIM) Working Paper ERS-2001-06-F&A

Kuosmanen, T. and G.T. Post (2001b): Measuring Economic Efficiency under Imperfect Price Information: With an Application to European Commercial Banks, *European Journal of Operational Research* (to appear)

Kuosmanen, T., G.T. Post, and S. Scholtes (2000): Testing for productive efficiency with errors-in-variables: With an application to the Dutch electricity sector, Working Paper.

Madden, P. (1986): Concavity and Optimization in Microeconomics, Oxford : Blackwell.

McCall, J.J. (1969): Competitive Production for Constant Risk Utility Functions, *Review of Economic Studies* 34, 417-420.

Moschini, G.-C. (2001): Production risk and the estimation of ex-ante cost functions, *Journal of Econometrics* 100(2), 357-380

Neumann, J. Von, and O. Morgenstern (1944): *Theory of Games and Economic Behavior*, Wiley, New York.

Nerlove, M. (1965): *Estimation and Identification of Cobb-Douglas Production Functions*, Rand McNally, Chicago, Illinois.

Olesen, O.B., and N.C. Petersen (1995): Chance Constrained Efficiency Evaluation, *Management Science* 41, 442-457.

Petersen, N. C. (1990): Data Envelopment Analysis on a relaxed set of assumptions. *Management Science* 36 (3), 305-314.

Post, G.T. (2001a): Performance benchmarking in stochastic environments using Mean-Variance Data Envelopment Analysis, *Operations Research*, March-April 2001.

Post, G. T. (2001b): Transconcave Data Envelopment Analysis, *European Journal of Operational Research* 132(2), 131-146

Post, G.T. (2001c): Estimating Non-convex Production Sets Using Transconcave DEA, *European Journal of Operational Research* (to appear)

Post, G. T. and J. Spronk (2000): Evaluating Productive Performance under Uncertainty: Combining Data Envelopment Analysis, Mean-Variance Analysis, and Multi-Factor Risk Models. In: Shi, Y. and M. Zeleny (2000): "*New Frontiers of Decision Making for the Information Technology Era*", World Scientific Publishing Co., pp. 249-269.

Sandmo, A. (1971): On the Theory of the Competitive Firm under Price Uncertainty, *American Economic Review* 61, 65-73.

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

Starmer, C. (2000), Developments in Non-Expected Utility Theory: the Hunt for a Descriptive Theory of Choice under Risk, *Journal of Economic Literature* 38, 332-382.

Thompson, R.G., F.D. Singleton Jr., R.M. Thrall, and B.A. Smith (1986): Comparative Site Evaluations for Locating High Energy Lab in Texas, *TIMS Interfaces*, 1380-1395

Thompson, R.G., L.N. Langemeier, C.-T. Lee, E. Lee, and R.M. Thrall (1990): The Role of Multiplier Bounds in Efficiency Analysis with Application to Kansas Farming, Journal of Econometrics 46, 93-108

Tulkens, H. (1993): On FDH Efficiency Analysis: Some Methodological Issues and Application to Retail Banking, Courts, and Urban Transit, *The Journal of Productivity Analysis* 4 (1), 183-210.

Varian, H (1984): The nonparametric approach to production analysis, *Econometrica* 52 (3), 579-597.

Varian, H (1985): Nonparametric Analysis of Optimizing Behavior with Measurement Error, *Journal of Econometrics* 30 (1/2): 445-458.