MEASURING ECONOMIC EFFICIENCY WITH INCOMPLETE PRICE INFORMATION

With an Application to European Commercial Banks

Timo Kuosmanen
Helsinki School of Economics and Business Administration
Department of Economics and Management Science
P.O. Box 1210, FIN-00101 Helsinki, **FINLAND**

Thierry Post
Erasmus University Rotterdam
Faculty of Economics
3062 PA Rotterdam, **THE NETHERLANDS**

Rotterdam Institute for Business Studies (RIBES)
Report 9934

This revision: November 1999

ISBN 90-5086-346-9
ISSN 0921-0180

Under Revision for *European Journal for Operational Research*

Abstract
Measuring economic efficiency requires complete price information, while resorting to technical efficiency exclusively does not utilise any price information. In most studies, at least some information on the prices is available from theory or practical knowledge of the industry under evaluation. In this paper we derive upper and lower bounds for Farrell's overall economic efficiency**,** assuming incomplete price information. The bounds typically give a better approximation for economic efficiency than technical efficiency measures that use no price data whatsoever. Standard linear programming techniques can obtain nonparametric estimates for these bounds from empirical data. From operational point of view, the approach developed in this paper both extends the scope of the weight restricted Data Envelopment Analysis (DEA) models to more general technologies, and explicitly links these techniques to the classic theoretical framework of efficiency measurement. The approach is illustrated with an empirical application to large European Union commercial banks.

Key words: Economic efficiency measurement, imperfect price information, Data Envelopment Analysis, weight restricted models

1. Introduction

In the recent literature of productivity and efficiency analysis, much discussion has focused on economic efficiency measures, as well as on the economic justification of technical efficiency measures (see i.e. Ray, 1997; Färe and Grosskopf, 1997; Briec and Lemaire, 1999; Cooper et al. 1999; among others). Following Farrell's (1957) seminal paper, *economic efficiency* can be decomposed into two components*: allocative efficiency* and *technical efficiency*. The technical component requires quantitative volume data of inputs and outputs only, while prices are also necessary for measuring allocative efficiency.

Already Debreu (1951) and Farrell (1957) expressed their concern about the ability to measure prices accurately enough to make good use of economic efficiency measurement. For example, accounting data can give a poor approximation for economic prices (i.e. marginal opportunity costs), because of debatable valuation and depreciation schemes. Several authors, including Charnes and Cooper (1985) cite this concern as a motivation for emphasizing technical efficiency measurement. Consequently, many studies in the field of productivity and efficiency analysis, including the seminal articles on Data Envelopment Analysis (DEA) by Charnes, Cooper and Rhodes (1978), Banker, Charnes and Cooper (1984), and Charnes et al. (1985), assess efficiency solely in terms of technical efficiency.

Interestingly, the radial (i.e. Debreu-Farrell) technical efficiency measure provides a theoretical upper bound for economic efficiency, as Debreu (1951) already pointed out. However, this measure need not be a particularly good approximation for economic efficiency, as it does not utilize any price information whatsoever. In numerous empirical studies, at least some rough information on the economic prices is available from theory or practical knowledge of the industry under evaluation. One source of price information is prior knowledge on the quality or risk of the different inputs and outputs. For example, primary inputs are typically more expensive than secondary inputs. Therefore, the unit of labour input of key personnel (i.e. detectives, dentists, university professors, surgeons, teachers) is more expensive than that of assisting staff (i.e. secretaries, Ph.D. students, cleaners, janitors). As capital inputs are concerned, the unit cost of equity capital exceeds that of debt, because equity involves more risk for the capital suppliers than debt does. Consequently, there exist both need and opportunities to include incomplete price information in efficiency analysis, so as to improve approximation of economic efficiency concepts.

The above considerations are frequently cited as the motivation for imposing weight restrictions in DEA models, as initiated by Thompson et al. (1986).[1] In contrast to common belief, however, we think that incorporating price information in DEA is *not* desirable *per se*. Firstly, by augmenting the DEA reference set by price information we can lose valuable information on marginal substitution and transformation properties contained in the shadow prices of the production frontier. Secondly, the reference or target points given from the augmented reference set can be technically infeasible, and hence provide poor information for benchmarking purposes. Finally, and perhaps most importantly, the desirability of price information ultimately depends on the notion of economic efficiency considered, which is inseparable from the choice of the efficiency metric or orientation of measurement. For example, cost efficiency studied by Farrell does not depend on output prices at all. Hence, if production set is augmented in this case using both input and output prices as in Charnes et al. (1990) and Thompson et al. (1990), among others, we not only get infeasible reference units, but also poor estimates for Farrell's overall economic efficiency in terms of cost savings. Clearly, upgrading any DEA model with price data is not sufficient for measuring overall efficiency: A theoretical foundation is necessary for interpretation of results. Therefore, also from this perspective we see that it worth to extend the theory of efficiency measurement to allow for incomplete price data so as to gain insight to the use of the existing operational techniques (see also Thompson et al., 1990, p. 98).

In this paper, we extend the classic Farrell's theoretical framework to include incomplete price information. Assuming that a convex polyhedral cone can represent the price domain, we derive both upper and lower bounds for economic efficiency. The bounds typically give a better approximation for true economic efficiency than technical efficiency measures that use no price data whatsoever. For sake of generality, we only assume production technologies are closed. In particular, we do not impose convexity or monotonicity of the production set.

In addition to the theoretical discussion, we show how DEA estimators for these bounds can be computed from empirical data using standard linear programming techniques. In contrast to weight restricted DEA models, we use separate coefficients for shadow prices of the production set and for economic prices. This allows us to remedy some serious shortcomings of the weight restricted DEA models; most notably, the debatable *ad hoc* assumption of convex production set (see e.g. Kuosmanen, 1999, and Cherchye et al., forthcoming, for critical discussion). Moreover, we use the separate coefficients for distinguishing between production technology related information on one hand, and

[1] For a detailed survey of the weight-restricted DEA models, see Allen et al. (1997).

price or value related information on the other. Furthermore, the separate coefficients allow one to incorporate price information in both envelopment and multiplier problems.

The remainder of this paper is organised as follows. In Section 2 we present the Farrell decomposition along with some definitions of various efficiency measures. Section 3 discusses how incomplete price information can be used for deriving an interval for economic efficiency. Section 4 discusses standard mathematical programming techniques can obtain nonparametric estimators for these bounds from empirical data. Section 5 illustrates the approach with an empirical application for large European Union commercial banks, and show that even a single restriction on prices can substantially improve the economic efficiency estimation. Section 6 draws some conclusive remarks.

2. Farrell decomposition

The Farrell decomposition is a fundamental cornerstone of the theory of efficiency measurement. Farrell (1957) explicitly decomposed overall *economic efficiency* into components of *technical efficiency* and *allocative efficiency*.[2] Farrell's general framework applies to a number of different ways to define and measure economic efficiency. In efficiency analysis, the organisations under study differ both in terms of their objectives and their environment (compare i.e. business enterprises versus non-for-profit organisations). Therefore, alternative measures are needed for a meaningful assessment of economic efficiency for different organisations. Following Farrell, we focus on cost minimising behaviour of producers. In this perspective, the appropriate measure for economic efficiency is *cost efficiency*, associated with the *input oriented* technical efficiency measure (i.e. output is held constant). However, the proposed approach is directly applicable to analogous *revenue efficiency* and *profit efficiency* (Nerlove, 1965) measures, and the associated *output* and *graph* oriented technical efficiency measures respectively (see e.g. Banker and Maindiratta, 1988; Chambers et al., 1998; and Cherchye et al., 1999).

It is first necessary to introduce some notation. Throughout the paper, production inputs are represented by the input vector $x = \begin{pmatrix} x_1 & \cdots & x_q \end{pmatrix} \in \Re^q_+$, whereas outputs are denoted by $y = \begin{pmatrix} y_1 & \cdots & y_q \end{pmatrix} \in \Re^p_+$. Input prices in terms of marginal opportunity cost are represented by the price vector $w = \begin{pmatrix} w_1 & \cdots & w_q \end{pmatrix}^T \in \Re^q_+$. The empirical production data of the production units are represented by the output matrix $Y = (y_1 ... y_n)^T$, with $y_j = (y_{1j} \cdots y_{pj})$, and the input matrix $X = (x_1 ... x_n)^T$, with $x_j = (x_{1j} \cdots x_{qj})$. In addition, we use the index set $S = \{1,...,n\}$.

There are many alternative ways to characterise the production technology (see e.g. Färe (1988) for an elaborate discussion). The most general representation is the production possibilities set *T*:

$$\textbf{(2.1)} \quad T := \left\{ (x, y) \in \Re^{q+p}_+ \middle| x \text{ can produce } y \right\}$$

[2] In fact, Farrell originally used the term “price efficiency” instead of allocative efficiency, but we agree with Kopp (1981) who pointed out that the assumed price-taking nature of production units that is better captured by the term allocative efficiency.

For the purposes of Sections 2 and 3, however, a more convenient representation of the technology is the *input correspondence*

(2.2) $L(y) := \{x | (x, y) \in T\}$,

which maps outputs *y* into subsets *L*(*y*) of inputs. In other words, the input set *L*(*y*) denotes all input vectors *x* that yield output *y*. For sake of generality, in Sections 2 and 3 we only postulate *T* and *L* to be closed. In the empirically oriented section 4 and 5 we will further assume both *T* and *L* are monotonous, i.e. inputs and outputs are assumed freely disposable, so as to enable empirical approximation. However, we do not assume convexity of either *T* or *L*. To the best of our knowledge, there are no valid theoretical or empirical arguments for such convexity assumptions (see e.g. Kuosmanen (1999) and Cherchye et al. (forthcoming) for critical discussion). In addition, as demonstrated below, convexity assumptions are not required for duality results, nor for deriving tractable model formulations.

Cost efficiency is conventionally defined as the ratio of the minimum cost of producing the output of the evaluated production unit to the actual cost at the given input prices and technology.[3] Using the standard notion of *cost function*, i.e.

(2.2) $C^L(y; w) := \underset{x}{Min}\{xw | x \in L(y)\}$

cost efficiency can be formally defined as

(2.3) $CE^L(x, y; w) := \dfrac{C^L(y; w)}{xw}$.

In the subsequent section, we shall utilise an equivalent formulation of cost efficiency that views cost efficiency as the cost function at normalised prices, i.e.

(2.4) $CE^L(x, y; w) = \underset{x' \in L(y)}{Min}\{x'w' | w' = aw, a \in \Re_+; xw' = 1\}$

In this measure, prices are normalised such that the normalised cost of the evaluated production vector equals unity, i.e. $xw' = 1$. It is worth to note at least the following properties of this measure: Firstly, cost efficiency is homogeneous of degree one in input prices. Furthermore, the measure is restricted by construction to the closed interval [0,1], the value of unity representing total efficiency.

Number of alternative measures for *technical efficiency* have been proposed in the literature (see De Borger et al., 1998, for an up-to-date presentation). Nevertheless, it turns out that the only measure with a tractable dual formulation is the radial measure proposed by Debreu (1951) (see also Shephard (1953)), and adopted by Farrell. The *Debreu-Farrell input efficiency measure* defines inefficiency as the maximum equiproportionate reduction of inputs that is attainable without reducing any of the outputs, formally

[3] It is worth to note that we evaluate efficiency of technically feasible production vectors $(x,y) \in T$ only, so we use *minimum* and *maximum* rather than *infimum* and *supremum*.

(2.5) $DF^L(x,y) := \underset{\theta}{Min}\{\theta \mid \theta x \in L(y)\}$.

Analogous output and graph measures are also available (see i.e. Farrell, 1957; Chambers et al., 1998; and Cherchye et al.,1999)

Allocative efficiency can be seen as the cost efficiency measure (or the overall economic efficiency in general) applied to the technically efficient reference production plan, i.e. $(DF^L(x,y)x, y)$. In case of cost savings, allocative efficiency is formally defined as

(2.6) $AE^L(x,y;w) := \dfrac{C^L(y;w)}{DF^L(x,y)xw}$.

Although the above definition associates allocative efficiency intimately with cost efficiency, it is worth to note that allocative efficiency can be equivalently defined without any such reference by using the radial Debreu-Farrell input gauge and the minimum isocost hyperplane $H^L(y;w) = \{x' \mid x'w = C^L(y;w)\}$, i.e.

(2.7) $AE^L(x,y;w) = DF^H(DF^L(x,y)x, y)$.

It is immediate from (2.6) that the allocative efficiency measure has by definition the homogeneity property in input prices. By construction, allocative efficiency is restricted to the interval [0,1]. Moreover, it is obvious from (2.6) that the product of allocative efficiency and technical efficiency always equals cost efficiency, i.e.

(2.8) $CE^L(x,y;w) = AE^L(x,y;w) \times DF^L(x,y) \quad \forall x \in L(y), w \in \Re^q_+$.

Equation (2.8) is the fundamental Farrell decomposition. Among other things, this decomposition implies that technical efficiency is a necessary, but not sufficient condition for economic efficiency. This result is also known as the *Mahler inequality* (see e.g. Färe and Primont, 1995), i.e.

(2.9) $CE^L(x,y;w) \le DF^L(x,y) \quad \forall x \in L(y), w \in \Re^q_+$.

Figure 1 illustrates the Farrell decomposition graphically with two inputs, at a fixed level of output *y*. The production unit *k* lies in the interior of the (non-convex) input set *L* so it is inefficient both in technical and economic sense. The Debreu-Farrell technical efficiency measure is the ratio OR_{DF}/Ok. Allocative efficiency is the ratio OR'_{CE}/OR_{DF}. Finally, cost efficiency is OR'_{CE}/Ok, which we obtain by multiplying the technical efficiency and allocative efficiency indices.

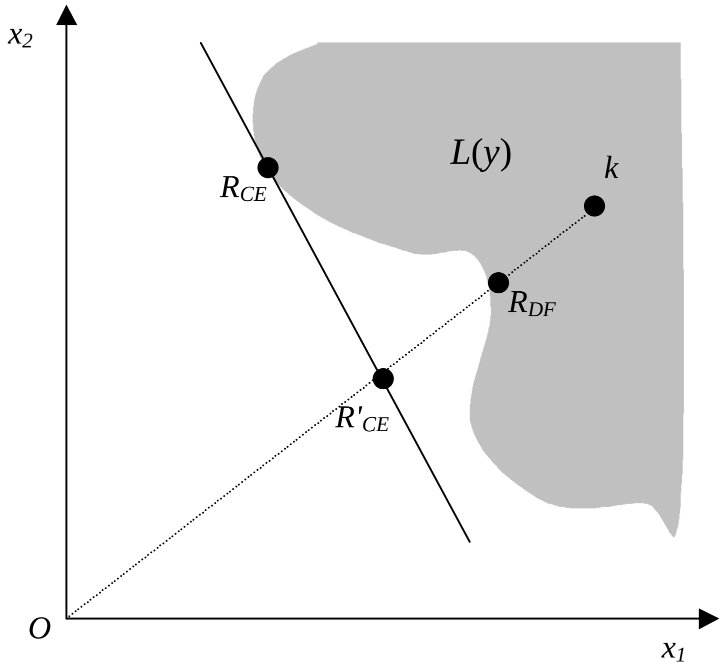

Figure 1: Farrell decomposition of cost efficiency (OR'_{CE}/Ok) as a product of technical efficiency (OR_{DF}/Ok) and allocative efficiency (OR'_{CE}/OR_{DF}) components.

3. Farrell framework under incomplete price information

The economic and technical efficiency measures are extreme cases with respect to the requirement and the utilisation of price information. The economic efficiency measures require complete and accurate information. On the other hand, no prices appear in the definitions of technical efficiency measures, cf. e.g. (2.6). However, we believe many research situations are best described by the intermediate case where some incomplete price information is available. In that case, the economic approach can not be used, because the price information is incomplete. However, the technical approach leaves potentially valuable price information unutilised. Below, we demonstrate how incomplete price information can be utilised to obtain better approximations for economic efficiency than available from technical efficiency measures.

In this paper we relax the information requirement of prices by only assuming that the factor prices belong to domain $W \subseteq \Re_+^q$. For simplicity, we use a polyhedral convex cone to represent the price domain. More specifically, we assume the following structure for the price domain:

(3.1) $W = \left\{ w \in \Re_+^q \middle| Aw \geq 0 \right\}.$

This cone represents the price domain in terms of l linear inequalities. A is a $l \times q$ matrix. It should be noted that 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. For example, the information that equity capital is more expensive than debt capital can be represented by a single linear inequality. In addition, the limiting cases of exact and complete price information on 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, polyhedral convex cones are computationally convenient, because they can be included in Linear Programming models.

Figure 2 illustrates a graphical example of a price domain as represented by the convex cone W. This price domain restricts the relative price $W1/W2$ of the two inputs to the closed interval [0.5, 2].

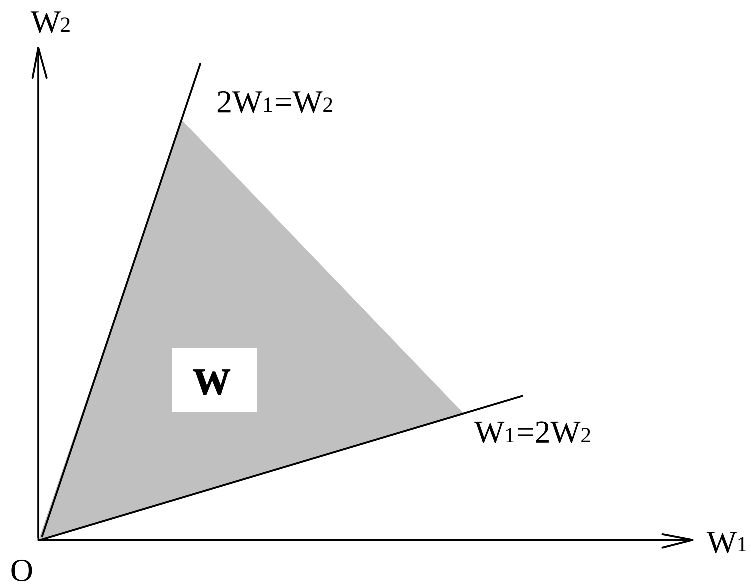

Figure 2: Example of the price domain as represented by the convex cone W

To obtain an upper bound for cost efficiency, we propose to use the *maximum* value of cost efficiency over the price domain, i.e.:

(3.2) $$\overline{CE}^L(x, y; W) = \max_{w \in W} CE^L(x, y; w) = \max_{w \in W} \min_{x' \in L(y)} \{x'w \mid xw = 1\}.$$

Using duality theory, we can relate this upper bound to the Debreu-Farrell measure (2.6).

THEOREM 3.1: For all closed input sets $L(y)$ and non-negative price vectors w, the upper bound of cost efficiency $\overline{CE}^L(x, y; W)$ is equal to the Debreu-Farrell input measure $DF^U(x, y)$ relative to the augmented input set $U(y) = \{x \in \Re^q_+ \mid x \geq ax' + (1-a)x'' + zA; z \in \Re^l_+; x', x'' \in L(y); a \in [0,1]\}$.

PROOF 3.1 In problem (3.2), the inputs set L can be harmlessly substituted by the convex hull of L, i.e. $CHL(y)=\{x\in\Re_+^q \mid x\geq ax'+(1-a)x''; x', x''\in L(y); a\in[0,1]\}$, because the objective function is linear in inputs. Hence, problem (3.2) can equivalently be defined as:

(i) $\overline{CE}^L(x,y;W)=\max_{w\in W}\min_{x'\in CHL(y)}\{x'w \mid xw=1\}=\min_{x'\in CHL(y)}\max_{w\in W}\{x'w \mid xw=1\}.$

This problem embeds the following Linear Programming problem:

(ii) $\max_{w\in W}\{x'w \mid xw=1\}=\max_{w\in\Re_+^q}\{x'w \mid xw=1; Aw\geq 0\}.$

The dual associated with this problem is:

(iii) $\min_{q,z\in\Re_+^l}\{q \mid x'+zA\leq qx\}.$

Substituting (iii) for (ii) in (i) gives:

(iv) $\overline{CE}^L(x,y;W)=\min_{x'\in CHL(y)}\min_{q,z\in\Re_+^l}\{q \mid x'+zA\leq qx\}=DF^U(x,y).$

The above theorem shows that the weight restricted DEA approach can in fact be given a sound economic interpretation, provided that the empirical reference set and the efficiency measure are correctly specified, and the price domain contains the true input prices. In addition, the above theorem extends the use of augmented reference sets to cases where the true production set is non-convex. The augmented reference set $U(y)$ represents all input vectors x that involve a minimum cost of xw at most, for *all* prices $w\in W$. As $L(y)$ is always contained within $U(y)$, the upper bound $\overline{CE}^L(x,y;W)$ can never exceed the value of $DF^L(x,y)$, but depending on the price domain W, it can be substantially lower. Thus, this upper bound can give a substantially better approximation for economic efficiency than the Debreu-Farrell measure alone.

Note that the elements of $U(y)$ need not be included in $L(y)$, and hence they need not suffice to produce output y. For example, $U(y)$ has convex isoquants even if the isoquants of the true production set are non-convex. However, the non-feasible elements can be harmlessly used for cost comparisons, because $L(y)$ necessarily contains at least one input vector that can produce y at a cost of xw at most. As $DF^U(x,y)$ is computationally convenient, we will elaborate on how to obtain empirical approximations for $U(y)$ in Section 4 below.

Figure 3 illustrates graphically the upper bound of cost efficiency relative to $L(y)$, which equals the Debreu-Farrell measure relative to set $U(y)$. The slopes of the two isocost lines drawn in Figure 3 represent the extreme scenarios of relative prices included in a hypothetical price domain W. Notice that the radial reference point R'_{CE} is an infeasible production plan, as it lies outside the input correspondence $L(y)$. However, $L(y)$ contains two production plans R_{CE1} and R_{CE2} that produces the output y at the same cost as the reference point R'_{CE}. Note that while $L(y)$ is non-monotonous and non-convex, $U(y)$ is both monotonous and convex. Therefore, for the purpose of evaluating cost efficiency, the input set can be 'monotonised' and 'convexified' without harm. However, this does neither imply that the entire production set T could be convexified, nor that the convexification would be allowed for measuring technical or allocative efficiency.

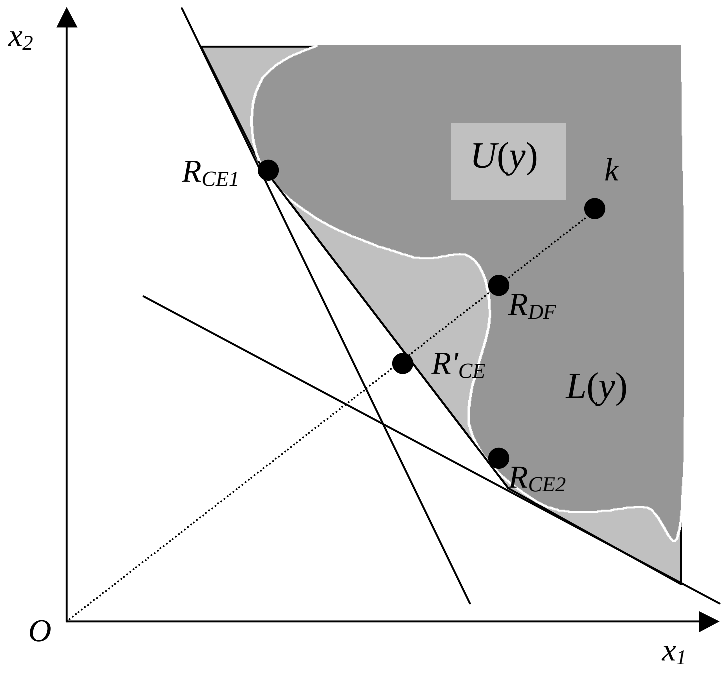

Figure 3: Illustration of the upper bound of cost efficiency, which is the Debreu-Farrell measure relative to set the augmented set U rather than L(y).

Interestingly, if price information is complete, i.e. the price domain *W* only contains multiples of the true price vector, the upper bound (3.2) always equals the cost efficiency measure (2.3). On the other hand, if no price information is available, sets *U*(*y*) and *L*(*y*) coincide, and (3.2) equals the Debreu-Farrell measure (2.5) in the special case of monotonous and convex input set. This, however, need not be the case in general. Therefore, the upper bound (3.2) can give a better approximation for cost efficiency in case of non-convex technologies even when no price information is available, i.e. price domain *W* contains the whole non-negative orthant of real valued vector space. Finally, the properties of measure (3.3) effectively induce a gradual transition from Debreu-Farrell efficiency (associated with no price information) to economic efficiency (associated with full complete price information).

A lower bound for cost efficiency can be obtained in a similar way. More specifically:

(3.4) $$\underline{CE}^{L}(x,y;W) = \min_{w \in W} \min_{x' \in L(y)} \left\{ x'w \middle| xw = 1 \right\}.$$

Using duality theory, we can also relate this lower bound to the Debreu-Farrell measure (2.6).

THEOREM 3.2: For all closed input sets *L*(*y*) and non-negative price vectors *w*, the lower bound of cost efficiency $\underline{CE}^{L}(x,y;W)$ is equal to Debreu-Farrell input measure $DF^{V}(x,y)$ defined with respect to the augmented input set $V(y) = \left\{ x \in \Re_{+}^{q} \middle| x \geq ax' + (1-a)x'' - zA;\, z \in \Re_{+}^{l};\, x', x'' \in L(y);\, a \in [0,1] \right\}$.

PROOF 3.2 Since the objective function is linear in the inputs, the inputs sets of *T* can be 'convexified' without harm for solving problem (3.4). Hence, problem (3.4) can equivalently be defined as:

(i) $\underline{CE}^{L}(x,y;W)=\min_{w\in W}\min_{x'\in CHL(y)}\{x'w \mid xw=1\}=\min_{x'\in CHL(y)}\min_{w\in W}\{x'w \mid xw=1\}.$

This problem embeds the following Linear Programming problem:

(ii) $\min_{w\in W}\{x'w \mid xw=1\}=\min_{w\in\Re_{+}^{q}}\{x'w \mid xw=1;\ Aw\geq 0\}.$

The dual associated with this problem is:

(iii) $\max_{q,z\in\Re_{+}^{l}}\{q \mid x-zA\geq qx_{k}\}.$

Substituting (iii) for (ii) in (i) gives:

(iv) $\underline{CE}^{L}(x,y;W)=\min_{x'\in CHL(y)}\min_{q}\{q \mid x'-zA\leq qx;\ z\in\Re_{+}^{l}\}=DF^{V}(x,y)$.

Combined with the upper bound, the lower bound can be used to construct an interval approximation for the true cost efficiency. Theorem 3.2 implies that the weight restriction approach can in fact be used for computing the whole interval for economic efficiency, whereas the outcomes of this type of models have been predominantly interpreted as "optimistic" point estimates. We think that also the worst case scenario, as reflected by the augmented reference set *V*(*y*) gives valuable information in many applications. The set *V*(*y*) represents all input vectors that produce output *y* at a minimum cost of *xw* at most, for *some* prices $w\in W$. Obviously, the reference set *U*(*y*) is always contained within *V*(*y*). Again, the elements of *V*(*y*) need not be contained within *L*(*y*). However, *L*(*y*) contains input vectors that can produce *y* at a cost of *xw* for some prices. As $DF^{V}(x,y)$ is computationally convenient, we will elaborate on how to obtain empirical approximations for *V*(*y*) in Section 4 below.

Figure 4 illustrates graphically the lower bound relative to set *V*(*y*). Notice that the radial reference point in the boundary of *V*(*y*), denoted by R'_{CEL}, is associated with the nonradial reference point in the boundary of *L*(*y*), denoted by R_{CEL}. In fact, R_{CEL} represents the optimal *x* to the minimization problem (3.3). Interestingly, the upper and lower bounds of cost efficiency provide some economic justification for the general idea of non-radial projections to the technical frontier. Unfortunately, any well-known non-radial measure does not generally coincide with these reference points. As shown by Figure 4, achieving this reference may actually require increase in some inputs.

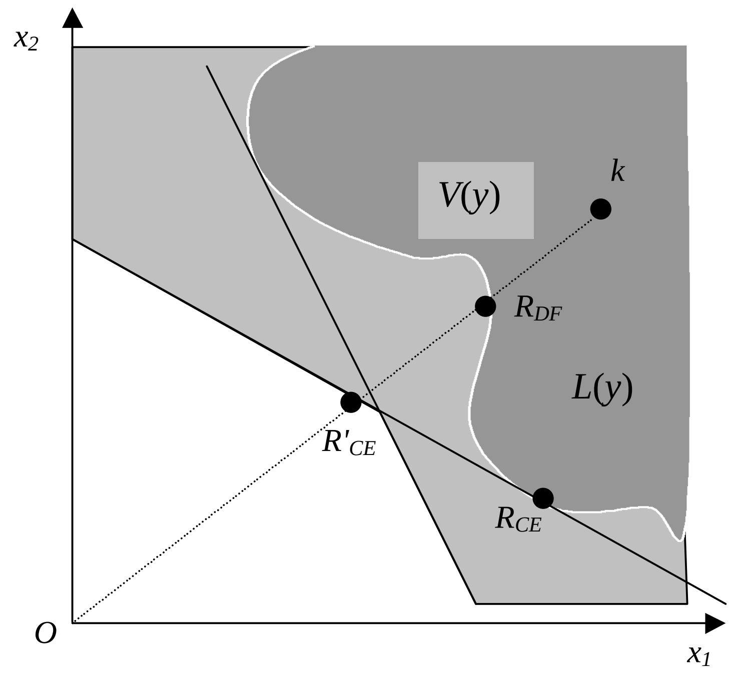

Figure 4: Illustration of the lower bound of cost efficiency, which is the Debreu-Farrell measure relative to set V

Consider the special case of full information. As could be expected the sets $U(y)$ and $V(y)$ coincide, and the lower bound equals the upper bound. On the other hand, in case of no prior price information, the lower bound can fall considerably short of the upper bound. Still, the lower bound typically exceeds zero, which suggests that the theoretical lower bound can give new insight even when prices are totally unavailable.

Finally, using the Farrell decomposition (2.4) and the cost efficiency bounds derived above, we can construct a theoretical interval for allocative efficiency, i.e.

(3.5) $$AE^L \in \left[\underline{AE}^L, \overline{AE}^L\right],$$

where the lower bound is given by

(3.6) $$\underline{AE}^L = \underline{CE}^L / DF^L,$$

and the upper bound is respectively

(3.7) $$\overline{AE}^L = \overline{CE}^L / DF^L.$$

It is immediate from (3.6) and (3.7) that the interval for allocative efficiency is simply cost efficiency interval normalised by technical efficiency. Nevertheless, it can be highly informative to decompose the estimated cost efficiency measures into its two components so as to gain insight of the underlying sources of inefficiency, i.e. to what extent the observed overall inefficiency is due to inefficient planning versus inefficient operation. In this respect, even a rough interval for allocative efficiency can provide valuable information in comparison to the level of technical efficiency. Moreover, the interval generally diminishes by a more accurate price information.

4. Empirical estimators

Although the above measure can deal with incomplete price information, complete information on the input set $L(y)$ is still required. Unfortunately, one typically faces imperfect information on the production possibilities in the empirical analysis of production data. However, empirical inference is possible by approximating the true input correspondence by particular empirical input sets constructed from empirical data.

Various approximations can be employed, depending on the assumptions imposed on production possibilities and the data used for approximating it. In the nonparametric literature, the most popular approximation of production possibilities set T is the *convex monotone hull* of observations used i.e. by Banker *et al.* (1984) among others for estimating Debreu-Farrell input efficiency, and Banker and Maindiratta (1988) among others for estimating overall economic efficiency. Unfortunately, this set imposes strong *ad hoc* assumptions of convex input correspondence, convex output correspondence, and convex graph set that in general cannot be justified by economic theory (see Kuosmanen (1999) and Cherchye et al. (forthcoming) for critical discussion).

Recall that in the previous sections of this study we only assumed sets $L(y)$ and T to be closed. However, we need to impose some additional structure for the production possibilities in order to obtain a non-empty empirical reference set. Therefore, we henceforth resort to the *free disposable hull* (henceforth FDH) technology (see i.e. Deprins *et al.* (1984), and Tulkens (1993)). FDH approximation of the input correspondence L can be formally represented by:

(4.1) $$\hat{L}_{FDH}(y) = \left\{ x \in \Re^q_+ \left| \begin{pmatrix} y \\ -x \end{pmatrix} \le \begin{pmatrix} l^T Y \\ -l^T X \end{pmatrix}; l^T e = 1; l_j \in \{0,1\}; j \in S \right. \right\}.$$

If the production set T is monotonous, i.e. inputs and outputs are freely disposable, and all the observations included in the data set are technically feasible (as they should be if data is reliable; after all, these production plans are observed), then $\hat{L}_{FDH}(y) \subseteq L(y)$. Hence, measuring technical efficiency relative to the FDH approximation gives an upper bound for the true Debreu-Farrell efficiency, i.e.

(4.2) $$\overline{DF}^L(x,y) = DF^{\hat{L}_{FDH}}(x,y) = \min_{q,l} \left\{ q \left| l^T X \le qx; l^T Y \ge y; l^T e = 1; l_j \in \{0,1\} \; \forall j \in S \right. \right\}.$$

Solving a Mixed Integer-Linear Programming problem for each DMU can obtain these estimates. Tulkens (1993) has also provided a simple and effective enumeration algorithm for this problem.

Unfortunately, a lower bound for technical efficiency cannot be obtained when *L* is unknown, except for the value of zero, which provides a theoretical infimum when no free lunch is allowed, i.e. production of positive output bundles is impossible with non-positive input levels.

FDH approximation also forms a good starting point for cost efficiency evaluation. For this purpose, we construct an FDH approximation for the price-augmented reference set *U(y)* as

(4.3) $$\hat{U}_{FDH}(y) = \left\{ x \in \Re^q \middle| x \geq l^T c(y) + A^T z;\ l^T e = 1;\ l_j \geq 0,\ j \in S;\ z \in \Re^l_+ \right\} ,$$

where $c(y)$ denotes the input matrix of such observed production plans that produce higher output than *y*, i.e.

(4.4) $$c(y) = \left\{ x_j \in X \middle| y_j \geq y \right\} .$$

Using Theorem 3.1 we obtain the following upper bound for cost efficiency

(4.5) $$\overline{CE}^{\hat{L}_{FDH}}(x,y) = DF^{\hat{U}_{FDH}}(x,y) = \min_{q,l,z} \left\{ q \middle| l^T c(y) + A^T z \leq qx;\ l^T e = 1;\ l_j \geq 0\ \ j \in S;\ z \in \Re^l_+ \right\} .$$

Similarly, for computing the empirical lower bound for cost efficiency we first construct the empirical reference set $\hat{V}_{FDH}(y)$:

(4.6) $$\hat{V}_{FDH}(y) = \left\{ x \in \Re^l \middle| x \geq l^T c(y) - A^T z;\ l^T e = 1;\ l_j \geq 0\ \ j \in S;\ z \in \Re^l_+ \right\} .$$

Using Theorem 3.2, the empirical lower bound for cost efficiency can be computed as

(4.7) $$\underline{CE}^{\hat{L}_{FDH}}(x,y) = DF^{\hat{V}_{FDH}}(x,y) = \min_{q,l,z} \left\{ q \middle| l^T c(y) - A^T z \leq qx;\ l^T e = 1;\ l_j \geq 0\ \ j \in S;\ z \in \Re^l_+ \right\} .$$

Note that both problem (4.5) and (4.7) can be solved by standard Linear Programming algorithms, despite the fact that we built on non-convex FDH technology. As pointed out in the previous section, the convexification of input correspondence is harmless due to the linearity of the cost function. Therefore, the price-augmented FDH input set $\hat{U}_{FDH}(y)$ can be equivalently obtained as the input correspondence of such empirical production possibility sets that only impose convexity in input space (see i.e. Bogetoft (1996), and Post (1999). In the above formulas, potential non-convexity in graph or output space is taken into account by first filtering out such observations that produce smaller amount of any output than the evaluated production plan.

The empirical bounds of cost efficiency derived above have following properties. If FDH input set is contained within the true input correspondence, then the empirical upper bound is always greater than or equal to the theoretical upper bound, i.e.

$\overline{CE}^{\hat{L}^{FDH}} \geq \overline{CE}^{L}$. Moreover, as the theoretical upper bound is greater than or equal to the true cost efficiency, i.e. $\overline{CE}^{L} \geq CE^{L}$, it follows that the true cost efficiency can never exceed the empirical upper bound, i.e. $\overline{CE}^{\hat{L}^{FDH}} \geq CE^{L}$. In addition, under some generally accepted assumptions of the distribution of inefficiency terms, the FDH converges to the true production set as the sample size is increased (see Park, Simar and Weiner, 1997). Therefore, the empirical upper bound asymptotically converges to the theoretical upper bound, i.e. $\overline{CE}^{\hat{L}^{FDH}} \rightarrow \overline{CE}^{L}$.

Similarly, the empirical lower bound is always greater than or equal to the theoretical lower bound, i.e. $\underline{CE}^{\hat{L}^{FDH}} \geq \underline{CE}^{L}$. However, as $\underline{CE}^{L} \leq CE^{L}$, the empirical lower bound can generally be greater or less than the true cost efficiency. Thus, the lower bound is essentially a sample property. We would underline that the true economic efficiency can in some cases fall short of the empirical lower bound i.e. due to the finite sample error. However, the empirical lower bound asymptotically converges to the corresponding theoretical value, i.e. $\underline{CE}^{\hat{L}^{FDH}} \rightarrow \underline{CE}^{L}$. Furthermore, the lower bound estimate relative to a sample can still be a valuable information for practical purposes, as there is no evidence of DMUs operating with lower cost than the reference cost on the lower bound.

Unfortunately, a more precise theoretical interval than the natural [0,1] range cannot be obtained for allocative efficiency, because we can only characterize upper bounds for both cost and technical efficiency. Still, it can be interesting to consider allocative efficiency interval observed in the *sample* relative to the empirical technology, i.e.

(4.4) $$AE^{\hat{L}_{FDH}} \in \left[\underline{AE}^{\hat{L}_{FDH}}, \overline{AE}^{\hat{L}_{FDH}} \right].$$

The upper and lower bound of allocative efficiency can be computed from (3.6) and (3.7) by substituting *L* by its empirical approximation. In a large enough data set the sample interval can give a good approximation due to the asymptotic properties of FDH approximation.

5. Empirical illustration

The nonparametric approach to efficiency measurement has seen extensive application for studying the financial industry. For example, Berger and Humphrey (1997) find that 69 out of 122 frontier efficiency studies for financial institutions use the nonparametric approach. To illustrate the use of incomplete price information in economic efficiency analysis, we performed an empirical application for commercial banks. Specifically, we used a data set with 1997 financial statement data of the 453 largest commercial banks in the European Union[4]. For convenience, we use a simplified representation of the bank technology, which involves a single output, *total earning assets*, and a three inputs, *equity capital*, *debt capital* and *operational costs*

[4] In this article, we use BankScope data provided by Bureau van Dijk Nederland.

(which aggregates all inputs apart from equity and debt). All variables are measured in millions of Euro. Figure 1 illustrates graphically the employed technology. Furthermore, Table 1 presents some descriptive statistics for the data set.

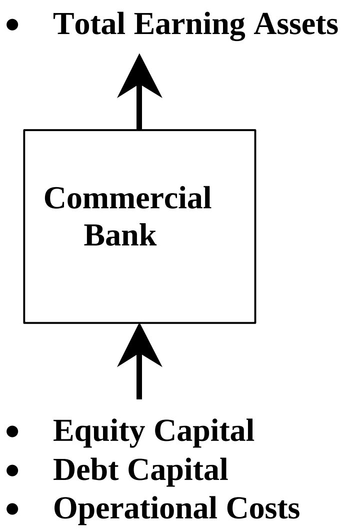

Figure 5: Graphical representation of the bank technology.

Table 1: Descriptive statistics for the bank sample.

	Tot. Earn. Ass.	Eq. Cap.	Debt cap.	Op. Costs
Mean	1027.88	468.49	23389.64	24283.46
Median	222.25	64.34	4005.60	3964.36
Maximum	30808.54	10433.01	568501.80	574533.80
Minimum	8.35	1.23	500.70	533.43
St. Dev.	2547.49	1286.28	60211.62	62961.11
Skewness	5.82	4.83	4.84	4.77
Kurtosis	51.52	29.90	31.24	29.91

To measure cost efficiency and allocative efficiency, information is required on the prices of the inputs, say $w = (w_1 \quad w_2 \quad w_3)$, where w_1 denotes the price of equity capital, w_2 the price of debt capital and w_3 the price of other inputs. Our data set includes operational cost, which is computed using the price of other inputs, w_3. However, the prices of capital inputs are not available. The price of debt capital is not available, because interest revenues and interest expenditures are aggregated as *net interest income*. In addition, the cost of capital for equity relates to the *ex ante* required rate of return from future dividend payments and capital gains. This rate of return can't be inferred from *ex post* financial statements. In addition, the required rate of return on equity generally depends on the debt/equity ratio, and therefore can not be assumed to apply to reference units that have a different debt/equity ratio than the evaluated DMU. These complications exclude the possibility of direct measurement of cost efficiency and allocative efficiency.

As discussed above, the input-oriented Debreu-Farrell measure for technical efficiency can be used as an upper bound for cost efficiency. However, prior price information can derive better bounds for cost efficiency. For example, economic theory assumes that the cost of equity capital exceeds that of debt capital. The rationale behind this assumption is that equity involves more risk for the capital supplier than debt does, because debt claims are senior to equity claims. Combined with the assumption that prices are non-negative, this assumption gives the following *A* matrix:

$$\textbf{(5.1)} \quad A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix}.$$

Additional information on the prices could be found, e.g. using empirical estimates from capital market data or subjective assessments by bank managers. However, for the purpose of illustration, we resort to this single assumption on the relative price of equity and debt capital. Our purpose is to demonstrate by this example that valuable information can be gained by this simple restriction on input prices.

For these purposes we computed the Debreu-Farrell measure (4.2), and the empirical upper and lower bounds for cost efficiency associated with *A*, i.e. (4.4) and (4.7). We subsequently computed the empirical bounds on allocative efficiency, i.e. (3.6) and (3.7). Table 2 reports the summary statistics of the cost efficiency bounds and Debreu-Farrell technical efficiency.[5] The statistics of the allocative efficiency bounds are reported in Table 3.

Table 2: Statistics on the cost efficiency bounds and Debreu-Farrell measures

	Cost efficiency Lower bound	**Cost efficiency Upper bound**	**Debreu-Farrell**
Mean	0.114	0.978	0.988
Median	0.048	1	1
Maximum	1	1	1
Minimum	0.007	0.312	0.369
St. Dev.	0.176	0.048	0.041
Skewness	3.374	-6.962	-9.141
Kurtosis	12.574	83.382	123.774

[5] More detailed estimation results can be obtained from the authors upon request.

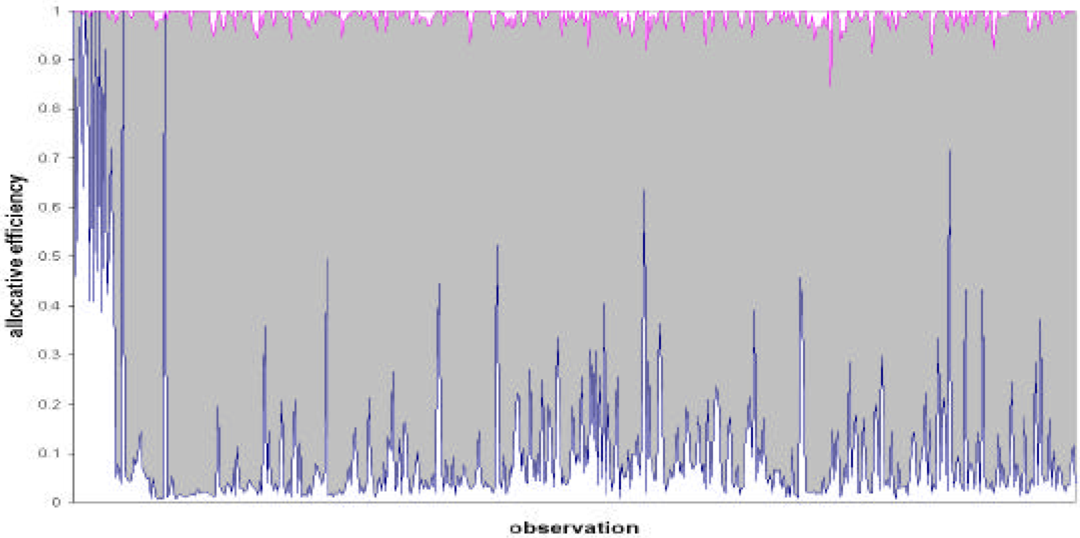

Figure 6: White area represents the improvement in cost efficiency bounds associated with the information that the cost of capital of equity exceeds that of debt.

Figure 6 graphically illustrates the relative improvement in the cost efficiency bounds achieved by incorporating price restriction (5.1), in other words, the computed cost efficiency interval is presented at the normalised scale [0, $DF^{\hat{L}_{FDH}}(x,y)$]=[0,1]. Alternatively, The grey area between the upper line and the lower line in Figure 6 can be viewed as the allocative efficiency interval $\left[\underline{AE}^{\hat{L}^{FDH}}, \overline{AE}^{\hat{L}^{FDH}}\right]$ for the observations. As allocative and cost efficiency intervals generally improve at the same rate, in what follows we refer to intervals and bounds in both meanings. Table 3 further summarises the improvements from using prior price information relative to interval from [0, $DF^{\hat{L}_{FDH}}(x,y)$], i.e. the interval associated with no prior information.

Table 3: Statistics on the improvements due to price information

	Lower Bound	Upper Bound
Mean	0.114	0.989
Median	0.049	1.000
Maximum	1.000	1.000
Minimium	0.007	0.846
St. Dev.	0.176	0.018
Skewness	3.371	-2.210
Kurtosis	12.554	11.122

It is clear both from Figure 6 and Table 3 that for a number of banks, incorporating price information substantially narrow intervals from [0,1], i.e the interval associated with no prior information. We see that the reduction in the intervals achieved by the restriction (5.1) solely, as represented by the white area, is quite substantial relative to the additional information content in (5.1). Obviously, for some banks, the improvements are more substantial than for others. For example, for a number of banks the lower bound equals unity. These banks are demonstrably cost efficient in this sample at all prices within the price cone. Note that in the classic framework, when full price data was unavailable, the cost efficiency interval for a Debreu-Farrell efficient bank was the maximal [0,1]. For some of these banks, we were able to provide evidence of their efficiency that was previously unavailable.

In addition, for many banks the upper bound is substantially lower than the Debreu-Farrell measure. For these banks, (1) the radial projection point is on a non-convex part of the isoquant, or alternatively (2) the marginal rates of technical substitution at the projection point differ from the input prices contained within the price cone.

For the majority of banks, however, the improvement in the interval was due to the introduction of the lower bounds. There certainly is a correlation between upper and lower bounds, so that when upper bounds are relatively close to unity, like with most banks in this data set, the lower bounds capture a greater proportion of the narrowing of the bounds, and vice versa. However, notice also that the above results may be subject to small sample bias. The empirical measures are biased above the theoretical bounds if the empirical production set is only a subset of the true production set, as is true in small samples. As discussed above, this positive bias preserves the status of the empirical upper bound as an upper bound for true cost efficiency. However, the empirical lower bound is not necessarily a lower for true cost efficiency, and should be interpreted as a sample property rather than a population property. This needs to be taken into consideration in the interpretation of the empirical bounds. Still, the empirical and theoretical bounds converge as the sample size increases.

We conclude by pointing some interesting opportunities to extend this type of empirical analysis. Naturally, collecting more accurate price restrictions can help to further improve the efficiency intervals. Moreover, when more accurate price data is not available, it can be worth to consider alternative price scenarios, speculate what would happen if the price domain would be specified in a different way, or how particular production assumptions (i.e. convexity, constant returns to scale, etc.) would affect the results. However, these speculations are beyond the scope of the current application. Our purpose was to show that the theoretical framework proposed in this paper provides an operationally tractable basis that requires little additional information, and most remarkably, rests of minimal assumptions on the unknown production process.

6. Concluding remarks

Economic efficiency measures require complete price data, while technical efficiency measures do not utilise any price information whatsoever. In many empirical studies, price information is considered highly relevant, but price data available is considered incomplete for using economic efficiency concepts. We derived upper and lower bounds for overall economic efficiency and allocative efficiency assuming incomplete price data in the form of a convex polyhedral cone. Standard linear programming techniques can obtain nonparametric estimates for these bounds from empirical data.

The application for European commercial banks presented in Section 5 shows that these bounds can give a substantially better approximation for economic efficiency than technical efficiency measures; a single theoretically sound assumption about the relative prices of equity and debt capital generated a substantial improvement.

In the Introduction, we noted that our approach has a strong similarity with so-called weight restricted DEA models. Now, how do the empirical bounds we derived from Farrell's theoretical model compare to the weight restricted DEA models? In contrast to the weight restricted models, we distinguish between the marginal properties (shadow prices) of the production set on one hand, and the economic prices on the other. This distinction has at least three important implications. First, it allows one to incorporate information in both envelopment and multiplier problems. Hence, including economic prices does not interfere with the elicitation of information on the shadow prices. Second, the distinction reveals that different notions of economic efficiency require different kinds of price information. For example, it is immediately clear that cost efficiency measures require information on the input prices, but not on the output prices. Third, the distinction facilitates the analysis of non-convex production technologies, which we consider desirable e.g. to account for increasing marginal productivity. In addition to distinguishing between technical and economic information, we compute both upper and lower bounds so as to present the whole interval of possible efficiency values, whereas weight restricted DEA models usually consider an upper bound at 'most favourable' weights only.

Regardless of the limitations of weight restricted DEA models for measuring economic efficiency, proper restrictions of technology parameters can help to improve the approximation of the production set. However, it should be understood that weight restrictions reflect technical information on marginal products or marginal substitution or transformation rates, rather than information on economic prices. In fact, weight restrictions have been utilized to reflect the former information in DEA right from the beginning, for example to impose particular assumptions of returns-to-scale (see e.g. Banker et al, 1984). Therefore, it is useful to distinguish between technology related information on one hand, and price or value related information on the other, both of which can be included in the model by parameter restrictions. We conclude by calling for more attention to the interpretation of weight restricted DEA models. We hope that the constructive critique presented in this paper increases the awareness of the interdependence between efficiency measures, production information, and price information.

References

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, No. 9, 1078 – 1092.

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

Berger, A. N. and D. B. Humphrey (1997): 'Efficiency of Financial Institutions: International Survey and Directions for Future Research', *European Journal of Operational Research* 98: 175-212.

Briec, W., and B. Lemaire (1999): Technical Efficiency and Distance to a Reverse Convex Set, *European Journal of Operational Research* 114, 178 – 187.

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

Charnes, A. and W. W. Cooper (1985): 'Preface to topics in data envelopment analysis.', *Annals of Operations Research* 2: 59-94.

Charnes, A., W. Cooper, B. Golany, L. Seiford, J. Stutz (1985): Foundations of Data Envelopment Analysis for Pareto-Koopmans Efficient Empirical Production Functions, *Journal of Econometrics*, 30(1), 91 – 107

Charnes, A., W. Cooper, Z. M. Huang and D. B. Sun (1990): Polyhedral cone-ratio DEA models with an illustrative application to large commercial banks, *Journal of Econometrics* 46, 73-91.

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 G.T. Post (1999): *Extending the Farrell Framework to Graph Space: Synthesis of Radial Efficiency Measurement*, unpublished manuscript submitted to Economics Letters

Cherchye, L., T. Kuosmanen, and G.T. Post (forthcoming): What is the Economic Meaning of FDH? A Reply to Thrall, *Journal of Productivity Analysis*

Cooper, W.W., K. S. Park, and J. T. Pastor (1999): RAM: A Range Adjusted Measure of Inefficiency for Use with Additive Models, and Relations to Other Models and Measures in DEA, *Journal of Productivity Analysis* (to appear in 1999)

De Borger, B., G. D. Ferrier, and K. Kerstens (1998): The Choice of a Technical Efficiency Measure on the Free Disposal Hull Reference Technology: A Comparison Using US Banking Data, *European Journal of Operational Research* 105, 427-446

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.

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

Färe, R. (1988): *Fundamentals of Production Theory*, Springer-Verlag, Berlin

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

Färe R., and D. Primont (1995): *Multi-Output Production and Duality: Theory and Applications*, Kluwer Academic Publishers, Boston.

Koopmans, T. (1951): *Activity Analysis of Production and Allocation*, John Wiley & Sons Inc., New York.

Kopp, R. J. (1981): The Measurement of Productive Efficiency: A Reconsideration, The Quarterly Journal of Economics, 96(3), 477-503

Kuosmanen, T. (1999): *Data Envelopment Analysis of Non-convex Technology: With an Application to Finnish Super League Pesis Players*, Helsinki School of Economics and Business Administration, Working Papers, W-224

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

Park, B.U., L. Simar and C. Weiner (1997): 'FDH efficiency scores from a stochastic point of view', Discussion Paper 9715, Institut de Statistique, Universtite Catholique de Louvain, under revision for *Econometric Theory*.

Post, G.T. (1998): 'Estimating Non-Convex Production Sets: Imposing Convex Input Sets and Output Sets in Data Envelopment Analysis', Rotterdam Institute for Business Economics Studies (RIBES) Report, R9847, ISBN 90-5086-303-5, under revision for *European Journal of Operations Research.*

Ray, S. (1997): Weak Axiom of Cost Dominance: A Nonparametric Test of Cost Efficiency without Input Quantity Data, *Journal of Productivity Analysis*, 8(2), 151-165

Russell, R. (1985): Measures of Technical Efficiency, *Journal of Economic Theory* 35, 109-126

Shephard, R.W. (1953): *Cost and Production Functions*, Princeton University Press, Princeton

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*