Journal of Productivity Analysis, 15, 201–215, 2001

FDH Directional Distance Functions

with an Application to European Commercial Banks

LAURENS CHERCHYE laurens.cherchye@econ.kuleuven.ac.be
Center for Economic Studies, Catholic University of Leuven, Naamsestraat 69, 3000 Leuven, Belgium

TIMO KUOSMANEN kuosmane@hkkk.fi
Department of Economics and Management Science, Helsinki School of Economics and Business Administration, P.O. Box 1210, FIN-00101 Helsinki, Finland

THIERRY POST gtpost@few.eur.nl
Erasmus Research Centre of Management (ERIM) and Department of Finance and Accounting, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands

Abstract

We extend Free Disposable Hull (FDH) efficiency analysis towards the general directional distance function framework. The profit interpretation of directional distance functions is extended to the non-convex FDH technologies. In addition, we derive an efficient enumerative algorithm for computing distance measures in Free Disposable Hull (FDH) technologies, which applies to the entire (infinitely large) family of directional distance functions. A simple numerical example and an application to European commercial banks illustrate the algorithm.

Keywords: Free Disposable Hull (FDH), Data Envelopment Analysis (DEA), directional distance function, enumeration, banks

1. Introduction

*F*ree *D*isposable *H*ull (FDH) (Deprins, Simar, and Tulkens, 1984) efficiency analysis has been advocated e.g. by Tulkens (1993) as an alternative approach to the popular Data Envelopment Analysis (DEA: Charnes, Cooper, and Rhodes, 1978). Deviating from most DEA models, FDH does not restrict itself to convex technologies, but builds solely on a minimal assumption of free disposability. This orientation seems attractive since it is frequently difficult to find a good theoretical or empirical justification for postulating convex production sets in efficiency analysis. For example, Farrell (1959) stressed indivisibility of inputs and outputs and economies of scale and specialization as possible violations of convexity. More recent discussion and empirical evidence can be found e.g. in Deprins et al. (1984), Tulkens (1993), Simar and Wilson (2000), Cherchye, Kuosmanen, and Post (2000), and Kuosmanen (2001).

The FDH approach is nowadays well-known in the DEA field. For example, in a survey conducted among DEA experts by Seiford (1996), Tulkens' 1993 article was listed as one of the most influential DEA papers. Still, FDH has not enjoyed a similar success as DEA. It only takes a marginal position in applied work. For example, the Berger and Humphrey (1997) survey finds that 62 out of 67 nonparametric efficiency studies for financial institutions used convex DEA approximations, whereas only the remaining 5 employ non-convex FDH approximations. Also, most standard textbooks treat FDH in the margin (e.g. Färe, Grosskopf and Lovell, 1994), or completely ignore it (e.g. Charnes *et al.*, 1994). Some authors have even raised explicit doubts about the economic meaning of FDH (see e.g. the recent exchange between Thrall, 1999, and Cherchye *et al.*, 2000). Yet another reservation often related to the practical implementation of FDH is the non-linear nature of the associated programming problems. DEA models can typically be solved through standard Linear Programming techniques. By contrast, FDH problems generally take a Mixed Integer Linear Programming (MILP) form, and MILP problems are usually more costly from a computational point of view. In this respect, an obstacle that may prevent practitioners from choosing FDH can be that commercial software packages do not include FDH as an option (see e.g. the review by Hollingsworth, 1997).[1]

However, not all MILP problems are complicated to solve as they can often be conveniently enumerated. As for FDH, Tulkens (1993) first presented tailored enumerative formulations for efficient computation of the classic Farrell (1957) input and output measures. Fried, Lovell, and Vanden Eeckaut (1995) and Lovell (1994, 1995) have proposed other enumeration algorithms. From a computational point of view, these FDH enumeration algorithms have proved far less demanding than the general MILP codes available, and even superior to the LP codes available for standard DEA (see e.g. Lovell and Vanden Eeckaut, 1993). However, enumerative formulations specially tailored for FDH are currently available for a limited number of efficiency and distance gauges only (such as the input and output oriented Farrell measures). Computing a non-standard efficiency measure, such as McFadden's (1978) radial *gauge function*, may become an insurmountable task for a non-specialist practitioner who lacks a good integer programming software or programming skills.

To eliminate some of these barriers, we generalise the FDH approach towards the general *directional distance function* framework recently introduced by Chambers, Chung and Färe (1996, 1998).[2] The general directional distance function encompasses almost all known efficiency or distance measures as special cases. Notable examples include the classic Farrell input and output measures and the McFadden gauge function. To demonstrate that FDH technical efficiency measures are not void of economic meaning, we will reformulate the profit interpretation of directional distance functions derived in Chambers et al. (1998) for non-convex FDH approximations. In addition, we derive an efficient FDH algorithm that applies to the whole (infinitely large) family of efficiency and distance measures that can be obtained as special cases of the general directional distance function. An attractive feature, which is of particular importance to students and non-specialists, is that enumerating FDH distance measures by this simple formulation requires neither advanced integer programming software nor particular programming skills—spreadsheet software typically suffices. In addition to computational convenience, the enumerative formulation is useful for sensitivity analysis. Specifically, as a by-product we see how sensitive the distance measure is with respect to changes in input-output data or in the direction of measurement.

The rest of the paper unfolds as follows. In the next section we introduce the necessary notation and define the basic concepts. We extend the dual interpretation of the directional distance function to the FDH approximation in Section 3. In Section 4 we derive the enumeration formula. In Section 5 we illustrate the usefulness of the proposed formula by a simple numerical example. To further illustrate the power of these FDH techniques, Section 6 presents an empirical application to European commercial banks. That application also suggests how the directional distance function results might be used for obtaining empirical evidence of the convexity hypothesis. Finally, Section 7 summarizes our conclusions.

2. Preliminaries

For expositional convenience, we will compare the general directional distance function with the classic input distance function of Shephard (1953). The input distance function is defined as

$$D_T(y, x) = \underset{\theta}{Sup}\{\theta \mid (y, x/\theta) \in T\}, \tag{1}$$

where T is the *production possibility set*[3]

$$T = \left\{(y, x) \in \Re^{m+l} \mid \text{input } x \in \Re^l \text{ can produce output } y \in \Re^m\right\}. \tag{2}$$

The frequently employed Farrell input efficiency measure is simply the inverse of (1).

A more general distance metric, which contains (1) as its special case, is the *directional distance function* introduced by Chambers et al. (1996, 1998). These authors propose to measure distance to the production frontier in some pre-assigned input-output direction $(g^y, g^x) \in \Re^{m+l}$. This directional distance function can be defined as

$$\delta_T(y, x; g^y, g^x) = \underset{\theta}{Sup}\{\theta \mid (y + \theta g^y, x - \theta g^x) \in T\}. \tag{3}$$

The input distance function (1) is obtained from (3) as $D_T(y, x) = (1 - \delta_T(y, x; 0, x))^{-1}$. Similarly, the corresponding output distance function is obtained as $1 + \delta_T(y, x; y, 0)$.

Yet another interesting special case of the directional distance function (which will return in our numerical example (Section 5) and application (Section 6) below) is the McFadden (1978) gauge function, obtained as $(1+\delta_T(y, x; y, -x))^{-1}$. Somewhat counter-intuitively, this direction of measurement simultaneously augments both the input and the output vector. However, this gauge function is in some instances an attractive alternative to the oriented Farrell input and output measures due to its dual relationship to profit functions (see McFadden, 1978). In the context of efficiency measurement, the duality result implies that this orientation of measurement is consistent with the notion of *profit efficiency* (defined as a ratio of the actualized profit to the maximum profit) á la Nerlove (1965), and hence also for decomposing measured profit inefficiency into components of allocative and technical efficiency in the sense of Farrell (1957). In fact, it can be shown that the McFadden gauge function yields a theoretical upper bound for profit efficiency of profit maximizing firms which are known to earn strictly positive profit at given non-negative prices. Intuitively, equiproportionate scale augmentation is desirable for profitable (but inefficient) firms, since multiplying inputs and outputs by some number will then immediately increase the profit by the same factor.

In empirical studies, the production set T is typically unknown. In the nonparametric literature, empirical approximations for T are constructed from a set of observations drawn from the technology. The production data is henceforth represented by the output matrix $Y = (y^1 \cdots y^n)^T$, with $y^j = (y_1^j \cdots y_l^j)$, and the input matrix $X = (x^1 \cdots x^n)^T$, with $x^j = (x_1^j \cdots x_m^j)$. In addition, we use the index sets $L = \{1, \ldots, l\}$ for inputs, $M = \{1, \ldots, m\}$ for outputs, and $S = \{1, \ldots, n\}$ for production units.

The FDH approximation is based on two minimal assumptions concerning T. First, T is assumed to envelop all observed data. Second, T is assumed to be monotonous, i.e. *all* inputs and outputs are freely (or strongly) disposable. In other words, the marginal products of inputs, marginal rates of substitution between inputs, and marginal rates of transformation between outputs are assumed to be non-negative. Of course, congestion of production factors can violate this assumption. But as Färe and Grosskopf (1983) have pointed out, imposing monotonicity can be interpreted as a congestion adjustment of the production set, i.e. distance functions for 'monotonized' technologies include both pure technical efficiency and congestion components. Alternatively, monotonicity can be motivated by the fact that it does not interfere with the Pareto-Koopmans classification of technical efficiency.

The minimal set that complies with these two assumptions gives the FDH approximation, formally defined as

$$\hat{T}_{FDH}^{Y,X} = \{(y, x) \mid x \geq \lambda X; y \leq \lambda Y; \lambda e = 1; \lambda^j \in \{0, 1\} \quad \forall j \in S\}. \tag{4}$$

Note the similarity between $\hat{T}_{FDH}$ and the frequently employed convex monotone hull, which differs from FDH only in that it allows for convex combinations of production plans. Specifically, we obtain the convex monotone hull from (4) by substituting the constraint $\lambda^j \in \{0, 1\} \forall j \in S$ by $\lambda \geq 0$.

Distance functions and efficiency gauges can be measured relative to the FDH approximation when the true technology is unknown. For instance, substituting T by the empirical FDH approximation in (3) gives the following empirical distance function for some direction $(g^y, g^x) \in \Re^{m+l}$

$$\begin{aligned}&\hat{\delta}_{FDH}^{Y,X}(y, x; g^y, g^x)\\ &\quad = \underset{\theta}{Sup}\{\theta \mid x - \theta g^x \geq \lambda X; y + \theta g^y \leq \lambda Y; \lambda e = 1; \lambda^j \in \{0, 1\} \forall j \in S\}.\end{aligned} \tag{5}$$

Provided that the assumptions of envelopment and monotonicity are satisfied, the FDH set is contained within the true production set, i.e. $\hat{T}_{FDH}^{Y,X} \subseteq T$, and consequently $\hat{\delta}_{FDH}^{Y,X}(y, x; g^y, g^x) \leq \delta_T(y, x; g^y, g^x)$ for all (g^y, g^x) and $(y, x) \in T$. In other words, using the FDH approximation gives a lower bound for the true distance function value, which is an important property for many practical purposes.

3. Some Duality Results

One of the attractive features of the standard input distance function is its dual interpretation in terms of cost efficiency. Chambers et al. (1998) derived a similar dual interpretation of the directional distance function as a normalised profit difference. However, these authors restricted attention to convex technologies. We here extend this result for the monotonous but non-convex FDH set.

We call the set $M(j) = \{(x, y) \mid x \geq x_j, y \leq y_j\}$ the monotone hull of observation j. Using the fact that $\hat{T}_{FDH}^{Y,X}$ is simply the union of monotone hulls of all observations, i.e. $\hat{T}_{FDH}^{Y,X} = \bigcup_{j \in S} M(j)$, expression (5) can equivalently be written as

$$\hat{\delta}_{FDH}^{Y,X}(y, x; g^y, g^x) = \underset{j \in S}{Sup}\left[\underset{\theta}{Sup}\left\{\theta \mid \theta g^x \leq x - x^j; \theta g^y \leq y^j - y\right\}\right]. \tag{6}$$

Using duality theory, one could also express the embedded problem for each $j \in S$ as

$$\begin{aligned}&\underset{\theta}{Sup}\left\{\theta \mid \theta g^x \leq x - x^j; \theta g^y \leq y^j - y\right\}\\ &\quad = \underset{(\mu,v)\in\Re_+^{m+l}}{Inf}\left\{\mu(x - x^j) + v(y^j - y) \mid \mu g^x + v g^y = 1\right\}.\end{aligned} \tag{7}$$

As a result, we can re-express (6) as

$$\begin{aligned}&\hat{\delta}_{FDH}^{Y,X}(y, x; g^y, g^x)\\ &\quad = \underset{j \in S}{Sup}\left[\underset{(\mu,v)\in\Re_+^{m+l}}{Inf}\left\{(v y^j - \mu x^j) - (v y - \mu x) \mid \mu g^x + v g^y = 1\right\}\right].\end{aligned} \tag{8}$$

This expression yields an economic interpretation of the FDH distance function as a normalised difference between reference and actual profit.[4] It is to be recalled that FDH can account for a broad range of (possibly non-linear) objective functions (see e.g. Cherchye, Kuosmanen and Post (2000) and Wunsch (1996)).[5] Evidently, due to the specific construction of the FDH reference set, all shadow weight will usually be attributed to only one input or output (i.e. only one component of the vector (μ, v) will be strictly positive). However, if some additional (potentially incomplete) information about possible (relative) price, cost or revenue levels is available, we can impose additional restrictions to the shadow prices in problem (8) (see e.g. Pedraja-Chaparro, Salinas-Jimenez and Smith (1997) for a review of alternatives, and Kuosmanen and Post (1999) for a specific application).

4. Enumeration

An attractive feature of the FDH set from a practical point of view is that for each inefficient producer (situated within the strict interior of the FDH set) a physically existing reference unit can be identified that proves superior in all input and output dimensions. This is not generally true for the standard (convex) DEA models. However, the constraint $\lambda^j \in \{0, 1\} \forall j \in S$ that distinguishes FDH from convex DEA models also causes computational inconvenience. In general, computing FDH-based distance functions involves Mixed Integer Programming, which is usually computationally more demanding than Linear Programming. To remedy this problem, Tulkens (1993) first provided a simple and effective enumeration algorithm tailored for the input and output distance functions in FDH. In this procedure, we first identify the reference set, which is the set of observations weakly dominating the evaluated production vector (y, x), i.e.

$$R(y, x) = \left\{ i \in S \mid y^i \geq y; x^i \leq x \right\}. \tag{9}$$

Following Tulkens (1993), the input distance function can be computed as

$$\hat{D}^{Y,X}_{FDH}(y, x) = \underset{i \in R(y,x)}{Max} \left[\underset{j \in L}{Min} \left\{ \frac{x_j}{x^i_j} \right\} \right]. \tag{10}$$

This can be solved easily by enumeration. For a similar formulation for the output distance function, see Tulkens (1993).

We next extend this enumerative principle to the more general directional distance function. As a preliminary step, note that in general the value of the directional distance function for a production vector contained within the production set lies in the half-open interval $[0, \infty)$, and hence need not necessarily be finite. In the following exposition, we will call the directions $(g^y, g^x) \leq 0$ "trivial" since free disposability immediately implies that the concomitant value of $\hat{\delta}^{Y,X}_{FDH}(y, x; g^y, g^x)$ will be infinitely large. The following lemma proves finiteness in the remaining "non-trivial" cases.

LEMMA 1 *For any real-valued $(y, x) \in \hat{T}^{Y,X}_{FDH}$, the value of the distance function $\hat{\delta}^{Y,X}_{FDH}(y, x; g^y, g^x)$ will be a finite nonnegative real number for all non-trivial direction vectors $(g^y, g^x) \not\leq 0$.*

Proof. The reference set $\hat{T}^{Y,X}_{FDH}$ is bounded from above in output space by a finite vector $\bar{Y} \geq y^j \forall j \in S$ and is bounded from below in input space by a finite vector $\bar{X} \leq x^j \forall j \in S$. If there exists an output direction k with $g^y_k > 0$, then the condition $y_k + \vartheta g^y_k \leq \bar{Y}_k$ is satisfied only for a finite nonnegative ϑ. Alternatively, if there exists an input direction i with $g^x_i > 0$, then the condition $x_i - \vartheta g^x_i \geq \bar{X}_i$ is satisfied only for a finite nonnegative ϑ. As $\hat{T}^{Y,X}_{FDH}$ is contained within $(\bar{Y} - \Re^l_+, \bar{X} + \Re^m_+)$, we thus have that $\hat{\delta}^{Y,X}_{FDH}(y, x; g^y, g^x) \leq \vartheta$, and ϑ was already shown to be a finite nonnegative real number, if at least one component of the vector (g^y, g^x) is strictly positive, or $(g^y, g^x) \not\leq 0$. Hence, $\hat{\delta}^{Y,X}_{FDH}(y, x; g^y, g^x)$ must be finite. ■

Suppose we evaluate a vector $(y, x) \in \hat{T}^{Y,X}_{FDH}$, which by Lemma 1 suffices to guarantee that the distance function value is computable in all non-trivial directions. This allows us to write (6) equivalently as

$$\hat{\delta}^{Y,X}_{FDH}(y, x; g^y, g^x) = \underset{\theta \in \Re;\, j \in S}{Max} \left\{ \theta \mid y + \theta g^y \leq y^j; x - \theta g^x \geq x^j \right\}. \tag{11}$$

We can now prove the following:

PROPOSITION 1 *For all non-trivial direction vectors $(g^y, g^x) \not\leq 0$ and $(y, x) \in \hat{T}^{Y,X}_{FDH}$, the value of the directional distance function can be computed as:*

$$\hat{\delta}^{Y,X}_{FDH}(y, x; g^y, g^x) = \underset{j \in Q(y,x;g^y,g^x)}{Max} \alpha^j, \tag{12}$$

where

$$Q(y,x;g^y,g^x) = \begin{cases} S, & if\, g_k^y \neq 0 \forall k \in M \wedge g_i^x \neq 0 \forall i \in L, \\ V, & if\, \exists k \in M\colon\ g_k^y = 0 \vee i \in L\colon\ g_i^x = 0 \end{cases}$$

$$V = \left\{ j \in S \mid y_k^j - y_k \geq 0 \forall g_k^y = 0, k \in M;\, x_i - x_i^j \geq 0 \forall g_i^x = 0, i \in L \right\}$$

$$\alpha^j = \begin{cases} \beta^j, & \beta^j \geq \gamma^j \\ 0, & \beta^j < \gamma^j, \end{cases}$$

$$\beta^j = \underset{\substack{k \in M: g_k^y > 0 \\ i \in L: g_i^x > 0}}{Min} \left\{ \left(\frac{y_k^j - y_k}{g_k^y} \right), \left(\frac{x_i - x_i^j}{g_i^x} \right) \right\},$$

and

$$\gamma^j = \begin{cases} \underset{\substack{k \in M: g_k^y < 0 \\ i \in L: g_i^x < 0}}{Max} \left\{ \left(\frac{y_k^j - y_k}{g_k^y} \right), \left(\frac{x_i - x_i^j}{g_i^x} \right) \right\}, & if\ \exists k \in M\colon\ g_k^y < 0 \vee i \in L\colon\ g_i^x < 0 \\ 0, & otherwise. \end{cases}$$

Proof. If $g_k^y = 0$, then the restriction $y_k + \theta g_k^y \leq y_k^j$ reduces to $y_k \leq y_k^j$. Similarly, if $g_i^x = 0$, then $x_i - \theta g_i^x \geq x_i^j$ reduces to $x_i \geq x_i^j$. Hence, the reference unit is restricted to the DMUs in $Q(y,x;g^y,g^x)$. In addition, the restrictions $y_k + \theta g_k^y \leq y_k^j, k \in M\colon\ g_k^y > 0$, and $x_i - \theta g_i^x \geq x_i^j, i \in L\colon\ g_i^x > 0$, can be expressed as $\theta \leq \beta^j$. Similarly, the restrictions $y_k + \theta g_k^y \leq y_k^j, k \in M\colon\ g_k^y < 0$, and $x_i - \theta g_i^x \geq x_i^j, i \in L\colon\ g_i^x < 0$, can be expressed as $\theta \geq \gamma^j$. Combining these insights gives

$$\hat{\delta}_{FDH}^{Y,X}(y,x;g^y,g^x) = \underset{\theta \in \Re;\, j \in Q(x,y;g^y,g^x)}{Max} \left\{ \theta \mid \gamma^j \leq \theta \leq \beta^j \right\}. \tag{i}$$

If $\gamma^j \leq \beta^j$, the solution is simply β^j. By contrast, if $\gamma^j > \beta^j$, there is no feasible solution for observation j. However, in this case α^j can harmlessly be set equal to zero since that is the minimal distance secured by the assumption $(y,x) \in \hat{T}_{FDH}^{Y,X}$. Combining these insights, we find that (i) equals (12). ■

This result characterizes the most general FDH enumerative formulation for computing directional distance measures in any non-trivial direction. Note that the generality of this formulation adds a considerable degree of complexity. The direction vectors $(g^y, g^x) > 0$ constitute a notable special case, where the reference set $Q(y,x;g^y,g^x) = S$ and $\gamma^j = 0 \forall j \in S$, and hence problem (12) reduces to the simpler form

$$\underset{j \in S}{Max}\, \beta^j. \tag{13}$$

The analogy with the Tulkens formulation (10) follows immediately.

At least two substantial improvements to the existing FDH enumeration formulations should be noted. Firstly, this same formulation applies to the entire (infinitely large) family

Table 1. The example data set.

Observation j	A	**B**	C	D	E	F	G	H	I
Input x	3	6	4	6	5	8	12	14	18
Output y	4	5	6	7	8	9	11	13	14

of directional distance measures, whereas existing formulations are limited to particular orientations of measurement. In fact, problem (12) encompasses all known FDH enumeration schemes as special cases. Secondly, the formulation (12) does not depend on the sets of dominating observations like all earlier formulations. This feature allows for computing distance measures with respect to those parts of the FDH frontier that do not dominate the evaluated input- output vector in all input-output dimensions, i.e. directions $(g^y, g^x) \ngeq 0$ and $(g^y, g^x) \nleq 0$. In practice, reduction of some output or increase of some input could be favored e.g. in situations characterized by undesirable outputs (such as pollution) or desirable inputs (such as labor input during severe unemployment). Another interesting example of this class is the McFadden gauge function, corresponding to $(g^y, g^x) = (y, -x)$ (discussed in Section 2). Furthermore, this feature is convenient for assessing sensitivity of the computed direction function values with respect to changes in the input-output data or the direction vector, as we will illustrate in our next section.

5. Numerical Example

We illustrate the practical use of the enumeration formulation (12) by a simple numerical example. Consider the single-input single-output data set of 9 observations reported in Table 1.

For sake of brevity, we focus on observation B. To illustrate the use of enumeration for other measures than the standard Farrell input and output measures, we consider the direction vector $(g^y, g^x) = (5, -6)$, corresponding to the McFadden gauge function (see Section 2 for further discussion on the economic interpretation of this measure). That is, we look for the maximal equiproportionate augmentation of both the input and the output of B. A GAUSS computer code for calculating the values of this gauge function by enumeration is reported in the Appendix.

Figure 1 illustrates the McFadden projection path for observation B on the FDH frontier. The optimal reference production plan is H. The example demonstrates how tricky the computation of the McFadden gauge function with respect to the FDH technology can be, even in this simple setting. For example, the optimal reference H does not dominate observation B in the input dimension, and hence the heuristics underlying the Tulkens enumerative principle do not apply. Furthermore, the gauge function has two additional local optima (at observations F and G), not dominated by the global optimum at H. In addition, observation I attains globally the highest output level as well as a higher input level than B, but this does not yet make it the correct reference point in this case. We suspect that these features may render usual interior point methods for solving MILP problems useless.

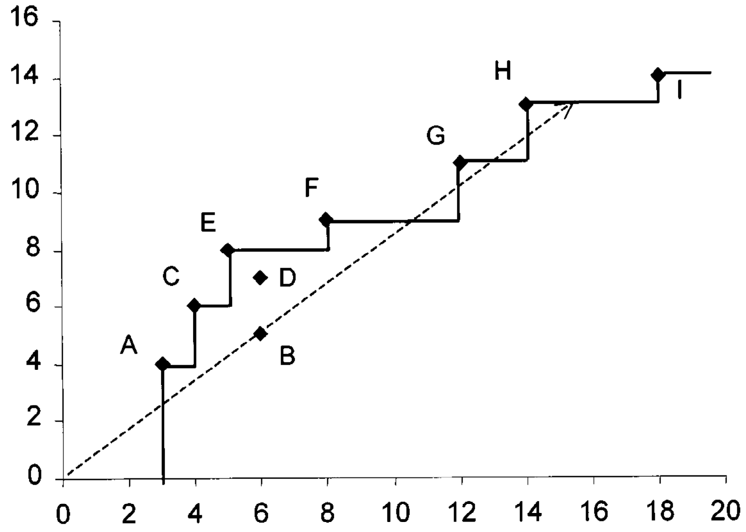

Figure 1. McFadden gauge of B relative to FDH.

As the direction vector is composed of non-zero elements only, the reference set $Q(\cdot)$ includes all observations. Table 2 reports the auxiliary indicators α^j, β^j, and γ^j, $j = A, \ldots, I$. Recall that β involves the positive directions (here output) and γ involves the negative directions (here input). Specifically, γ^j is obtained as $(6 - x^j)/-6$, i.e. the normalized difference between the inputs of observation B and observation j. Similarly, β^j is obtained as $(y^j - 5)/5$, i.e. the normalized difference between the outputs of observation j and observation B. In the bottom row, α^j equals β^j if $\beta^j \geq \gamma^j$ and is set equal to zero in the other case. In this example, only for observation I we have that $\beta^I < \gamma^I$ (compare with Figure 1). The distance function value is then obtained as the maximum of the α^j in the bottom row. This yields $8/5$ as an outcome, obtained relative to observation H.

Interestingly, the auxiliary indicators reported in Table 2 can be used for sensitivity analysis. Suppose for example that we are interested in the robustness of this result to changes in the input-output data. Using the insights of Proposition 1, we observe directly from Table 1 that the computed distance measure is relatively robust to changes in observations A, C, D, E and F. By contrast, small changes in observations G and I, and of course B itself, could

Table 2. Auxiliary indicators for enumeration.

Observation j	A	B	C	D	E	F	G	H	I
γ^j	$-1/2$	0	$-1/3$	0	$-1/6$	$1/3$	1	$4/3$	2
β^j	$-1/5$	0	$1/5$	$2/5$	$3/5$	$4/5$	$6/5$	$8/5$	$9/5$
α^j	$-1/5$	0	$1/5$	$2/5$	$3/5$	$4/5$	$6/5$	**$8/5$**	0

influence the result at least in principle. To analyse how sensitive the efficiency result is with respect to such changes, we can scrutinise these observations, hereby paying special attention to the difference quotients underlying the auxiliary indicators tabulated in Table 2.

This investigation reveals that if observation I only increased its output by 1, the distance function value would change to 2. Similarly, if the same observation decreased its input by 1.2 then the resulting distance measure would equal 9.5. However, for observation I marginal input and output changes do not generate a direct effect. For observation H, on the other hand, marginal output changes influence the result immediately by factor $1/5$, while input could harmlessly decrease by any number, or increase by 1.6, without affecting the distance measure value. The input of B itself could change within tolerance limits $(-0.61, +0.42)$ without influencing the distance measure. Finally, the distance is inversely proportional to the output of B within the tolerance limits $(-0.33, +0.57)$.

It should be clear that changes in the direction vector values have an analogous impact on the directional distance function value as changes in the input-output data. In a straightforwardly similar way, one could use the auxiliary results in Table 2 to investigate sensitivity of the resulting distance function values with respect to changes in the direction vector. As we have demonstrated, such sensitivity analysis may reveal interesting information and is made tractable by the specific enumerative formulation (12), which in contrast to the Tulkens enumerative principle is not based on the set of dominating observations.

6. Application to European Commercial Banks

The nonparametric approach to efficiency measurement has seen extensive application in the financial industry. See e.g. Berger and Humphrey (1997) for a survey. To illustrate the use of our enumeration formulation (12) in large samples, we performed an empirical application for European commercial banks. Specifically, we used a data set with 1997 financial statement data of the 100 largest commercial banks in the European Union.[6] This application extends and complements the study initiated in Kuosmanen and Post (1999).

For convenience, we use a simplified representation of the bank technology, which involves a single output, *total earning assets*, and three inputs, *equity capital*, *debt capital* and *operational costs* (which aggregates all inputs apart from equity and debt). All variables are measured in millions of Euro. Figure 2 summarises the technology, and table 3 presents some descriptive statistics for the data set.

Using the relevant prices, our data set aggregates all earning assets and physical inputs into measures of total earning assets and total operating costs respectively. 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 cannot be inferred from *ex post* financial statements. Therefore, debt and equity capital are included as separate variables.

Like in our previous example we have computed the McFadden gauge function values for all observations in the sample. For sake of comparison, we did so for both the FDH and the convex monotone hull (CMH) approximations of the production set (see Section 2). Recall

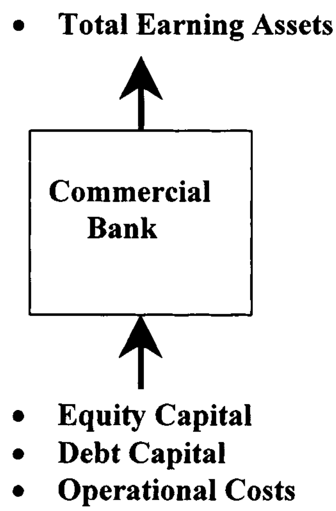

Figure 2. Graphical representation of the bank technology.

Table 3. Descriptive statistics for the bank sample.

	Tot. Earn. Ass.	Debt Cap.	Eq. cap.	Op. Costs
Mean	3884.77	88939.14	1858.21	92841.60
Median	2320.70	48494.25	944.27	49734.25
St. Dev.	4351.41	104015.20	2241.25	108802.00
Skewness	3.17	2.21	2.13	2.14
Kurtosis	14.64	5.40	4.52	4.85
Minimum	819.92	6741.91	5.76	6317.96
Maximum	30808.54	568501.80	10433.01	574533.80

that both these technologies exhibit variable returns to scale, and differ only with respect to the convexity property. Table 4 reports the summary statistics for the resulting FDH and CMH efficiency estimates.[7]

On average, the efficiency estimates are quite low in both models. This is at least partially explained by the fact that the equiproprtionate scale augmentation built in the McFadden measure takes into account both pure technical inefficiencies (i.e. the deviations from the best-practice) and scale inefficiencies (i.e. under-utilisation of resources). In the earlier study (Kuosmanen and Post, 1999) we analysed the same banks in terms of pure technical efficiencies (i.e. Farrell input measure) using the FDH technology. In that study, the efficiency scores of almost all banks included in the current sample were found to lie between 0.9 and 1 (mean 0.988). This finding suggests that most of the inefficiencies found in this study are due to scale inefficiencies, and that there exists considerable potential to utilise economies of scale. This conclusion is also consistent with the recent scale increases by mergers and acquisitions in the financial sector in Europe.

Table 4. Statistics on the FDH and CMH efficiency estimates.

	FDH	**CMH**
Mean	0.560	0.460
Median	0.510	0.373
St. Dev.	0.336	0.314
Skewness	0.131	0.529
Kurtosis	−1.524	−1.074
Minimum	0.052	0.052
Maximum	1.000	1.000

The FDH results deviate quite substantially from the CMH results. This observation, and the possible presence of economies of scale (as discussed above), together suggest that the convexity assumption (used for constructing the CMH set but not the FDH set) is questionable. It is therefore interesting to formally test whether or not the efficiency scores of the FDH model are significantly different from the CHM results. Therefore, we performed the five tests that were analysed by Kittelsen (1999), i.e. the two F-tests and the Kolmogorov-Smirnov-type test proposed by Banker (1993), and two t-tests. All five tests reject the convexity postulate at a significance level of at least 90 percent. Notwithstanding the problems associated with comparing the results of two nested models that are applied to a common data set (see Kittelsen, 1999), these results clearly put the convexity postulate into question.

7. Conclusions

We have generalised the directional distance function framework towards FDH efficiency analysis. We have reformulated the profit interpretation of directional distance function for the non-convex FDH approximation. In addition, we have derived a general enumeration formula for computing directional distance functions relative to the FDH reference technology. This enumeration formula can compute distance functions in any meaningful direction by enumeration, and hence can help to avoid computational problems associated with FDH. Apart from computational convenience, the enumerative formulation proves useful for sensitivity and robustness analysis with respect to data changes and changes in the direction vectors. Our application shows that the enumeration algorithm is easily performed for large samples. Furthermore, we have illustrated how the directional distance measures can be used for making inference about the validity of convexity assumptions.

APPENDIX

In this appendix we present a GAUSS computer code for computing

$$\delta_{FDH}^{Y,X}(y^j, x^j; y^j, -x^j)\, j \in S.$$

```
PROC(1)=McFadden(Y,X);

LOCAL a,b,c,d,e,j,n;

n=ROWS(X);
d=ONES(n,1)*9999;
j=1;
DO WHILE j .LE n;

    b=MINC(((Y-Y[j,.])./Y[j,.])');
    c=MAXC(((X-X[j,.])./X[j,.])');
    e=b~c;
    e=SELIF(e, e[.,1] .GE e[.,2]);
    a=e[.,1];
    d[j]=MAXC(a);
    j=i+1;

ENDO;
RETP(d);
ENDP;
```

The explanation proceeds as follows. The vector d contains the directional distance function value for each observation $j \in S$, i.e. $d = (\hat{\delta}^{Y,X}_{FDH}(y^1, x^1; y^1, -x^1) \cdots \hat{\delta}^{Y,X}_{FDH}(y^n, x^n; y^n, -x^n))^T$. In addition, we use the following notation: $b = (\beta^1 \cdots \beta^n)^T$ and $c = (\gamma^1 \cdots \gamma^n)^T$. Finally, $e = (b \;\; c)$ is an auxiliary matrix, used to construct a, which is a column vector containing α^j for all observations for which $\beta^j \geq \gamma^j$.

Acknowledgments

We thank two anonymous referees for helpful comments and suggestions.

Notes

1. A (non-commercial) software code that includes FDH is the EMS code developed by Holger Scheel (URL: http://www.wiso.uni-dortmund.de/lsfg/or/scheel/ems/).
2. Similar constructions have been introduced in another context by Luenberger (1992), and in efficiency analysis by Briec (1997), Chavas and Cox (1999), and Halme et al. (1999). For applications of directional distance functions see e.g. Chambers, Färe, and Grosskopf (1996), and Chung, Färe, and Grosskopf (1997).
3. For sake of generality, we do not restrict the sign of the inputs and outputs, but allow for ‘negative’ inputs or outputs. We thereby allow a production factor to simultaneously be input for one firm and output for another. The distinction between inputs and outputs thus becomes somewhat arbitrary, but for consistency with the standard notation and terminology, we choose to preserve it.
4. Chambers et al. (1998) obtained an *Inf-Sup* structure whereas we obtain a *Sup-Inf* structure. Indeed, when the technology would be truly convex then the order of the *Inf* and *Sup* operators could be reversed.

5. Of course, the precise interpretation of distance function in a more general setting depends on whether—and in what sense—the objective function value reacts in a *homogenous* way to input and output changes. The appropriate homogeneity condition further depends on the direction vector.
6. We use BankScope data provided by Bureau van Dijk Nederland.
7. More detailed estimation results can be obtained from the authors upon request.

References

Banker, R. D. (1993). "Maximum Likelihood, Consistency and Data Envelopment Analysis: A Statistical Foundation." *Management Science* 39, 1265–1273.

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. (1997). "A Graph-type Extension of Farrell Technical Efficiency Measure." *Journal of Productivity Analysis* 8, 95–110.

Chambers, R. G., Y. Chung, and R. Färe. (1996). "Benefit and Distance Functions." *Journal of Economic Theory* 70, 407–418.

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.

Chambers, R. G., R. Färe, and S. Grosskopf. (1996). "Productivity Growth in APEC Countries." *Pacific Economic Review* 1, 181–190.

Charnes, A., W. Cooper, A. Y. Lewin, and L. M. Seiford (eds.). (1994). *Data Envelopment Analysis: Theory, Methodology and Applications*. Kluwer Academic Publishers.

Chavas, J.-P., and T. Cox. (1999). "A Generalised Distance Function and the Analysis of Production Efficiency." *Southern Economic Journal* 66, 294–318.

Chung, Y., R. Färe, and S. Grosskopf. (1997). "Productivity and Undesirable Outputs: A Directional Distance Function Approach." *Journal of Environmental Management* 51, 229–240.

Cherchye, L., T. Kuosmanen, and G. T. Post. (2000). "What is the Economic Meaning of FDH? A Reply to Thrall." *Journal of Productivity Analysis* 13 (3), 259–263.

Deprins, D., L. Simar, and H. Tulkens. (1984). "Measuring Labor Efficiency in Post Offices." In M. Marchand, P. Pestieau, and H. Tulkens (eds.), *The Performance of Public Enterprises: Concepts and Measurements*, pp. 243–267. North Holland.

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

Farrell, M. J. (1959). "Convexity Assumption in Theory of Competitive Markets." *Journal of Political Economy* 67, 377–391.

Fried, H. O., C. A. K. Lovell, and P. Vanden Eeckaut. (1995). "Service Productivity in U.S. Credit Unions." In P. T. Harker (ed.), *The Service Productivity and Quality Change*, pp. 265–390. Boston: Kluwer.

Färe, R., and S. Grosskopf. (1983). "Measuring Congestion in Production." *Zeischrift fur Nationalökonomie* 43, 257–271.

Färe, R., S. Grosskopf, and C. A. K. Lovell. (1994). *Production Frontiers*. Cambridge: Cambridge University Press.

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.

Hollingsworth, B. (1999). "Data Envelopment Analysis and Productivity Analysis: A Review of the Options." *The Economic Journal* 109, 458–462.

Kittelsen, S. A. C. (1999). "Monte Carlo Simulations of DEA Efficiency Measures and Hypothesis Tests." Memorandum No. 9/1999, Department of Economics, University of Oslo.

Kuosmanen, T. (2001). "DEA with Efficiency Classification Preserving Conditional Convexity." *European Journal of Operational Research* (forthcoming).

Kuosmanen, T., and G. T. Post. (1999). "Measuring Economic Efficiency with Incomplete Price Information, With an Application to European Commercial Banks." Rotterdam Institute for Business Economic Studies, Report 9911, forthcoming in *European Journal of Operational Research*.

Lovell, C. A. K. (1994). "Linear Programming Approaches to the Measurement and Analysis of Productive Efficiency." *TOP* 2, 175–224.

Lovell, C. A. K. (1995). "Measuring the Macroeconomic Performance of Taiwanese Economy." *International Journal of Production Economics* 39, 165–178.

Lovell, C. A. K., and P. Vanden Eeckaut. (1993). "Frontier Tales: DEA and FDH." In E. W. Diewert, K. Spremann, and F. Stehlings (eds.), *Mathematical Modelling in Economics. Essays in Honor of Wolfgang Eichhorn*, pp. 446–457. Berlin: Springer-Verlag.

Luenberger, D. G. (1992). "Benefit Functions and Duality." *Journal of Mathematical Economics* 21, 261–481.

McFadden, D. (1978). "Cost, Revenue and Profit Functions." In M. Fuss and D. McFadden (eds.), *Production Economics: A Dual Approach to Theory and Applications*. Amsterdam: North-Holland.

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

Pedraja-Chaparro, F., J. Salinas-Jimenez, and P. Smith. (1997). "On the Role of Weight Restrictions in Data Envelopment Analysis." *Journal of Productivity Analysis* 8, 215–230.

Seiford, L. M. (1996). "Data Envelopment Analysis: The Evolution and the State of the Art (1978–1995)." *Journal of Productivity Analysis* 7, 99–137.

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

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.

Thrall, R. M. (1999). "What is the Economic Meaning of FDH?" *Journal of Productivity Analysis* 11(3), 243–250.

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

Wunsch, G. P. (1996). *Peer Comparison and Regulation*. Doctoral dissertation, CORE, Louvain-la-Neuve.