

Distinguishing Technical and Scale Efficiency on Non-Convex and Convex Technologies: Theoretical Analysis and Empirical Illustrations

Kristiaan Kerstens^{*}
Philippe Vanden Eeckaut^{}**

Abstract:

This paper defines a decomposition of technical efficiency for a series of nonparametric deterministic reference technologies related to the Free Disposal Hull. More specifically, introducing several returns to scale assumptions into this non-convex production model allows one to distinguish between technical and scale inefficiencies.

These technologies and the resulting efficiency decomposition are illustrated with several data sets and contrasted with results based on the traditional, convex Data Envelopment Analysis models. In particular, data on UK rates departments are extensively analysed. Furthermore, samples of French urban transit companies and of Belgian municipalities serve to verify certain empirical regularities.

JEL classification: D24.

Keywords: decomposition of technical efficiency; FDH; DEA.

An early version of this paper benefited from presentation at the conference “La méthode DEA (Data Envelopment Analysis) et l’analyse des performances des entreprises et des organisations” (Marseille, June 1997) and from the “Fifth European Workshop on Efficiency and Productivity Measurement” (Copenhagen, October 1997). In particular, comments by W. Briec, B. Dervaux, R. Färe, S. Grosskopf, H. Leleu and H. Tulkens are gratefully acknowledged. We bear complete responsibility for any errors.

* LABORES, Université Catholique de Lille, B.P. 109, 60 Boulevard Vauban, F-59016 Lille Cédex, France. Corresponding author.

** CORE, Université Catholique de Louvain, 34 Voie du Roman Pays, B-1348 Louvain-la-Neuve, Belgium.

December 1997
This revision: June 1998

1. INTRODUCTION

Efficiency and productivity measurements are nowadays popular tools in applied production analysis. Several types of frontier methodologies have been proposed and various ways of decomposing efficiency are available (see Lovell (1993)). Knowledge of the sources of efficiency and productivity is useful from both a policy-oriented and a theoretical point of view. These refined measurement schemes point to possible causes of inferior or superior performance. This paper determines two sources of inefficient behaviour using nonparametric deterministic frontier technologies that are derived from the Free Disposal Hull (FDH) (see Deprins, Simar and Tulkens (1984)). In particular, introducing several returns to scale assumptions into this non-convex production model allows identifying technical and scaling efficiencies. This decomposition is identical to the one proposed in Banker, Charnes and Cooper (1984), but more limited than the one presented in Färe, Grosskopf and Lovell (1983, 1985), since it does not contain a congestion component.

This paper also extends recent work aimed at partially relaxing the convexity assumptions (see Färe, Grosskopf and Njineu (1988), Petersen (1990), Bogetoft (1996) and Ray (1997)).¹ By dropping convexity altogether, the technologies and the resulting efficiency decomposition presented enlarge the methodological choices open to practitioners in the field. We believe this is very important, since there are indications that the convexity assumption has a great deal of impact in efficiency measurement. This is a consequence of the direct link between the volume of the production possibility set and the amount of technical inefficiency that can be revealed. Exploiting the relationship between efficiency measures and goodness-of-fit measures used for hypothesis testing (see, e.g., Färe and Grosskopf (1995) and Varian (1990)), we are able to derive a component summarising exactly the impact of the convexity axiom on efficiency results.

For example, using the same US banking sample as Ferrier and Lovell (1990), Ferrier, Kerstens and Vanden Eeckaut (1994) and De Borger, Ferrier and Kerstens (1998) use a traditional DEA model with variable returns to scale respectively a non-convex FDH model. While the former study obtains a mean radial input efficiency score of 0.370 the latter yields a 0.944 score.² Two studies systematically comparing several parametric and

¹ Färe, Grosskopf and Lovell (1994: 52-53) link the different piecewise reference technologies presented in the literature.

² Both studies did not utilise a series of environmental variables included in Ferrier and Lovell (1990).

nonparametric (including DEA and FDH models) methodologies confirm that efficiency results are consistent among models of the same “family”, that there can be large differences between parametric and nonparametric models, and –surprisingly-- that DEA and FDH models differ widely (see Cummins and Zi (1998) and De Borger and Kerstens (1996)). While convexity has traditionally been invoked as part of a calculus-based approach to economic and decision theory and measurement, we suggest this empirical evidence, as inconclusive as it may seem, should make us reconsider its use in performance gauging.

This paper unfolds as follows. Section 2 introduces the decomposition of technical efficiency based on convex nonparametric, deterministic frontier technologies (known as Data Envelopment Analysis (DEA) models). It also presents technologies based on the FDH non-convex production model and derives an analogous decomposition. Computational aspects of the decomposition are discussed in details in Section 3. This leads, among others, to a new, non-linear formulation of existing convex technologies that offers pedagogical advantages, since the role of the different assumptions is highlighted. The non-convex production models yield binary mixed integer, non-linear programming problems when determining radial efficiency. We prove that computations can make use of an implicit enumeration algorithm based upon a simple vector comparison procedure.

The next section (Section 4) offers, by way of illustration, empirical applications contrasting the traditional convex and the newly derived non-convex decompositions. The data used are a sample of UK rates departments. In addition, French urban transit companies and Belgian (Walloon region) municipalities provide two additional data sets to document empirical regularities. A final section concludes and provides directions for future research.

2. TECHNICAL AND SCALE EFFICIENCY IN CONVEX AND NON-CONVEX TECHNOLOGIES

2.1. Technical and Scale Efficiency: Definitions

The boundary of the production possibility set can be reconstructed using several methodologies (see, e.g., Lovell (1993)). Extensive efficiency decompositions have been developed for convex nonparametric deterministic reference technologies. In this field, the first operational procedure to measure technical and allocative efficiency originates in Farrell (1957). Färe, Grosskopf and Lovell (1983, 1985) offered a more elaborate

taxonomy of efficiency concepts. This paper focuses mainly on the distinction between technical and scale efficiencies.

Informally defined, technical efficiency (TE) requires production on the boundary of the production possibility set. If production occurs in the interior of a production possibility set, then a producer is technically inefficient. It is a private goal defined in terms of the best interest of the producer. A producer is scale efficient (SCE) if its size of production corresponds to a long run zero, profit competitive equilibrium configuration; it is scale inefficient otherwise. This social goal relates to a possible divergence between the actual and ideal configuration of inputs and outputs, the latter requiring production along constant returns to scale (CRS).

The production technologies are based on k observations (DMUs) using a vector of inputs $x \in \mathfrak{R}_+^n$ to produce a vector of outputs $y \in \mathfrak{R}_+^m$. Technology is represented by its production possibility or transformation set $GR = \{(x,y): x \text{ can produce } y\}$, i.e., the set of all feasible input-output vectors. An alternative representation of technology is the input requirement set: $L(y) = \{x: (x,y) \in GR\}$, where it follows: $x \in L(y) \Leftrightarrow (x,y) \in GR$.

Technologies can have several scaling assumptions. One crucial scaling assumption for our purpose is CRS: technology exhibits CRS if $\delta GR = GR$, $\delta > 0$.

Following Farrell (1957), efficiency is measured in a radial or equiproportional way. The radial input efficiency measure $DF_i(x,y)$ is:

$$DF_i(x,y) = \min \left\{ \lambda \mid \lambda \geq 0, \lambda x \in L(y) \right\}.$$

It indicates the maximum amount by which inputs can be decreased while still producing a given vector of outputs. $DF_i(x,y)$ is bounded above by one, with efficient production on the isoquant of $L(y)$ represented by unity. This efficiency measure can be used relative to different production models to obtain the above efficiency taxonomy.

Technical efficiency is usually measured relative to a production technology with the least restrictive returns to scale assumption. Scale efficiency is evaluated relative to the CRS technology, since this technology provides a long run competitive equilibrium benchmark. Efficiency measurement relative to the latter technology thus conflates scale and technical efficiencies. Traditionally, scale efficiency is defined as the ratio of two efficiency measures: one calculated on a CRS technology ($DF_i(x,y|CRS)$), and one computed on a variable returns to scale technology ($DF_i(x,y|VRS)$). Formally, input oriented scale efficiency measure ($SCE_i(x,y)$) is defined:

$$SCE_i(x,y) = DF_i(x,y \mid CRS)/DF_i(x,y \mid VRS).$$

This ratio indicates the lowest possible input combination able to produce the same output in the long run as a technically efficient combination situated on the VRS technology. Since $DF_i(x,y \mid CRS) \leq DF_i(x,y \mid VRS)$, evidently $0 < SCE_i(x,y) \leq 1$.³

Both of these static efficiency concepts are mutually exclusive and exhaustive and their radial measurement yields a multiplicative decomposition (see Färe, Grosskopf and Lovell (1985: 188-191)). Overall technical efficiency (OTE) differs from technical efficiency (TE) in that it is always measured relative to a strongly disposable CRS technology. This yields the identity: $OTE = TE \cdot SCE$, or: $DF_i(x,y \mid CRS) = DF_i(x,y \mid VRS) \cdot SCE_i(x,y)$.

Technology exhibits CRS at the observation under evaluation or at its input-oriented projection point when $SCE_i(x,y) = 1$. If $SCE_i(x,y) < 1$, then the evaluated observation is not located or projected on a piecewise linear segment where CRS prevail. For the latter observations, it is possible to determine the exact nature of returns to scale at its bounding hyperplane. Several methods exist to obtain this qualitative information regarding local scale economies. Computational aspects, however, are deferred to Section 3.

2.2. Measuring Technical and Scale Efficiency on Convex and Non-Convex Deterministic, Nonparametric Technologies

We focus on implementing the above decomposition using deterministic, nonparametric technologies based on activity analysis (see Koopmans (1951)). We first introduce the more traditional convex production models (DEA models), and then present the non-convex technologies.⁴

The first reference technology is a convex CRS production model:

$$GR^{DEA-CRS} = \{(x,y): M'z \geq y, \quad N'z \leq x, \quad z \in \mathfrak{R}_+^k\},$$

³ For the initial proposal, see Førsund and Hjalmarsson (1974, 1979). Färe, Grosskopf and Lovell (1983) stress that technical optimal scale, and not a price-dependent notion of profit-maximising scale, is used as the benchmark. See also Banker, Charnes and Cooper (1984).

⁴ The convex technologies are defined in Banker, Charnes and Cooper (1984), Färe, Grosskopf and Lovell (1983, 1985), among others. The FDH and FDH-based, nonconvex technologies have been described partly in Bogetoft (1996), Deprins, Simar and Tulkens (1984), and Tulkens (1993), among others. Other nonconvex technologies include, for instance, the Free Replicability Hull mentioned in Tulkens (1993) and the Free Affordability Hull in Ray (1997).

where M is the $k \times m$ matrix of observed outputs, N is the $k \times n$ matrix of observed inputs, z is a $k \times 1$ vector of intensity or activity variables, and y and x are $m \times 1$ and $n \times 1$ vectors of outputs respectively inputs. Since there is no restriction on the intensity vector z , CRS are imposed on this technology. It assumes strong input and output disposability.

In addition to this long run ideal, another more flexible model is needed to evaluate TE. This is done using a modified model imposing VRS. It results from the previous definition by restricting, in addition, the intensity variables to sum to one ($I'_k z = 1$, where I_k is a $k \times 1$ unity vector).

To determine returns to scale two more technologies need to be introduced. Convex technologies assuming non-increasing (NIRS) and non-decreasing returns to scale (NDRS) also add a constraint relative to the CRS model. In particular, NIRS and NDRS technologies require restricting the sum of the activity vector to be smaller respectively larger than or equal to unity ($I'_k z \leq 1$, respectively $I'_k z \geq 1$).⁵

The corresponding non-convex technologies can be defined as follows.⁶ A CRS, non-convex technology is defined as:

$$GR^{FDH-CRS} = \{(x, y): M'w \geq y, \quad N'w \leq x, \quad I'_k z = 1, \quad z_i \in \{0,1\}, \quad w_i = \delta z_i, \quad \delta \geq 0\}.$$

There is now one activity vector z operating subject to a non-convexity constraint and one rescaled activity vector w allowing for any scaling of the observations spanning the frontier. The scaling parameter (δ) is free because of the CRS assumption.

This model is closely related to the traditional FDH technology, initially proposed by Deprins, Simar and Tulkens (1984), that imposes strong disposability assumptions, but no specific returns to scale hypothesis:

$$GR^{FDH} = \{(x, y): M'z \geq y, \quad N'z \leq x, \quad I'_k z = 1, \quad z_i \in \{0,1\}\}.$$

FDH results from the previous model as a special case by fixing the scaling parameter (δ) at 1.

Similar to the convex case, it is also possible to define two more technologies based on the FDH assuming NIRS and NDRS. This simply requires modifying the scaling parameter included in the $GR^{FDH-CRS}$ technology to be smaller than or equal to unity ($\delta \leq 1$) respectively larger than or equal to unity ($\delta \geq 1$). In addition, a non-convex VRS model

⁵ The VRS technology, defined before, in fact satisfies NDRS and NIRS in different regions (see Färe, Grosskopf and Lovell (1994)).

⁶ Some of these have also been outlined in Bogetoft (1996: 464), but without any details.

can be defined as the intersection of these non-convex NIRS and NDRS technologies ($GR^{FDH-VRS} = GR^{FDH-NIRS} \cap GR^{FDH-NDRS}$). Obviously, $GR^{FDH-VRS} \subseteq GR^{FDH-NIRS}$ and $GR^{FDH-VRS} \subseteq GR^{FDH-NDRS}$. Furthermore, $GR^{FDH} \subseteq GR^{FDH-VRS}$. $GR^{FDH-VRS}$ is comparable to a convex VRS model, except for the convexity assumption.

While the DEA models are well known, these non-convex models have never been implemented empirically. To provide some intuition, these four non-convex technologies, partly proposed in Bogetoft (1996), together with the traditional FDH are illustrated in the input output space on Figure 1. While $GR^{FDH-CRS}$ allows for an unrestricted scaling of activities, $2GR^{FDH-NIRS}$ and $GR^{FDH-NDRS}$ only permit a lower respectively an upper proportionality of observed activities by means of the scaling parameter (δ). $GR^{FDH-NIRS}$ and $GR^{FDH-NDRS}$ are related to technologies proposed in Petersen (1990), but the latter author maintains the convexity assumption in input space and in output space.⁷ $GR^{FDH-VRS}$, being the intersection of $GR^{FDH-NIRS}$ and $GR^{FDH-NDRS}$, clearly contains the traditional FDH model.

Figure 1a: Non-convex CRS model

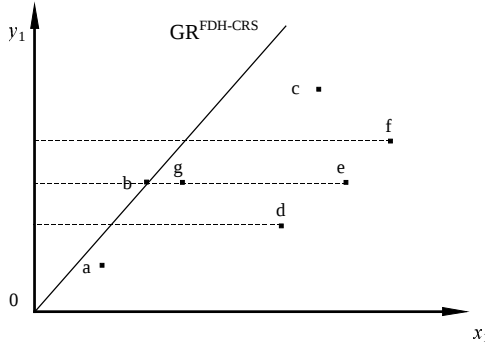


Figure 1b: Non-convex FDH and VRS model

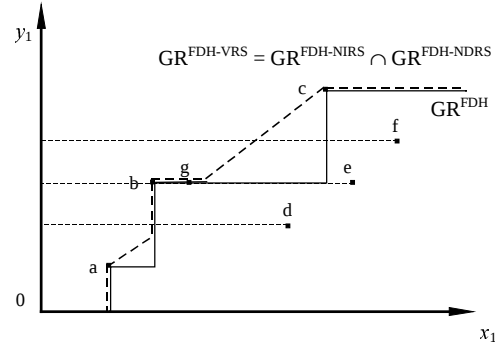
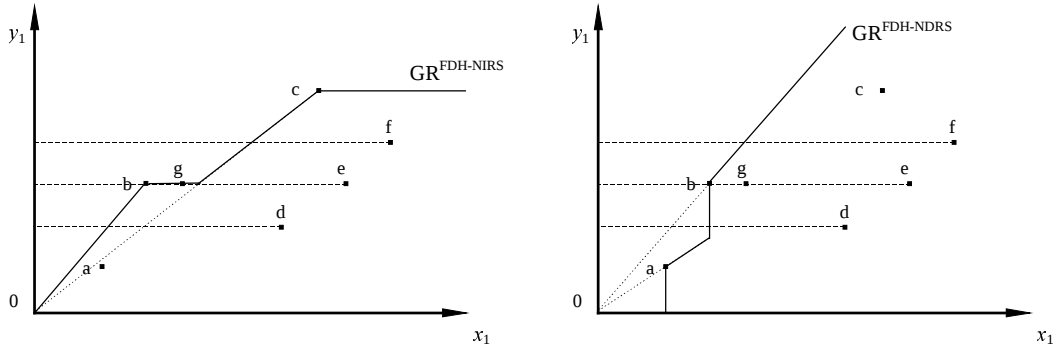


Figure 1c: Non-convex NIRS model

Figure 1d: Non-convex NDRS model

⁷ In the single input single output case, however, $GR^{FDH-NIRS}$ and $GR^{FDH-NDRS}$ coincide with the Petersen (1990) technologies. Obviously, also $GR^{FDH-CRS}$ and $GR^{DEA-CRS}$ coincide.



A comparison between a convex DEA model and a non-convex CRS model is depicted in Figure 2. This is a three-dimensional space with two inputs and a single output. Observations d_1 , d_2 , and d_3 span the frontiers. There is an input section at the output level y' where the rescaled observations d_1 and d_2 span the DEA isoquant, while the rescaled observations d_1 , d_2 , and d_3 determine the FDH level set. The darkly shaded surface represents the DEA hypersurface spanned by the origin, d_1 and d_2 . The lighter shaded surfaces are the FDH hypersurfaces spanned by the origin, d_1 , d_2 , and d_3 . Note that point d_3 is efficient relative to the non-convex model, but that is inefficient in the inputs with respect to the convex model (by an amount $y'd_3/y'd_3$).

Figure 2: Convex and Non-convex CRS models in Three Dimensions

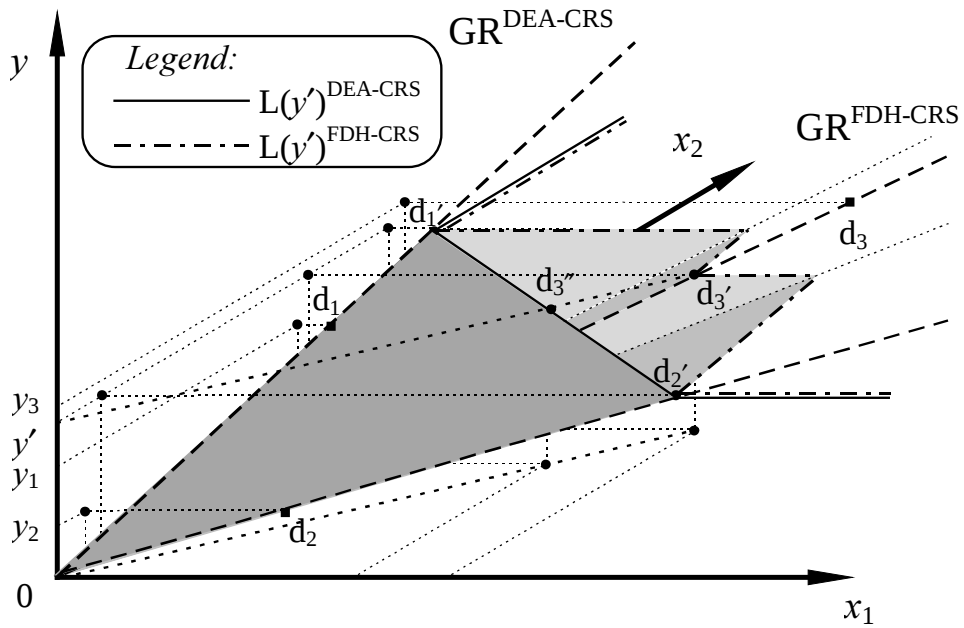


Figure 3 illustrates both static efficiency concepts on convex technologies. It shows two production possibilities sets and their boundaries, each showing the input combinations able to produce the output vector. The set $GR^{DEA-CRS}$ is characterised by CRS, in line with the competitive equilibrium ideal. The set $GR^{DEA-VRS}$ assumes VRS. Obviously, these technologies are nested: $GR^{DEA-VRS} \subseteq GR^{DEA-CRS}$.

Both static efficiency concepts and their radial way of measurement can be depicted by commenting on the positions of observation f (in the interior of the set $GR^{DEA-VRS}$) and its projection points relative to the two technologies.

First, observation f is technical inefficient (TE) because it uses more input to produce exactly the same output as, e.g., its projection point f_1 on the boundary of the set $GR^{DEA-VRS}$. Technical inefficiency in the Farrell (1957) sense is represented by the ratio of distances Of_1/Of measured relative to this strongly disposable technology ($GR^{DEA-VRS}$). Second, scale inefficiency (SCE) is illustrated using unit f_2 , which is another radial projection point of observation f . This point on the boundary of the set $GR^{DEA-VRS}$ is scale inefficient because it needs more inputs to deliver the same output level as, e.g., projection point f_2 on the boundary of the set $GR^{DEA-CRS}$. The latter point operates on a CRS technology. Scale inefficiency is captured by the ratio Of_2/Of_1 , i.e., by comparing strongly disposable short run ($GR^{DEA-VRS}$) and long run ($GR^{DEA-CRS}$) technologies.

On Figure 4 a similar decomposition is illustrated for the non-convex technologies. Also these production models are nested: $GR^{FDH} \subseteq GR^{FDH-CRS}$.

Figure 3: OTE: Taxonomy Illustrated on Convex Technologies

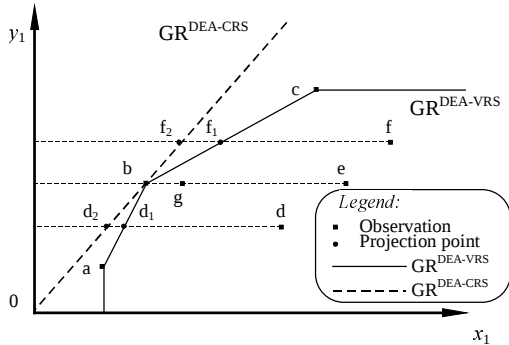
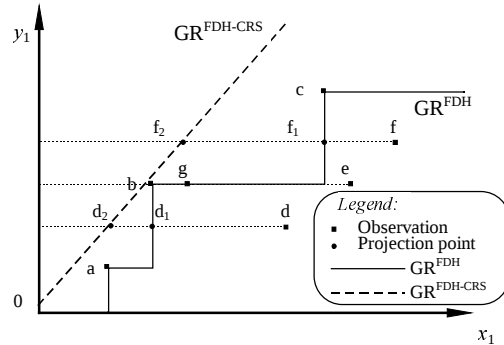


Figure 4: OTE: Taxonomy Illustrated on Non-Convex Technologies



2.3. Relation between Convex and Non-Convex Decompositions

The relation between convex and non-convex decompositions requires some clarification. As a preliminary remark, observe that all non-convex technologies are nested in their convex counterparts: $GR^{FDH} \subseteq GR^{DEA-VRS}$, and $GR^{FDH-CRS} \subseteq GR^{DEA-CRS}$.

The non-convex OTE component cannot be larger than its convex counterpart ($OTE^{Convex} \leq OTE^{Non-Convex}$). The difference between both can be completely attributed to the convexity assumption. Recall that both technologies impose CRS and strong input and output disposability. Therefore, it is possible to add a convexity-related efficiency component (CRE) to the non-convex decomposition to make it coincide with the convex decomposition: $OTE^{Convex} = CRE \cdot OTE^{Non-Convex}$. This CRE component is computed as the ratio between convex and non-convex OTE: $CRE = OTE^{Convex} / OTE^{Non-Convex}$. Since $OTE^{Convex} \leq OTE^{Non-Convex}$, clearly $0 < CRE \leq 1$. Both OTE components coincide (hence $CRE=1$) in two cases: (i) when an inefficient observation is projected on the hyperplanes parallel to the axes (with positive slack and surplus variables); (ii) when an efficient observation is part of these hyperplanes parallel to the axes or when it spans simultaneously FDH and DEA frontiers. Existing comparisons in the literature between FDH and VRS convex production models do not properly capture the convexity effect, because they mix up differences between technologies with regard to convexity and returns to scale assumptions.

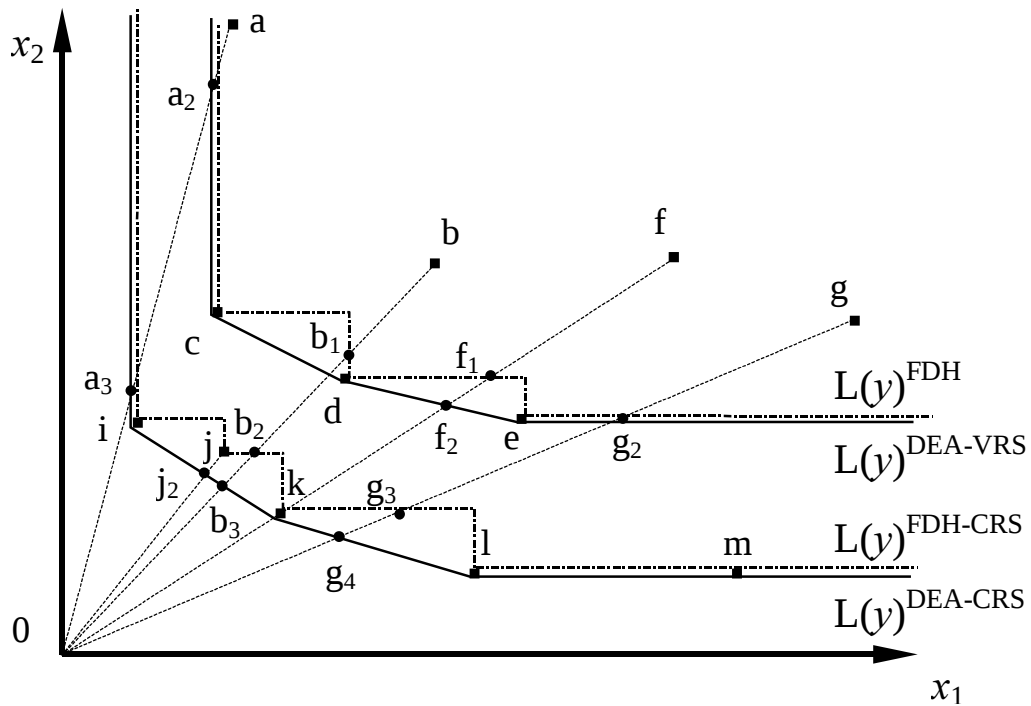
Next, it is clear that the TE component of OTE is larger when measured relative to non-convex technologies ($TE^{Convex} \leq TE^{Non-Convex}$), since the underlying technologies are nested. Finally, there is no a priori ordering between SCE components of both decompositions. While the underlying efficiency measures can be ordered (since again the non-convex technologies are subsets of the convex technologies), it is impossible to order the ratios between these efficiency measures evaluated relative to respectively non-convex and convex technologies ($SCE^{Convex} \leq SCE^{Non-Convex}$, or $SCE^{Convex} \geq SCE^{Non-Convex}$).

The major differences between both decompositions can be easily illustrated on Figure 5. First, non-convex OTE cannot be smaller since the technology $L(y)^{FDH-CRS}$ is nested in $L(y)^{DEA-CRS}$. For instance, the ratio Ob_3/O_b evaluated on $L(y)^{DEA-CRS}$ is smaller than the ratio Ob_2/O_b relative to $L(y)^{FDH-CRS}$. The same applies for the TE component. Second, the relative amount of SCE may either decrease or increase, depending on the sample. This can be seen from contrasting points f and g. While for observation f SCE increases under convexity (i.e., the ratio $Ok/O_{f_2} \geq Ok/O_{f_1}$), the reverse is true for g (convex

ratio $0g_4/0g_2 \leq \text{non-convex ratio } 0g_3/0g_2$). Third, CRE is unity only for three types of observations. For an inefficient observation (for instance, point a) that is projected on the hyperplanes parallel to the axes. For an efficient observation that is either part of these hyperplanes parallel to the axes (e.g., observation m), or that is simultaneously part of both FDH and DEA frontiers (e.g., observation k). In all other cases CRE is smaller than unity.

Which of the differences between convex and non-convex decompositions proves to be most important is an empirical matter. It may be possible to formulate some expectations, however. For instance, detailed studies of the convex decomposition reveal that especially the CRS technology (and hence the OTE component) is susceptible to outliers, as it is spanned by relatively few observations (see, e.g., Kerstens (1998)). The empirical section is a first attempt to systematically clarify the relative importance of these differences.

Figure 5: CRE and the Link between Convex and Non-Convex Decompositions



3. COMPUTATIONAL ASPECTS OF CONVEX AND NON-CONVEX DECOMPOSITIONS

3.1. Computing Technical and Scale Efficiency on Convex and Non-Convex Technologies

The efficiency computations on the convex DEA models require solving linear programming (LP) problems for each observation (x^o, y^o) being evaluated (see Färe, Grosskopf and Lovell (1994)). We repeat them for convenience and for comparative purposes. Radial input efficiency relative to $GR^{DEA-CRS}$, $GR^{DEA-VRS}$, $GR^{DEA-NIRS}$, $GR^{DEA-NDRS}$ requires solving for each evaluated observation (x^o, y^o) the following LP (P.1):

$$\begin{aligned}
DF_i(x, y) &= \underset{\lambda, z}{\text{Min}} \quad \lambda \\
\text{subject to} \quad & \sum_{k=1}^K y_{km} z_k \geq y_{km}^o, \quad m = 1, \dots, M, \\
& \sum_{k=1}^K x_{kn} z_k \leq \lambda x_{kn}^o, \quad n = 1, \dots, N, \\
& \lambda \geq 0, z_k \in \Gamma(s), \quad k = 1, \dots, K,
\end{aligned}$$

where (i) $\Gamma(s) = \{z_k: z_k \geq 0\}$ for $s = \text{DEA} - \text{CRS}$;

(ii) $\Gamma(s) = \{z_k: \sum_{k=1}^K z_k = 1, z_k \geq 0\}$ for $s = \text{DEA} - \text{VRS}$;

(iii) $\Gamma(s) = \{z_k: \sum_{k=1}^K z_k \leq 1, z_k \geq 0\}$ for $s = \text{DEA} - \text{NIRS}$;

(iv) $\Gamma(s) = \{z_k: \sum_{k=1}^K z_k \geq 1, z_k \geq 0\}$ for $s = \text{DEA} - \text{NDRS}$.

Proposition 1:

The input efficiency computations using the LP problem (P.1) are equivalent to the non-linear programming problem (P.2):

$$\begin{aligned}
DF_i(x, y) &= \underset{\lambda, z, \delta}{\text{Min}} \quad \lambda \\
\text{subject to} \quad & \sum_{k=1}^K y_{km} \delta z_k \geq y_{km}^o, \quad m = 1, \dots, M, \\
& \sum_{k=1}^K x_{kn} \delta z_k \leq \lambda x_{kn}^o, \quad n = 1, \dots, N, \\
& \sum_{k=1}^K z_k = 1, \\
& \delta \in \Gamma(s), \\
& \lambda \geq 0, z_k \geq 0, \quad k = 1, \dots, K,
\end{aligned}$$

where (i) $\Gamma(s) = \{\delta: 0 < \delta\}$ for $s = \text{DEA} - \text{CRS}$;

(ii) $\Gamma(s) = \{\delta: \delta = 1\}$ for $s = \text{DEA} - \text{VRS}$;

(iii) $\Gamma(s) = \{\delta: 0 < \delta \leq 1\}$ for $s = \text{DEA} - \text{NIRS}$;

(iv) $\Gamma(s) = \{\delta: \delta \geq 1\}$ for $s = \text{DEA} - \text{NDRS}$.

These new formulations deviate from standard representations available in the literature to highlight the similarity with the FDH-based non-convex production models (see below). It is inspired by the formulation of the DEA-CRS model in Banker, Charnes and Cooper (1984: 1082) and Chang and Guh (1991: 220), and it has also been extensively developed in Leleu (1997). In these alternative formulations the sum constraint on z represents the convexity hypothesis; δ is a scaling factor reflecting the specific returns to scale assumption; and the inequality signs are due to the strong output and input disposability axioms. The equivalence between (P.1) and (P.2) is proven in the Appendix. The major advantage of the new formulation is pedagogical: the convexity and returns to scale assumptions are clearly separated.

For the non-convex technologies, by contrast, mixed integer programming problems are required. This section shows that it is possible to make use of algorithms of implicit enumeration based upon vector dominance reasoning. Input efficiency is computed relative to $GR^{FDH-CRS}$, GR^{FDH} , $GR^{FDH-NIRS}$, and $GR^{FDH-NDRS}$ by solving for each observation (x^o, y^o) the following binary mixed integer, non-linear programming problem (P.3):

$$\begin{aligned}
DF_i(x, y) &= \underset{\lambda, z, \delta}{\text{Min}} \quad \lambda \\
\text{subject to} \quad & \sum_{k=1}^K y_{km} \delta z_k \geq y_{km}^o, \quad m = 1, \dots, M, \\
& \sum_{k=1}^K x_{kn} \delta z_k \leq \lambda x_{kn}^o, \quad n = 1, \dots, N, \\
& \sum_{k=1}^K z_k = 1, \\
& z_k \in \{0, 1\}, \\
& \delta \in \Gamma(s), \\
& \lambda \geq 0, z_k \geq 0, \quad k = 1, \dots, K,
\end{aligned}$$

where (i) $\Gamma(s) = \{\delta: 0 < \delta\}$ for $s = \text{FDH} - \text{CRS}$;
(ii) $\Gamma(s) = \{\delta: \delta = 1\}$ for $s = \text{FDH}$;
(iii) $\Gamma(s) = \{\delta: 0 < \delta \leq 1\}$ for $s = \text{FDH} - \text{NIRS}$;
(iv) $\Gamma(s) = \{\delta: 1 \leq \delta\}$ for $s = \text{FDH} - \text{NDRS}$.

In this formulation the sum constraint on z together with the integer constraint on z represent the non-convexity assumption. This problem is non-linear and furthermore contains binary integer restrictions. Given the binary nature of the integers, it can be solved using an implicit enumeration algorithm based upon a simple vector comparison

procedure (see Garfinkel and Nemhauser (1972), § 10.1). This algorithm developed to solve for radial efficiency on these non-convex technologies is similar to the one already described for FDH in the literature (see Tulkens (1993) and below).

The algorithm derived for the FDH-based family of technologies can be developed along the following lines. In the first step a modified index set of better observations is defined allowing for a rescaling of the observations in the sample according to the specific returns to scale assumption postulated. The vector dominance comparison thus accounts for the possibility that observations may be rescaled within certain parameter bounds. For convenience, this will be coined “scaled vector dominance”. The “scaled better set” of the observation (x^o, y^o) is therefore conditional on the returns to scale assumptions:

$$B(x^o, y^o | s) = \{(x_k, y_k) : \delta x_k \leq x^o, \delta y_k \geq y^o, \delta \in \Gamma(s)\},$$

where

- (i) $\Gamma(s) = \{\delta : 0 < \delta\}$ for $s = \text{FDH - CRS}$;
- (ii) $\Gamma(s) = \{\delta : \delta = 1\}$ for $s = \text{FDH}$;
- (iii) $\Gamma(s) = \{\delta : 0 < \delta \leq 1\}$ for $s = \text{FDH - NIRS}$;
- (iv) $\Gamma(s) = \{\delta : \delta \geq 1\}$ for $s = \text{FDH - NDRS}$.

In the second step the input efficiency measure is computed given some knowledge about the scaling parameter. It is not necessary to test for all values of the scaling parameter (δ). Instead, for each observation being evaluated one only needs to find critical values for this scaling parameter depending on the selected orientation of measurement and the assumption made regarding returns to scale.

The critical value of the scaling parameter for the input oriented technical efficiency measure (δ^I) is defined as follows:

$$\delta^I = \text{Min}_{m=1, \dots, M} \left(\frac{y_{km}}{y_{km}^o} \right).$$

If this critical value respects the lower and upper bounds on the scaling parameter related to the postulated returns to scale assumption, then the rescaled observation is part of the scaled better set; otherwise it is not an element of the scaled better set:

$$\begin{aligned} \text{If } \delta^I \in \Gamma(s), \text{ then } (x_k, y_k) &\in B(x^o, y^o | s); \\ \text{otherwise, then } (x_k, y_k) &\notin B(x^o, y^o | s). \end{aligned}$$

Proposition 2:

The input oriented technical efficiency measure on the FDH-based family of technologies (mixed integer non-linear programming problem (P.3)) can be computed as follows:

$$DF_i(x^o, y^o) = \min_{(x_k, y_k) \in B(x^o, y^o | s)} \max_{n=1, \dots, N} \left(\frac{\delta^n x_{kn}}{x_{kn}^o} \right).$$

The validity of this vector dominance procedure for the input radial efficiency measure is proven in the Appendix. Proofs for other orientations are similar and left to the reader.

Since FDH involves no scaling, the scaling parameter (δ) is fixed at 1 in (P.3) and in all formulas. This yields the traditional FDH binary mixed integer LP problem. As shown in Tulkens (1993) (see also De Borger, Ferrier and Kerstens (1998), Fried, Lovell and Vanden Eeckaut (1995) and Lovell (1995)), this linear problem can be solved in two steps using an algorithm of implicit enumeration (see Garfinkel and Nemhauser (1972), § 4.5). In the first step vector dominance procedures determine for each observation its set of dominating observations ($B(x^o, y^o)$ is independent of any scaling). In the second step the efficiency measure is computed by directly applying its definition. This clearly fits into the algorithm derived above for the FDH-based family of technologies.

Finally, the computation of input efficiency relative to the VRS non-convex model requires some discussion. Since $GR^{FDH-VRS} = GR^{FDH-NIRS} \cap GR^{FDH-NDRS}$, input efficiency can be trivially computed relative to this technology as follows:⁸

$$DF_i(x, y | VRS) = \max \{DF_i(x, y | NIRS), DF_i(x, y | NDRS)\}.$$

Alternatively, a programming problem similar to (P.3) can be solved whereby the intersection of technologies $GR^{FDH-NIRS}$ and $GR^{FDH-NDRS}$ is simultaneously imposed.⁹

⁸ The convex VRS model can similarly be written: $GR^{DEA-VRS} = GR^{DEA-NIRS} \cap GR^{DEA-NDRS}$. Also in this convex case, input-oriented efficiency can be computed as: $DF_i(x, y | VRS) = \max \{DF_i(x, y | NIRS), DF_i(x, y | NDRS)\}$. In a similar vein, both convex and non-convex CRS models can alternatively be defined in terms of the union of NIRS and NDRS technologies. Then, input-oriented efficiency is: $DF_i(x, y | CRS) = \min \{DF_i(x, y | NIRS), DF_i(x, y | NDRS)\}$. These indirect ways of estimating VRS and CRS technologies went so far unnoticed in the literature.

⁹ To be precise, the $GR^{FDH-VRS}$ technology can be written explicitly as a conjunction of $GR^{FDH-NIRS}$ and $GR^{FDH-NDRS}$ as follows:

$$GR^{FDH-VRS} = \{(x, y): M' w_1 \geq y, N' w_1 \leq x, I'_k z_1 = 1, z_1 \in \{0,1\}, w_1 = \delta_1 z_1, 0 \leq \delta_1 \leq 1, \\ M' w_2 \geq y, N' w_2 \leq x, I'_k z_2 = 1, z_2 \in \{0,1\}, w_2 = \delta_2 z_2, \delta_2 \geq 1\}.$$

See, e.g., Ruys (1974) for definitions of operations on technologies.

3.2. Qualitative Returns to Scale Information

Important additional characteristics of production that can be analysed are the determinants of returns to scale. For both observations on the frontier and in the interior of the production possibility set it is possible to obtain qualitative information on scale economies locally, i.e., for its bounding hyperplane. For inefficient observations, this characterisation of returns to scale depends obviously on the chosen measurement orientation. The three traditional, equivalent methods to determine returns to scale in DEA technologies (see Banker, Chang and Cooper (1996)) do, however, not apply for the non-convex technologies defined above.¹⁰

A new procedure based on goodness-of-fit has been developed to determine the returns to scale assumption which best fits the data (see Kerstens and Vanden Eeckaut (1998)). Essentially, input efficiency is computed on three different technologies imposing respectively CRS, NIRS, and NDRS. Since these three technologies are not nested, it is possible to infer for any single observation whether it satisfies constant (CRS), increasing (IRS), or decreasing returns (DRS) to scale by simply identifying the technology yielding the maximal input efficiency score. The latter reflects for the given observation the returns to scale hypothesis that fits best. Using an input-oriented measurement and conditional on the optimal projection point, technology is characterised locally by:

$$\begin{aligned} \text{CRS} &\Leftrightarrow \text{DF}_i(x,y \mid \text{CRS}) = \max\{\text{DF}_i(x,y \mid \text{CRS}), \text{DF}_i(x,y \mid \text{NIRS}), \text{DF}_i(x,y \mid \text{NDRS})\}; \\ \text{IRS} &\Leftrightarrow \text{DF}_i(x,y \mid \text{NDRS}) = \max\{\text{DF}_i(x,y \mid \text{CRS}), \text{DF}_i(x,y \mid \text{NIRS}), \text{DF}_i(x,y \mid \text{NDRS})\}; \text{ or} \\ \text{DRS} &\Leftrightarrow \text{DF}_i(x,y \mid \text{NIRS}) = \max\{\text{DF}_i(x,y \mid \text{CRS}), \text{DF}_i(x,y \mid \text{NIRS}), \text{DF}_i(x,y \mid \text{NDRS})\}.^{11} \end{aligned}$$

This new procedure is more general than previous methods.

Remark finally that the use of NIRS and NDRS technologies is not limited to obtaining information on local scale economies, but that they also serve a purpose in determining the effect of the convexity assumption. Indeed, any difference between efficiency measures evaluated relative to convex and non-convex technologies with otherwise identical returns to scale assumptions can be completely attributed to the convexity assumption. Thus, in addition to CRE computed relative to two CRS technologies, it is also possible to compare two NIRS technologies: one convex, another

¹⁰ Note that the optimal value of the scaling parameter δ does not convey information regarding returns to scale. It is only instrumental in the construction of the frontier.

¹¹ Note that all three input efficiency measures coincide for observations subject to CRS. Remark also that the same formulas apply for any efficiency measure smaller than unity. Obviously, for radial output efficiency measures defined to be larger than unity a min operator should replace the max operator.

non-convex. The same applies to NDRS and also to VRS technologies. This provides a perfect base to compare the precise impact of the convexity assumption, conditional on a specific returns to scale assumption.

4. EMPIRICAL ILLUSTRATION: CONTRASTING CONVEX AND NON-CONVEX OTE DECOMPOSITIONS

To exploit the possibility of contrasting decompositions of OTE based on convex technologies with similar decompositions based on non-convex technologies, this section partially duplicates earlier research. First, the article of Thanassoulis, Dyson and Foster (1987) is selected for the empirical analysis, since this publication also contained the data set.¹² Next, two additional samples with a somewhat larger size are investigated to corroborate empirical regularities: one concerns French urban transit companies (Kerstens (1996)); the other one deals with Belgian municipalities located in the Walloon region (Vanden Eeckaut, Tulkens and Jamar (1993)).

4.1. Analysing UK Rates Departments

Thanassoulis, Dyson and Foster (1987) investigate the relative performance of 62 UK rates departments observed in the financial year 1982/1983. More in particular, the rate-collection function of London Boroughs and Metropolitan District Councils is evaluated. The data consist of four outputs: the number of non-council rates accounts administered, rate rebate applications granted, the number of summonses issued and distress warrants obtained, and the net present value of non-council rates collected. On the input side no details are available. Real resources used are captured by total outlays. Published findings focus on radial input efficiency relative to a DEA model with CRS. The reader is referred to the original sources for all details concerning the samples and their interpretation.

One consequence of using costs to compute SCE is that the results normally differ from the ones obtained when proper input quantities would have been available. As shown in Färe and Grosskopf (1985) dual and primal approaches to measuring SCE only coincide when (i) all organisations face the same input price vector, and (ii) allocative efficiency is identical when measured relative to CRS and VRS technologies.

¹² In fact, the data itself were published as Appendix 2 in Dyson and Thanassoulis (1988).

Our analysis provides results for both input and output orientations. In different tables the convex decomposition results of OTE are presented along with the new non-convex decomposition.

Summary statistics for both convex and non-convex decomposition results are presented in Table 1. To respect the multiplicative nature of the decomposition geometric averages are used. Frequencies for both decompositions are depicted in Figures 6 (a and b) and 7 (a and b) for the output respectively the input orientation. These distributions are markedly skewed to the right. The sources of inefficiency differ clearly between both methods. This can be systematically argued as follows.

First, in the convex case TE is on average smaller than SCE in both orientations. For the non-convex decomposition, by contrast, TE is on average larger than SCE in both cases. Thus both decompositions disagree on the major source of poor performance. Second, OTE is identical under both orientations on both convex and non-convex production models, because under CRS input and output oriented radial efficiency measures coincide (see Färe and Lovell (1978) or Deprins and Simar (1983) for proofs).

Table 1: Non-Convex and Convex Decomposition Results under Output and Input Orientations

	Output Orientation						
	Non-Convex Decomposition				Convex Decomposition		
	TE	SCE	OTE	CRE	TE	SCE	OTE
Average*	0.9041	0.8588	0.7764	0.8672	0.7197	0.9355	0.6733
Stand.Dev.	0.1367	0.1223	0.1764	0.1192	0.1921	0.0788	0.1870
Minimum	0.4796	0.5792	0.3975	0.5744	0.3625	0.5288	0.3280
Maximum	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
# Effic. Obs.	39	15	15	14	13	7	7
	Input Orientation						
	Non-Convex Decomposition				Convex Decomposition		
	TE	SCE	OTE	CRE	TE	SCE	OTE
Average*	0.9329	0.8323	0.7764	0.8672	0.7535	0.8935	0.6733
Stand.Dev.	0.1123	0.1333	0.1764	0.1192	0.1793	0.1031	0.1870
Minimum	0.4602	0.5349	0.3975	0.5744	0.4057	0.5288	0.3280
Maximum	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
# Effic. Obs.	39	15	15	14	13	7	7

* Geometric Average

Third, to make these decompositions comparable, a CRE component has been added. CRE links OTE under both decompositions: i.e., the identity $0.7764 \times 0.8672 = 0.6733$ holds using geometric averages. Because OTE is on average quite different under

both decompositions, one infers that convexity plays a non-negligible role in explaining efficiency in this sample. In particular, CRE amounts to 13% on average. Given an average convex OTE score of 0.67, this means that about half of the measured inefficiency (13% out of 33%) is solely attributed to the convexity assumption. Remark that CRE is identical under both orientations, since it is a ratio of two CRS efficiency measures.

Figure 6a: Frequencies of Non-Convex OTE Decomposition: Output Orientation

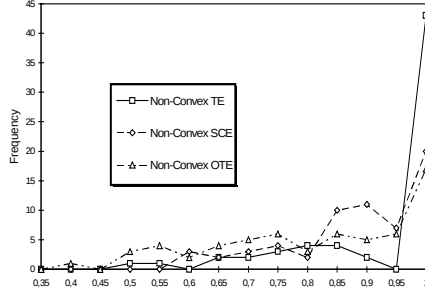


Figure 6b: Frequencies of Convex OTE Decomposition: Output Orientation

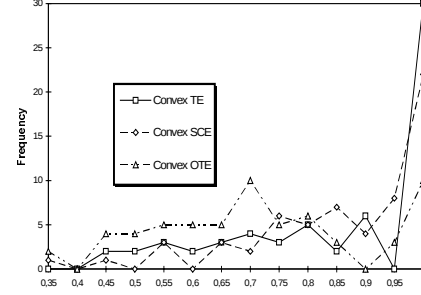


Figure 7a: Frequencies of Non-Convex OTE Decomposition: Input Orientation

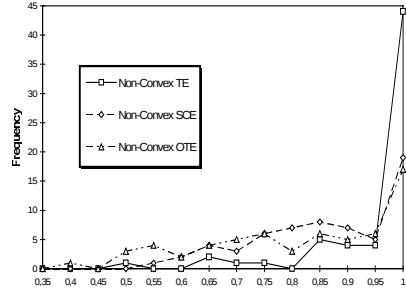
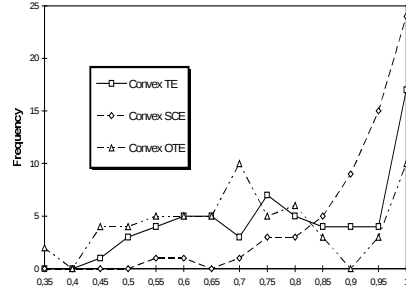


Figure 7b: Frequencies of Convex OTE Decomposition: Input Orientation



Simple nonparametric test statistics indicate that the distributions of both convex and non-convex decompositions differ in a statistically significant way. Furthermore, this result holds for both measurement orientations. In particular, using a Friedman two-way analysis of variance based on ranks reveals that convex and non-convex TE, OTE, and SCE components have no common distribution. Furthermore, a Wilcoxon signed rank test confirms these negative results per convex and non-convex decomposition component. Remark that a priori the SCE components of both decompositions cannot be ordered.

The bottom row of Table 1 indicates the number of efficient observations per component. Clearly, the number of technically efficient units triples in the non-convex compared to the convex case. In relative terms, about 63% of observations are technically efficient on the FDH. More importantly, the numbers for the scale and overall technical efficiency components are also doubles of one another for both types of decompositions. It

amounts to only 11% of the sample for the convex decomposition. Thus, the CRS FDH model is always spanned by more or the same number of observations as the convex CRS frontier. This fact may, in general, render the CRS FDH frontier less susceptible to outliers. Finally, for 14 observations the convexity assumption makes no difference for their overall technical efficiency evaluation: 7 are inefficient observations being projected onto hyperplanes parallel to the axes; 7 observations are efficient under both the convex and non-convex technologies.

As mentioned before, the impact of convexity can also be assessed in three additional ways: by comparing convex respectively non-convex NIRS, NDRS and VRS technologies. This provides additional information on the precise impact of the convexity assumption. This is the first empirical application that properly accounts for differences between technologies both with regard to convexity and returns to scale assumptions. By contrast, comparisons in the literature of FDH and VRS convex production models have not accurately documented the convexity effect.

Results are listed in Table 2, where for convenience CRE results reported in Table 1 are duplicated. On average, between 10% and 16% of measured inefficiency can be solely attributed to the convexity assumption. Clearly, the impact of convexity is conditional on specific returns to scale assumptions. In these cases, the effect of the convexity assumption obviously need not be identical under both orientations. It turns out, however, that differences between orientations are negligible.

Table 2: Impact of the Convexity Assumption: A Detailed Look

	Output Orientation				Input Orientation			
	NDRS	CRS	NIRS	VRS	NDRS	CRS	NIRS	VRS
Average*	0.8386	0.8672	0.8764	0.8476	0.8765	0.8672	0.8649	0.8742
Stand.Dev.	0.1219	0.1192	0.1137	0.1149	0.1076	0.1192	0.1212	0.1084
Minimum	0.5772	0.5744	0.5684	0.5711	0.6044	0.5744	0.5684	0.6044
Maximum	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
# Effic. Obs.	13	14	14	13	12	14	14	13

* Geometric Average

To study the effects on the ranking of individual observations product-moment correlations between the components of both decompositions have been computed. These correlations are reported in Table 3. For the OTE and TE components the rankings seems to be pretty similar. But they are quite diverging for the SCE components. There is little difference between measurement orientations.

Table 3: Product-Moment Correlations between Convex and Non-Convex Decomposition Components

Orientation	TE	SCE	OTE
Output	0.7172	0.3988	0.8285
Input	0.6230	0.5024	0.8285

Next, it is important to verify whether there are differences in the determination of returns to scale for individual observations. Table 4 summarises the results per decomposition, per orientation of measurement, and distinguishes between efficient and projected inefficient observations (according to the TE criterion). Table 5 reports results across decompositions and across measurement orientation (the latter type of results was first documented in Fukuyama (1996)).

Table 4: Returns to Scale Results per Technology

Technology	Orientation	Type of Observations	IRS	CRS	DRS
FDH	Output	Efficient Observations	19	15	5
		Projected Inefficient Observations	12	0	11
	Input	Efficient Observations	19	15	5
		Projected Inefficient Observations	16	0	7
DEA	Output	Efficient Observations	5	7	1
		Projected Inefficient Observations	30	0	19
	Input	Efficient Observations	5	7	1
		Projected Inefficient Observations	46	0	3

Starting with Table 4, DEA clearly projects the overwhelming majority of observations on an IRS part of technology, while the distribution of observations over the 3 categories is more even in the FDH case. Nevertheless, it is clear that the evaluated organisations mainly operate on a too small scale. Furthermore, the orientation of measurement has almost no impact on the classification under FDH. On convex technologies, by contrast, 16 observations are classified completely differently (IRS versus DRS) conditional on measurement orientation. This is probably related to the greater flexibility of returns to scale specifications under non-convex technologies. Finally, relatively speaking, there are quite some observations characterised by CRS on the non-convex technologies.

Comparing results across decompositions and measurement orientations in Table 5, the first three columns indicate the degree of agreement while the next three columns reveal any disagreement between classifications. Regarding DEA and FDH results, there is quite some disagreement. For instance, there are changes from IRS to CRS

or even DRS. Especially the results in the IRS-DRS column, where conclusions are exactly opposite, show that non-convex and convex decompositions can substantially disagree regarding returns to scale. Furthermore, the impact of measurement orientation is more drastic under the convex compared to the non-convex decomposition. While under the former model 16 observations change from IRS to DRS (or the reverse) when changing orientation, only 4 observations change under the non-convex models. Thus, different qualitative conclusions regarding returns to scale are possible for both efficient and inefficient observations.

Table 5: Returns to Scale Results between Technologies and Measurement Orientations

	IRS-IRS	CRS-CRS	DRS-DRS	IRS-DRS	IRS-CRS	DRS-CRS	Total
Convex versus Non-Convex Technologies							
Output	26	7	12	9	5	3	62
Input	34	7	3	10	8	0	62
Output versus Input Orientation of Measurement							
Non-Convex	31	15	12	4	0	0	62
Convex	35	7	4	16	0	0	62

4.2. Additional Evidence: French Urban Transit & Belgian Municipalities

The OTE in the samples of French urban transit companies, analysed in Kerstens (1996), and among Walloon municipalities, investigated in Vanden Eeckaut, Tulkens and Jamar (1993), are decomposed in this subsection. We first briefly summarise the production model specifications in both articles.

The French urban transit sample contains information on 114 single mode (buses) urban transport companies in 1990. The output, which is a pure supply indicator, is the number of vehicle kilometres, i.e., a classical units times distance per unit time concept. There are three traditional transport inputs: number of vehicles; number of employees; and fuel consumption. The data on the Walloon municipalities cover 235 observations for the year 1985 and mainly 1986. Total expenditures are used as inputs. Outputs include: number of beneficiaries of minimal subsistence grants; number of students enrolled in local primary schools; total population; fraction of population above 65; length of the public road system; and crime rate.

This subsection documents results for both input and output orientations for the French urban transit sample, but only input oriented findings for the Walloon municipalities. Limitations of space force us to report only on a selection of the previously discussed tables, with results being duplicated for both samples.

Starting with Table 6, it is apparent that both decompositions yield different conclusions with respect to the major causes of deviations from OTE: it is SCE for one; and TE according to the other. This is valid under both orientations for the urban transit sample. Furthermore, CRE amounts to about 9% to 12% in both samples, in line with the above reported findings. Finally, under the FDH decomposition the number of efficient observations tends at least to double for each efficiency source compared to the convex methodology.

Table 6: Non-Convex and Convex Decomposition Results under Output and Input Orientations

French Urban Transit Companies							
Output Orientation	Non-Convex Decomposition				Convex Decomposition		
	TE	SCE	OTE	CRE	TE	SCE	OTE
Average*	0.9755	0.7543	0.7358	0.9143	0.7421	0.9064	0.6727
Stand.Dev.	0.0582	0.1179	0.1184	0.0476	0.1411	0.1197	0.1172
Minimum	0.6527	0.5056	0.5056	0.8028	0.5192	0.4545	0.4545
Maximum	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
# Effic. Obs.	83	8	7	4	11	6	4
Input Orientation	Non-Convex Decomposition				Convex Decomposition		
	TE	SCE	OTE	CRE	TE	SCE	OTE
Average*	0.9723	0.7568	0.7358	0.9143	0.7372	0.9124	0.6727
Stand.Dev.	0.0609	0.1173	0.1184	0.0476	0.1385	0.1147	0.1172
Minimum	0.6667	0.5056	0.5056	0.8028	0.5114	0.4545	0.4545
Maximum	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
# Effic. Obs.	88	7	7	4	11	6	4
Belgian Municipalities							
Input Orientation	Non-Convex Decomposition				Convex Decomposition		
	TE	SCE	OTE	CRE	TE	SCE	OTE
Average*	0.9704	0.8370	0.8122	0.8755	0.7880	0.9025	0.7111
Stand.Dev.	0.0697	0.1370	0.1528	0.0719	0.1524	0.1099	0.1563
Minimum	0.5410	0.4859	0.4514	0.7111	0.4073	0.4832	0.4023
Maximum	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
# Effic. Obs.	189	64	63	17	46	19	17

* Geometric Average

The correlation analysis in Table 7 corroborates our previous results. While rankings are very similar for OTE, they are much lower for both of its components. As before, SCE yields the least degree of concordance.

Table 7: Product-Moment Correlations between Convex and Non-Convex Decomposition Components

Orientation	TE	SCE	OTE
French Urban Transit Companies			
Output	0.3797	0.3425	0.9484
Input	0.3866	0.3343	0.9484
Belgian Municipalities			
Input	0.5801	0.5699	0.9139

Turning to qualitative scale information in Tables 8 and 9, the majority of observations in the French transport sample is subject to DRS according to the non-convex decomposition. In the DEA case, this conclusion depends on measurement orientation. For the municipalities, both models agree that overall diseconomies of scale prevail. Furthermore, on the non-convex technologies quite some observations are subject to CRS,

especially in the municipalities sample. Taking both samples together, even 2 inefficient observations are projected onto the CRS part of the non-convex technology, although this depends on measurement orientation in the urban transit case. It is well-known that the latter phenomenon is extremely rare under DEA. In our samples none of the observations falls into this category.

Table8: Returns to Scale Results per Technology

Technology	Orientation	Type of Observations	IRS	CRS	DRS
French Urban Transit Companies					
FDH	Output	Efficient Observations	22	7	54
		Projected Inefficient Observations	11	1	19
	Input	Efficient Observations	25	7	56
		Projected Inefficient Observations	11	0	15
DEA	Output	Efficient Observations	4	4	3
		Projected Inefficient Observations	54	0	49
	Input	Efficient Observations	4	4	3
		Projected Inefficient Observations	61	0	42
Belgian Municipalities					
FDH	Input	Efficient Observations	10	63	116
		Projected Inefficient Observations	15	1	30
DEA	Input	Efficient Observations	4	17	25
		Projected Inefficient Observations	41	0	148

Table 9 confirms our previous conclusion about the possibility of diverging qualitative conclusions regarding returns to scale. Again there is quite some disagreement between non-convex and convex technologies, as witnessed in columns four to six. For instance, radical changes from IRS to DRS are, relatively speaking, a serious problem for the urban transit sample. Once more, the measurement orientation has a more drastic effect under the convex compared to the non-convex decomposition.

Table 9: Returns to Scale Results between Technologies and Measurement Orientations

	IRS-IRS	CRS-CRS	DRS-DRS	IRS-DRS	IRS-CRS	DRS-CRS	Total
French Urban Transit Companies							
Convex versus Non-Convex Technologies							
Output	31	4	49	26	3	1	114
Input	33	4	41	33	2	1	114
Output versus Input Orientation of Measurement							
Non-Convex	33	7	71	2	0	1	114
Convex	58	4	45	7	0	0	114
Belgian Municipalities							
Convex versus Non-Convex Technologies							
Input	18	17	137	16	18	29	235

5. CONCLUSIONS AND DIRECTIONS FOR FUTURE RESEARCH

Starting from an existing non-convex production model (FDH) several new nonparametric deterministic production models have been explored. In particular, several returns to scale assumptions have been introduced. The resulting series of non-convex production models yields a decomposition of OTE that is similar to the existing one based on DEA models. The formal relations between convex and non-convex decompositions and their respective TE and SCE components have been fully spelled out.

The empirical part of this contribution first used a sample of UK rates departments to explore the similarities and differences between both decompositions in great detail. The major conclusions are that the identification of the major sources of inefficiency differs strongly and that conflicting conclusions with regard to the returns to scale of individual observations are possible. Furthermore, scale information is more stable under the FDH decomposition for changes in measurement orientation. All findings are corroborated by two more samples for which results are briefly reported. The fact that there are always more observations spanning the CRS FDH model may contribute to these different decomposition results.

Overall, we believe these empirical results should make us more cautious about invoking the convexity assumption in performance gauging. Lacking proper statistical tests when comparing specifications, it is important that practitioners have a precise idea of the effect of each assumption. Making use of the relation between efficiency measures and goodness-of-fit tests, our convexity related efficiency component provides exactly such a tool.

Two methodological extensions may seem worthwhile pursuing in the future. First, this new decomposition could be integrated into the productivity measurement literature. When panel data are available, it is obviously possible to integrate this decomposition into a dynamic analysis of productivity change. For instance, it would be interesting to investigate the effect of using CRS non-convex instead of convex technologies on the computation of technical change with a Malmquist productivity index. Furthermore, in analogy with the decomposition of the technical efficiency change component of the Malmquist index (see Färe, Grosskopf and Lovell (1994)), this new non-convex decomposition can be integrated into this productivity index.¹³

¹³ Malmquist indexes and other productivity measures based on the traditional FDH model have been proposed in Tulkens and Vanden Eeckaut (1995) and applied in Tulkens and Malnero (1996).

Second, the presence of slacks and the use of nonradial instead of radial efficiency measures for decomposition purposes have been neglected in the literature (see Dervaux, Kerstens and Vanden Eeckaut (1998)). It has been argued that nonradial efficiency measures provide a better alternative, especially on FDH, (De Borger, Ferrier and Kerstens (1998)), since they comply with Koopmans (1951) definition of technical efficiency.

APPENDIX

Proof of Proposition 1:

To derive (P.1) starting from (P.2), substitute $w_k = \delta z_k$ (or, $z_k = w_k/\delta$) to obtain:

$$\begin{aligned} DF_1(x, y) &= \min_{\lambda, w, \delta} \lambda \\ \text{subject to } & \sum_{k=1}^K y_{km} w_k \geq y_{km}^0, \quad m = 1, \dots, M, \\ & \sum_{k=1}^K x_{kn} w_k \leq \lambda x_{kn}^0, \quad n = 1, \dots, N, \\ & \sum_{k=1}^K \frac{w_k}{\delta} = 1, \\ & \delta \in \Gamma(s), \\ & \lambda \geq 0, w_k \geq 0, \quad k = 1, \dots, K, \end{aligned}$$

- where (i) $\Gamma(s) = \{\delta: 0 < \delta\}$ for $s = \text{DEA} - \text{CRS}$;
(ii) $\Gamma(s) = \{\delta: \delta = 1\}$ for $s = \text{DEA} - \text{VRS}$;
(iii) $\Gamma(s) = \{\delta: 0 < \delta \leq 1\}$ for $s = \text{DEA} - \text{NIRS}$;
(iv) $\Gamma(s) = \{\delta: \delta \geq 1\}$ for $s = \text{DEA} - \text{NDRS}$.

The sum constraint on the activity vector can be rewritten as:

$$\sum_{k=1}^K w_k = \delta.$$

Concentrating on this sum constraint and the restrictions on the scaling factor δ yields the following series of final constraints:

- where (i) $\Gamma(s) = \{\delta: \sum_{k=1}^K w_k = \delta \geq 0\}$ for $s = \text{DEA} - \text{CRS}$;
(ii) $\Gamma(s) = \{\delta: \sum_{k=1}^K w_k = \delta = 1\}$ for $s = \text{DEA} - \text{VRS}$;
(iii) $\Gamma(s) = \{\delta: \sum_{k=1}^K w_k = \delta \leq 1\}$ for $s = \text{DEA} - \text{NIRS}$;
(iv) $\Gamma(s) = \{\delta: \sum_{k=1}^K w_k = \delta \geq 1\}$ for $s = \text{DEA} - \text{NDRS}$.

It is immediately clear that the resulting constraint is redundant for the CRS case (since both $w_k \geq 0$ and $\delta \geq 0$), and that in all other cases the constraint on the scaling factor δ is in fact integrated into the sum constraint on the activity vector (thus: δ can be dropped). Including these last modifications clearly yields the traditional LP formulation (P.1).

The reverse process of deriving (P.2) from (P.1) proceeds along similar lines.

Proof of Proposition 2:

We assume $x > 0$ and $y > 0$, i.e., strictly positive input and output vectors. This proof could be generalised for semi-positive vectors at the cost of notational simplicity.

In (P.3) the first m constraints on the output dimensions can be rewritten:

$$\begin{aligned} \delta^I &\geq \frac{y_{km}^o}{\sum_{k=1}^K y_{km} z_k}, \quad m = 1, \dots, M. \\ &\geq \text{Max}_{m=1, \dots, M} \left(\frac{y_{km}^o}{y_{km}} \right) = \text{Min}_{m=1, \dots, M} \left(\frac{y_{km}}{y_{km}^o} \right). \end{aligned}$$

The second line follows because only one component of z can equal 1 at the optimum. This can be written as an equality, i.e., binding constraints, for reasons explained below.

The second series of n input constraints in (P.3) yield:

$$\begin{aligned} \lambda &\geq \frac{\sum_{k=1}^K x_{kn} z_k}{x_{kn}^o} \delta^I, \quad n = 1, \dots, N \\ &\geq \text{Max}_{n=1, \dots, N} \left(\frac{x_{kn}}{x_{kn}^o} \right) \delta^I. \end{aligned}$$

Again the second line follows for the same reason mentioned above. This can be rewritten as an equality, since at the optimum at least one constraint must hold with equality.

The first set of constraints determines the critical value of the scaling parameters. If this critical value respects the lower and upper bounds related to the postulated returns to scale assumption, then the rescaled observation is part of the scaled better set; otherwise it is not. Formally:

$$\begin{aligned} \text{If } \delta^I \in \Gamma(s), \quad \text{then } (x_k, y_k) &\in B(x^o, y^o | s); \\ \text{otherwise, then } (x_k, y_k) &\notin B(x^o, y^o | s). \end{aligned}$$

Obviously, the first set of constraints determining the scaling parameter can be substituted into the second set of constraints. Minimising over all $(x_k, y_k) \in B(x^o, y^o | s)$

corresponds to the initial objective function and respects the integer constraints on z_k (i.e., the third and fourth constraints in problem (P.3)). Hence, $DF_i(x,y)$ is computed as:

$$\begin{aligned}\lambda^* = DF_i(x^o, y^o) &= \min_{(x_k, y_k) \in B(x^o, y^o | s)} \max_{n=1, \dots, N} \left(\frac{x_{kn}}{x_{kn}^o} \left(\min_{m=1, \dots, M} \left(\frac{y_{km}}{y_{km}^o} \right) \right) \right) \\ &= \min_{(x_k, y_k) \in B(x^o, y^o | s)} \max_{n=1, \dots, N} \left(\frac{\delta^I x_{kn}}{x_{kn}^o} \right).\end{aligned}$$

The first set of constraints determining δ^I can be rewritten as an equality to guarantee the optimality of the objective function solution. Any other solution would be larger and non-optimal: i.e., $\delta^{Io} > \delta^I$ would imply $\lambda^o > \lambda^*$.

REFERENCES

- Banker, R., H. Chang, W. Cooper (1996) Equivalence and Implementation of Alternative Methods for Determining Returns to Scale in Data Envelopment Analysis, *European Journal of Operational Research*, 89(3), 473-481.
- Banker, R., A. Charnes, W. Cooper (1984) Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis, *Management Science*, 30(9), 1078-1092.
- Bogetoft, P. (1996) DEA on Relaxed Convexity Assumptions, *Management Science*, 42(3), 457-465.
- Chang, K-P., Y-Y. Guh (1991) Linear Production Functions and the Data Envelopment Analysis, *European Journal of Operational Research*, 52(2), 215-223.
- Cummins, D., H. Zi (1998) Measuring Cost Efficiency in the U.S. Life Insurance Industry: Econometric and Mathematical Programming Approaches, *Journal of Productivity Analysis*, forthcoming.
- De Borger, B., G. Ferrier, 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(3), 427-446.
- De Borger, B., K. Kerstens (1996) Cost Efficiency of Belgian Local Governments: A Comparative Analysis of FDH, DEA, and Econometric Approaches, *Regional Science and Urban Economics*, 26(2), 145-170.
- Deprins, D., L. Simar (1983) On Farrell Measures of Technical Efficiency, *Recherches Economiques de Louvain*, 49(2), 123-137.
- Deprins, D., L. Simar, H. Tulkens (1984) Measuring Labor-Efficiency in Post Offices, in: M. Marchand, P. Pestieau, H. Tulkens (eds) *The Performance of Public Enterprises: Concepts and Measurement*, Amsterdam, Elsevier, 243-267.
- Dervaux, B., K. Kerstens, P. Vanden Eeckaut (1998) Radial and Nonradial Static Efficiency Decompositions: A Focus On Congestion Measurement, *Transportation Research*, 32B(5), 299-312.
- Dyson, R., E. Thanassoulis (1988) Reducing Weight Flexibility in Data Envelopment Analysis, *Journal of the Operational Research Society*, 39(6), 563-576.
- Färe, R., S. Grosskopf (1985) A Nonparametric Cost Approach to Scale Efficiency, *Scandinavian Journal of Economics*, 87(4), 594-604.

- Färe, R., S. Grosskopf (1995) Nonparametric Tests of Regularity, Farrell Efficiency and Goodness-of-Fit, *Journal of Econometrics*, 69(2), 415-425.
- Färe, R., S. Grosskopf, C.A.K. Lovell (1983) The Structure of Technical Efficiency, *Scandinavian Journal of Economics*, 85(2), 181-190.
- Färe, R., S. Grosskopf, C.A.K. Lovell (1985) *The Measurement of Efficiency of Production*, Boston, Kluwer.
- Färe, R., S. Grosskopf, C.A.K. Lovell (1994) *Production Frontiers*, Cambridge, Cambridge Univ. Press.
- Färe, R., S. Grosskopf, D. Njineu (1988) On Piecewise Reference Technologies, *Management Science*, 34(12), 1507-1511.
- Färe, R., C.A.K. Lovell (1978) Measuring the Technical Efficiency of Production, *Journal of Economic Theory*, 19(1), 150-162.
- Farrell, M. (1957) The Measurement of Productive Efficiency, *Journal of the Royal Statistical Society Series A: General*, 120(3), 253-281.
- Ferrier, G., K. Kerstens, P. Vanden Eeckaut (1994) Radial and Nonradial Technical Efficiency Measures on a DEA Reference Technology: A Comparison Using US Banking Data, *Recherches Économiques de Louvain*, 60(4), 449-479.
- Ferrier, G., C.A.K. Lovell (1990) Measuring Cost Efficiency in Banking: Econometric and Linear Programming Evidence, *Journal of Econometrics*, 46(1-2), 229-245.
- Førsund, F., L. Hjalmarsson (1974) On the Measurement of Productive Efficiency, *Swedish Journal of Economics*, 76(2), 141-154.
- Førsund, F., L. Hjalmarsson (1979) Generalised Farrell Measures of Efficiency: An Application to Milk Processing in Swedish Dairy Plants, *Economic Journal*, 89(354), 294-315.
- Fried, H.O., C.A.K. Lovell, P. Vanden Eeckaut (1995) Service Productivity in U.S Credit Unions, in: Patrick T. Harker (ed.) *The Service Productivity and Quality Challenge*, Boston, Kluwer, 365-390.
- Fukuyama, H. (1996) Returns to Scale and Efficiency of Credit Associations in Japan: A Nonparametric Frontier Approach, *Japan and the World Economy*, 8(3), 259-277.
- Garfinkel, R., G. Nemhauser (1972) *Integer Programming*, New York, Wiley.
- Kerstens, K. (1996) Technical Efficiency Measurement and Explanation of French Urban Transit Companies, *Transportation Research*, 30A(6), 431-452.
- Kerstens, K. (1998) Decomposing Technical Efficiency and Effectiveness of French Urban Transport, *Annales d'Économie et de Statistique*, forthcoming.
- Kerstens, K., P. Vanden Eeckaut (1998) Estimating Returns to Scale Using Nonparametric Deterministic Technologies: A New Method Based on Goodness-of-Fit, *European Journal of Operational Research*, 113(1), 210-218.
- Koopmans, T. (1951) Analysis of Production as an Efficient Combination of Activities, in: T. Koopmans (ed) *Activity Analysis of Production and Allocation*, New haven, Yale Univ. Press, 33-97.
- Leleu, H. (1997) Élaboration d'un outil d'évaluation de la performance hospitalière: Implications pour une politique d'allocation des ressources, Toulouse, Université Paul Sabatier (thèse doctorale).
- Lovell, C.A.K. (1993) Production Frontiers and Productive Efficiency, in: H. Fried, C.A.K. Lovell, S. Schmidt (eds) *The Measurement of Productive Efficiency: Techniques and Applications*, Oxford, Oxford Univ. Press, 3-67.
- Lovell, C.A.K. (1995) Measuring the Macroeconomic Performance of the Taiwanese Economy, *International Journal of Production Economics*, 39(1), 165-178.

- Petersen, N. (1990) Data Envelopment Analysis on a Relaxed Set of Assumptions, *Management Science*, 36(3), 305-314.
- 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.
- Ruys, P (1974) Production Correspondences and Convex Algebra, in: W. Eichhorn et al. (eds) *Production Theory*, Berlin, Springer, 231-252.
- Thanassoulis, E., R. Dyson, M. Foster (1987) Relative Efficiency Assessments Using Data Envelopment Analysis: An Application to Data on Rates Departments, *Journal of the Operational Research Society*, 38(5), 379-411.
- Tulkens, H. (1993) On FDH Efficiency Analysis: Some Methodological Issues and Applications to Retail Banking, Courts, and Urban Transit, *Journal of Productivity Analysis*, 4(1-2), 183-210.
- Tulkens, H., A. Malnero (1996) Nonparametric Approaches to the Assessment of the Relative Efficiency of Bank Branches, in: Mayes, D. (ed) *Sources of Productivity Growth*, Cambridge, Cambridge University Press, 223-244.
- Tulkens, H., P. Vanden Eeckaut (1995) Non-parametric Efficiency, Progress and Regress Measures for Panel Data: Methodological Aspects, *European Journal of Operational Research*, 80(3), 474-499.
- Vanden Eeckaut, P., H. Tulkens, M. Jamar (1993) Cost Efficiency in Belgian Municipalities, in: H. Fried, C.A.K. Lovell, S. Schmidt (eds) *The Measurement of Productive Efficiency: Techniques and Applications*, Oxford, Oxford University Press, 300-334.
- Varian, H. (1990) Goodness-of-Fit in Optimizing Models, *Journal of Econometrics*, 46(1-2), 125-140.

Filename: FDH-SCE.DOC
Directory: D:\Efficiency\Paper\Scale
Template: C:\Program Files\Microsoft Office\Templates\Normal.dot
Title: FDH-Based Decompositions of Overall Efficiency:
Subject:
Author: Kristiaan Kerstens
Keywords:
Comments:
Creation Date: 09/12/97 15:29
Change Number: 2
Last Saved On: 30/09/98 11:01
Last Saved By: XYQ
Total Editing Time: 10 Minutes
Last Printed On: 19/10/98 12:21
As of Last Complete Printing
 Number of Pages: 31
 Number of Words: 8.563 (approx.)
 Number of Characters: 48.814 (approx.)