RESUME

Nous présentons ici les résultats d'une étude portant sur les performances économiques d'une société de transports publics urbains. Des mesures d'efficacité productive et d'adéquation de l'offre à la demande (ou effectivité) sont calculées mensuellement. Nous observons une tendance claire à l'amélioration des performances techniques conjuguée à des résultats moins favorables en ce qui concerne l'effectivité.

SUMMARY

We present results from a study on the economic performances of a public mass transit firm. Measures of productive efficiency as well as of adequacy between supply and demand (effectiveness) are computed monthly. We observe a clear positive trend for technical performances but less favorable results in terms of effectiveness.

ZUSAMMENFASSUNG

In diesem Artikel präsentieren wir die Ergebnisse einer Studie über die ökonomischen Leistungen eines Unternehmens für öffentlichen Transport in Städten. Masse für die Produktionseffizienz sowie die Angemessenheit zwischen Angebot und Nachfrage (Effektivität) werden für jeden Monat berechnet. Wir stellen eine deutliche positive Entwicklung der technischen Leistungen, jedoch weniger gute Ergebnisse bezüglich der Effektivität fest.

A non-parametric Free Disposal Hull (FDH) approach to technical efficiency: an illustration of radial and graph efficiency measures and some sensitivity results

BRUNO DE BORGER, KRISTIAAN KERSTENS, WIM MOESEN and JACQUES VANNESTE*

INTRODUCTION

A variety of different approaches to the measurement of technical efficiency coexist in the literature[1]. Methodologically, they are categorized according to at least two criteria. First, one distinguishes between stochastic and deterministic methods. Whereas the former make explicit assumptions with respect to the stochastic nature of the data, the latter do not. A second classification differentiates between parametric and non-parametric methods. In the parametric approach it is assumed that the boundary of the production possibility set can be represented by a particular functional form with constant parameters. The non-parametric approach on the contrary concentrates on the regularity assumptions of the production possibility set itself. Imposing some plausible restrictions on the production process the latter methods directly construct a piecewise linear reference technology or best practice frontier on the basis of observed input-output combinations.

Recently, DEPRINS, SIMAR and TULKENS (1984) and TULKENS (1986, 1993) suggested the Free Disposal Hull (FDH) as a new deterministic and non-parametric reference technology for the evaluation of productive efficiency. Compared to other existing methods the FDH requires minimal assumptions with respect to the production technology. For example, it does not require convexity – in contrast to the popular Data Envelopment Analysis (DEA) models[2]. As there is no generally accepted model of governmental behavior, the minimal technical and behavioral assumptions underlying the FDH make it a particularly useful tool for analyzing public sector efficiency questions. Not surprisingly, since its introduction a number of empirical studies have appeared in which the approach is applied to evaluate the technical efficiency of a number of public service providers as well as a few private enterprises (for a review, see PESTIEAU

* For all correspondence, contact the first author: University of Antwerp, UFSIA-SESO, Prinsstraat 13, B-2000 Antwerpen, Belgium. Affiliations of the other authors: K.U. Leuven and UFSIA; and UFSIA and KULAK respectively. The authors thank J. COUDER for his superb programming assistance and I. JANSSENS, D. DE GRAEVE, the editor of this journal and two referees for their comments on an earlier version. The authors are responsible for any remaining errors.

1. For surveys of these approaches see, e.g., FØRSUND, LOVELL and SCHMIDT (1980), SCHMIDT (1986), and THIRY and TULKENS (1989).
2. See FÄRE, GROSSKOPF, and LOVELL (1985) and SEIFORD and THRALL (1990) for a review.

and TULKENS [1990] and TULKENS [1993]). Both the theoretical and empirical work have clarified the main advantages and disadvantages of the FDH reference technology.

This paper elaborates upon DE BORGER et al. (1994) and serves three purposes. First, using information on 589 local authorities we apply the FDH production technology to evaluate the technical efficiency of the provision of municipal services in Belgium. Second, we use this data set to illustrate in a systematic way the strengths and weaknesses of the FDH. Specifically, we assess the sensitivity of the efficiency results with respect to the number of input and output dimensions and with respect to sample size, and we consider the impact of the existence of outliers on efficiency scores. Third, contrary to common practice in non-parametric efficiency analyses, we argue in favour of graph efficiency measurement. This contrasts with the common use of radial input or output efficiency measures, as proposed in FARRELL (1957). Almost all existing empirical studies confine the attention to measuring either input or output efficiency (see, e.g., TULKENS [1993], FÄRE, GROSSKOPF and LOGAN [1985]). In this paper, however, we do not restrict the analysis to separate input and output indices but also calculate two graph efficiency measures that take account of all dimensions simultaneously.

The paper unfolds as follows. The first section deals with methodological issues. We review the FDH reference technology for measuring technical efficiency, and we systematically discuss its advantages and shortcomings. We then review the different input, output and graph efficiency measures that are used in the empirical analysis. In Section 2 we apply the suggested methodology to study the efficiency of local public service provision by Belgian municipalities. The sensitivity of the results with respect to sample size, the existence of outliers and the number of dimensions is illustrated in Section 3. Some further reflections and a conclusion are provided in Section 4.

1. THE FREE DISPOSAL HULL APPROACH TO PRODUCTIVE EFFICIENCY

A production unit is technically efficient if it produces the maximum output which is technically feasible for given inputs, or uses minimal inputs for the production of a given level of output. In other words, technical or productive efficiency of a production unit is defined in terms of the ability of the unit to produce on the boundary of its production set[3]. Consequently, any methodology for evaluating technical efficiency requires the complete specification of the production possibility set as well as some concept of distance to relate the observed input-output combinations to the boundary of the specified set[4]. We therefore first characterize the FDH reference technology by specifying its assumptions regarding the production set, and then present various efficiency measures which relate observations to the boundary of the FDH. The third subsection deals with the computation of efficiency measures on the FDH. We conclude this section with a review of the advantages and shortcomings of the FDH production technology.

3. A complete characterization of types of efficiency is found in FÄRE, GROSSKOPF and LOVELL (1985).
4. Recently TULKENS and VANDEN EECKAUT (1993) relaxed the need for the representation of the boundary of the production possibility set by defining technical efficiency solely in terms of pairwise dominance relations.

1.1 The FDH reference technology

Let $y = y(y_1,y_2,...,y_n)$ be the n non-negative outputs produced by using various combinations of the m non-negative inputs $x = x(x_1,x_2,...,x_m)$. The production possibility set Y is the set of all input and output combinations which are technically feasible:

$$Y = \{ (x,y) \mid x \in \mathbb{R}^m_+, y \in \mathbb{R}^n_+, (x,y) \text{ is feasable} \} \tag{1}$$

It is convenient to model the production technology by an input correspondence $y \to L(y) \subseteq \mathbb{R}^n_+$. For any output vector y, the level set L(y) denotes the subset of all input vectors x which yield at least the output vector y[5].

Different production technologies are defined by imposing various restrictions on L(y). The non-parametric approaches typically impose very weak assumptions. Although they vary widely, they are generally less restrictive than those used in the parametric approaches[6]. Moreover, it is fair to say that the FDH reference technology imposes one of the mildest assumptions among the deterministic, non-parametric alternatives. Specifically, the following axioms define the Free Disposal Hull[7]:

$$0 \notin L(y) \textit{ for } y \geq 0, \textit{ and } L(0) = \mathbb{R}^n_+ \tag{2.1}$$

$$\textit{If } \|y^l\| \to +\infty \textit{ as } l \to +\infty, \textit{ then } \bigcap_{l=1}^{+\infty} L(y^l) \textit{ is empty} \tag{2.2}$$

$$\textit{If } x \in L(y) \text{ and } x' \geq x, \textit{ then } x' \in L(y) \tag{2.3}$$

$$L(y) \textit{ is a closed correspondence} \tag{2.4}$$

$$\textit{If } y' \geq y, \textit{ then } L(y') \subseteq L(y) \tag{2.5}$$

5. See FÄRE, GROSSKOPF, and LOVELL (1985) and VARIAN (1984).
6. GROSSKOPF (1986) and SEIFORD and THRALL (1990) review the deterministic, non-parametric reference technologies.
7. See, e.g., DEPRINS, SIMAR and TULKENS (1984). Note that the notion of a free disposal hull originally referred to the property of strong free disposal and not to any particular reference technology (see MC FADDEN [1978]).

The intuition behind each of these axioms is straightforward. Axiom 1 states that a semipositive output cannot be obtained from a null input vector – thus excluding free production – and that any nonnegative input results at least in a zero output. The second axiom implies that finite inputs cannot produce infinite outputs. Axiom 3, known as strong free disposability of inputs or positive monotonicity, guarantees that an increase in inputs cannot result in a decrease in outputs. In axiom 4 it is stated that if a sequence of input vectors can each produce y and converges to x^*, then x^* can also produce y. Closedness is an axiom postulated for mathematical convenience which cannot be contradicted by any empirical observation[8]. Axiom 5, known as strong free disposability of outputs, implies that any reduction in outputs remains producible with the same amount of inputs. This assumption allows for variable returns to scale.

The FDH is now easily defined as a piecewise linear reference technology, constructed on the basis of observed input-output combinations, that satisfies the above axioms. The FDH input correspondence is specified as:

$$L(y)^{FDH} = \{ x \mid x \in \mathbb{R}^m_+, z'N \geq y, z'M \leq x, I'_k z = 1, z_i \in \{0, 1\} \} \tag{3}$$

where N is the k×n matrix of observed outputs, M is the k×m matrix of observed inputs, z is a k×1 vector of activity or intensity variables, and I_k is a k×1 vector of ones. Consistent with the idea of variable returns to scale, the activity vector is restricted to sum to unity. Since the activity vector contains either zeros or ones linear combinations of several observations are excluded. Clearly, the axioms did not impose convexity on the technology.

We have focused so far on the FDH input correspondence L(y). Obviously, this technology can equivalently be characterized using the output or the graph correspondence. The output correspondence is the subset of all output vectors y which are obtained from the input vector x. Based on similar axioms, the FDH output correspondence is given by:

$$P(x)^{FDH} = \{ y \mid y \in \mathbb{R}^n_+, z'N \geq y, z'M \leq x, I'_k z = 1, z_i \in \{0, 1\} \} \tag{4}$$

Finally, the FDH graph correspondence is defined with respect to either the input or the output correspondence:

$$\begin{aligned} GR^{FDH} &= \{(x,y) \mid x \in L(y)^{FDH}, x \in \mathbb{R}^m_+, y \in \mathbb{R}^n_+ \} \\ &= \{(x,y) \mid y \in P(x)^{FDH}, x \in \mathbb{R}^m_+, y \in \mathbb{R}^n_+ \} \end{aligned} \tag{5}$$

8. For further interpretation, see FÄRE, GROSSKOPF and LOVELL (1985, p. 25).

To develop some intuition for the graphical representation of the FDH reference technology, notice that, reflecting free disposal, each observed input-output combination adds one orthant, positive in the inputs and negative in the outputs, to the production set. The FDH reference technology is then the boundary to the union of all such orthants. Its graph section as well as its input isoquants, illustrated respectively in Figures 1 and 2, typically have a staircase form. To clarify the above definitions of the correspondences, observe that for an efficient or frontier observation the corresponding component of z equals one and the restrictions in (3) and (4) hold with equality. For an inefficient observation in an orthant spanned by a boundary observation j, the j-th component of z equals one and the inequalities hold, as the dominated observation uses more inputs to produce less outputs than the boundary observation j.

Otherwise formulated, the construction of the FDH boundary closely follows the definition of technical efficiency in that it is solely based on weak vector dominance reasoning. An observation is declared inefficient if it is possible to find at least one other observation which contains the same or more outputs but strictly less of at least one input, or which uses the same or less inputs to produce strictly more of at least one output. Input-output combinations which are undominated are declared efficient. However, efficient observations that never dominate other observations have been aptly called 'efficient by default'[9]. Due to the partial ordering implied in the weak vector dominance reasoning, one is unable to make precise statements concerning their technical efficiency. Nevertheless, it is useful to distinguish them from efficient observations that do dominate others. These concepts are illustrated in Figure 2. Observation a is dominated by observations 3 and 4, but dominates observation b. Observations 1,5,6 and 7 are efficient by default. Finally, the effect of not imposing convexity is easily indicated. Observation 5 is efficient although, had convexity been imposed, it would have been inefficient relative to the linear combination of observations 4 and 6.

1.2. Measures of technical efficiency: Definitions

Once the boundary of the reference technology has been determined, technical efficiency is measured as the distance between an observed production unit and the postulated boundary. In the non-parametric approach attention is often restricted to the measurement of either input or output efficiency. Furthermore, it is common to restrict the attention to radial or Farrell measures[10]. For ease of comparison we stick to the tradition of radial measurement. However, we do not limit the analysis to input and output efficiency, but also calculate graph technical efficiency measures.

9. See VANDEN EECKAUT, TULKENS, and JAMAR (1993).
10. Note, however, that due to the non-convex nature of the FDH, radial efficiency measures leave a lot of slacks. One could therefore easily argue in favour of non-radial efficiency measures: see RUSSELL (1988) for an overview.

Figure 1: FDH input section

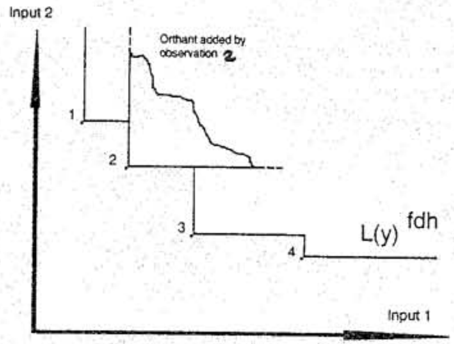
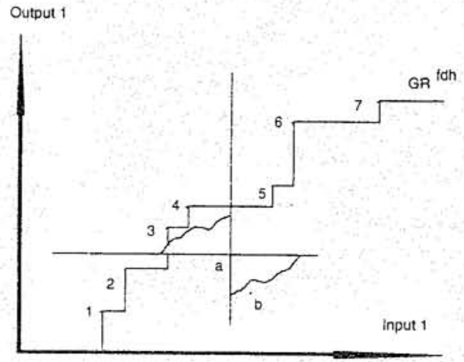


Figure 2: FDH graph section



1 to 7: Efficient observations
1, 5 to 7: Efficient by default
a: Dominated by 3 and 4; Dominates b
b: Dominated by 2, 3, 4 and a

This interest in graph technical efficiency measurement can be justified on several accounts. First, as the production possibility set is a primitive choice set of technological options, we object in general to imposing restrictive assumptions on the choice of an orientation of measurement without checking the specific behavioural assumptions. Most of the empirical literature focuses on radial efficiency measurement in either the inputs or the outputs, depending on whether the inputs or the outputs are the decision variables under the control of the production unit[11]. For example, if it can be assumed that for the public sector cost minimization is a more likely behavioural postulate than output maximization or any other objective, then restricting attention to input efficiency is considered legitimate. One may wonder whether in practice so few organizations have control over both their inputs and outputs. We argue instead that graph measurement is the adequate basic orientation, unless one has strong reasons, e.g. based on theoretical assumptions and statistical tests, to limit the attention to input or output efficiency measures. This is especially the case when there is ignorance or uncertainty on the issue of control over inputs and outputs, i.e. on organizational goals. Second, if the purpose of the analysis is to rank the production units according to technical efficiency, then a priori some overall measure may prove more informative than a detailed two part analysis of input and output efficiency. Restricting the measurement of technical efficiency to either the input or the output dimensions yields only a partial view on performance.

The various efficiency measures used in this paper are easily defined. The Farrell input measure of technical efficiency is given by:

$$F_i(x, y) = \mathit{min}\ \{\lambda \mid \lambda \geq 0, \lambda\, x \in L(y)\} \tag{6}$$

This efficiency measure determines the maximum equiproportionate reduction in all inputs which still allows production of the given outputs. It varies between zero and one, with unity representing efficient production. Analogously, the Farrell output measure is defined as:

$$F_o(x, y) = \mathit{max}\ \{\mu \mid \mu \geq 1, \mu\, y \in P(x)\} \tag{7}$$

It is no smaller than unity and determines the maximal proportional expansion in all outputs while still using the same input. Observe that in the empirical section we report its inverse, which is smaller than unity, to facilitate the comparison[12].

11. See Färe, Grosskopf and Lovell (1985), p. 16.
12. Or formally:

$$F_o(x, y)^* = \mathit{min}\ \{\mu^* \mid \mu^* \leq 1, \frac{y}{\mu^*} \in P(x)\} \tag{7'}$$

Various graph measures have been proposed in the literature. In this paper we use two measures of the Farrell type, both varying between zero and unity[13]. First, the Farrell graph measure of technical efficiency is defined as:

$$F_g(x, y) = \min \{ \lambda \mid \lambda \geq 0, (\lambda x, \lambda^{-1} y) \in GR \} \tag{8}$$

It looks for the maximal equiproportionate reduction of all inputs and increase of all outputs. Finally, the Generalized Farrell graph measure allows the proportional reduction of all inputs to differ from the proportional increase of all outputs and simply takes the average:

$$F_g^G(x, y) = \min \{ \frac{\lambda + \mu}{2} \mid \lambda \geq 0, \mu \geq 0, (\lambda x, \mu^{-1} y) \in GR \} \tag{9}$$

It has been shown that $F_i(x,y) = F_o(x,y)$ if and only if the technology satisfies constant returns to scale. As the FDH allows for variable returns to scale the Farrell input and output measures will generally differ. Also note that $F_g(x,y) \geq \max [F_i(x,y), F_o(x,y)]^{-1}$ and that $F_g(x,y) = 1$ if and only if either $F_i(x,y) = 1$ or $F_o(x,y) = 1$. Finally, observe that for $m = n = 1$, $F_g^G(x,y) = F_g(x,y)$ and that $F_g^G(x,y) < F_g(x,y)$ for $\lambda \neq \mu$[14].

1.3. *FDH and efficiency measures: Computational aspects*

To illustrate the ease of obtaining efficiency measures relative to an FDH production technology, we consider in this section the computation of the Farrell input measure of technical efficiency. The computation of efficiency measures on FDH normally requires solving one mixed integer programming problem for each observation, because the activity variables (z) in the definition of the production technology are constrained to be either zero or unity. However, this does not make its implementation more difficult than the standard linear programming problems solved in DEA, as it has been shown that a data classification algorithm based on simple vector dominance reasoning can do the job (see, e.g., TULKENS [1993])[15].

13. Observe that the Graph and the Generalized Graph Farrell efficiency measures are radial in the input and the output sections, but respectively rectangular hyperbolic and hyperbolic in the graph sections (see FÄRE, GROSSKOPF, and LOVELL [1985, p. 126] for details).
14. For details on the relations between these efficiency scores: FÄRE, GROSSKOPF, and LOVELL (1985), chapter 6. While on the FDH input, output and graph Farrell efficiency measures eliminate slacks in at least a single dimension, the generalized graph Farrell efficiency measure at least eliminates slacks in one input and one output dimension.
15. The algorithms for the other efficiency measures are in an appendix which is available upon request. We also selected this data classification algorithm because of its ease of programming and because it allows one to generate useful additional information. The program was developed in Turbo Pascal.

The procedure operates basically in two steps. (i) Define for each observation (x^o,y^o) to be evaluated an index set $DO(x^o,y^o)$ containing the observations which weakly dominate (x^o,y^o) in that they produce at least as much of each output with no more of any input. Or formally:

$$DO(x^o, y^o) = \{ (x_i, y_i) \mid x_i \leq x^o, y_i \geq y^o \} \tag{10}$$

(ii) Calculate the radial efficiency measure in the inputs $[F_i(x,y)]$ by applying the following algorithm:

$$F_i(x, y) = \min_{(x_i, y_i) \in DO(x^o, y^o)} \quad \max_{l = 1, \ldots, m} \left\{ \frac{x_{il}}{x_{o_l}} \right\} \tag{11}$$

where x_{il} denotes the l-th component of the input vector of observation i and x_{o_l} represents the same component of the observation (x^o,y^o) being evaluated. The first step constructs the FDH boundary and, consequently, provides the classification between efficient and inefficient observations mentioned earlier, while the second step computes the efficiency measures relative to this FDH boundary. The element of the index set $DO(x^o,y^o)$ relative to which the efficiency measure reaches its minimum is called the most dominating observation. Note that the most dominating observation may be a different observation depending on the efficiency measure being used. Identifying the most dominating observations provides useful information concerning the opportunities available for improving efficiency.

1.4. *FDH and efficiency measurement: Advantages and shortcomings*

The advantages and disadvantages of the FDH reference technology are summarized from two perspectives[16]. First, we evaluate the production technology from the theoretical and empirical point of view. Then we discuss its merits and inconveniences from the managerial viewpoint.

From a theoretical and empirical point of view, the FDH makes very weak assumptions regarding the production technology. Apart from FDH, the least restrictive technology used so far in the non-parametric approach assumes weak disposability instead of strong disposability. But these technologies always assumed convexity[17]. Furthermore, it can be argued that the assumptions of strong free disposal in inputs and outputs

16. See e.g. BÔS (1988), THIRY and TULKENS (1989), and especially TULKENS (1993).
17. See GROSSKOPF (1986, p. 504). But see PETERSEN (1990) who relaxes the assumption of convexity.

have a strong intuitive appeal since they are closest to the concept of technical efficiency itself. A dominated observation is inefficient due to its excessive usage of resources or due to its lack of outputs compared to another observation, irrespective of formal convexity or functional form considerations[18].

In FDH the problem of measuring the technical efficiency of the observed production units is separated from the issue of representing the boundary of the production possibility set. Because it is a multidimensional stepfunction, this reference technology is less useful in answering other questions on, e.g., the determination of factor productivity, of economies of scale and of scope, etc.[19]. These problems require focussing on the boundary of the production possibility set and are difficult to solve without resort to parametric production or transformation functions. Here the more restrictive technologies considered in the parametric approach may well be indispensable[20].

A second advantage of the FDH is its non-parametric nature. It is a general methodological requirement that the results of theoretical economic analysis should not depend on specific parametric forms chosen. However, in empirical work specific parametrizations are often crucial. It is then implicitly postulated that the parametric forms selected are good approximations for the true functional relationships. This maintained hypothesis is, however, not directly testable. Therefore, it has been argued that both theoretical and empirical work should stay as close as possible to the raw data[21]. Furthermore, it has recently been argued that the non-parametric reference technologies and the resulting efficiency measures are related to the results of the parametric approach: the former provide upper bounds to the latter[22]. With respect to both the parametric approaches and the non-parametric methods that impose convexity, the FDH is therefore considered conservative, as it yields an upper bound to technical efficiency measurement.

Like any methodology the FDH has some drawbacks. The most obvious problem is due to the partial ordering based on the vector dominance reasoning. It implies that the approach may be sensitive both to the number and distribution of the observations in the data set, and to the number of input and output dimensions considered. Increasing the sample size increases the possibility of dominance for any given observation, and therefore the probability of being denoted inefficient. Also a rather uniform distribution of the observations over the dimensions in the data set increases the possibility of dominance. On the other hand, an increase in the dimensions considered decreases this possibility. Therefore, one expects that incorporating more inputs or outputs into the analysis increases the probability of efficiency. Moreover, all deterministic approaches, which envelop the efficient observations as closely as possible, may be sensitive to

18. In the case of undesirable outputs the assumption of strong free disposal of outputs is disputable: see FÄRE, GROSSKOPF, LOVELL and PASURKA (1989) for details.
19. A point developed in TULKENS (1990).
20. The non-parametric approach can still be useful, viz. as a first step in the estimation of parametric frontiers. For applications of this method, see THIRY and TULKENS (1992) and TULKENS (1993).
21. See, e.g., VARIAN (1984).
22. See BANKER and MAINDIRATTA (1988).

efficient outliers (while they may be less sensitive to inefficient outliers). Notice, however, that among the non-parametric deterministic approaches the FDH is least sensitive to this defect. Each observation potentially only adds a small subset, i.e. an orthant, to the reconstruction of the production set. Therefore outliers only affect a small subset of observations.

From a managerial viewpoint, the major advantage of the FDH is that the resulting efficiency measures are related to an observed production unit. In most other methods the point of reference is a hypothetical construct. For example, an observation may be inefficient with respect to some convex combination of observations in the non-parametric DEA, or with respect to some fitted value on a postulated frontier in the stochastic frontier approach. It may be difficult to convince managers that they are outperformed by such a hypothetical unit. They can always object that these convex combinations of observed activities are not feasible, or that they cannot learn how to improve from an unobservable standard of comparison[23]. A final advantage is that additional information is readily available. For example, the set of dominating observations can provide useful information in designing stepwise improvements in the direction of a production unit on the frontier. The possibilities of the FDH to improve productivity, to reward production units, etc., are clear[24].

2. AN APPLICATION TO BELGIAN MUNICIPALITIES

In this section we determine technical efficiency of all 589 Belgian municipalities using the FDH production technology[25]. The choice of input and output indicators has been motivated both by the desire to account for the most important local public services provided, and by the availability of data. Our basic data set has one input indicator, defined as total municipal staff, and five output indicators. The latter capture important aspects of local production in the field of education, transportation, and social and recreational services. The following outputs were used:

(i) the surface of municipal roads
(ii) the number of beneficiaries of minimal subsistence grants
(iii) the number of students enrolled in local primary schools
(iv) the area of public recreational facilities
(v) a 'residual' output defined as total municipal outlays minus the identifiable outlays on outputs (i), (ii), (iii) and (iv).

23. See the remarks in EPSTEIN and HENDERSON (1989).
24. Its use in public sector management has been developed in PESTIEAU and TULKENS (1990).
25. VANDEN EECKAUT and TULKENS (1988) and VANDEN EECKAUT, TULKENS and JAMAR (1993) have reported results for the Belgian local authorities using FDH. This paper differs from their studies on four accounts. First, their sample is restricted to the Walloon region. Second, they use somewhat different input and output indicators. Third, their analysis does not consider graph efficiency measures. Finally, they do not engage in the kind of sensitivity testing reported in Section 3 below.

Several remarks are in order. First, some justification of the inclusion of the residual output is warranted. From municipal accounts we verified that the first four output indicators capture between 30% and 75% of municipal outlays. Therefore, the fifth output attempts to correct for other unobserved outputs. If it were not included, then municipalities that spend a large fraction of their budget for the production of outputs not captured by our first four indicators would be incorrectly assigned very low efficiency scores. The residual output should largely eliminate the possible bias in the efficiency ranking on this account. Second, observe that the available outputs only very crudely proxy for the underlying services provided by local authorities, and that no information on capital inputs was available. As a consequence, our study may have a limited scope and the results should be interpreted with care.

Table 1: Descriptive statistics: Farrell technical efficiency in the main data set

	Farrell Graph	Farrell Input	Farrell Output	Farrell Gen.Graph[c]
Mean	.927[a]	.861	.878	.893
	.855[b]	.724	.757	.788
Median	1.000	1.000	.999	.995
Minimum	.490	.235	.268	.466
Standard Deviation	.105	.189	.167	.130
	.107	.181	.160	.107
Kurtosis	4.732	3.245	3.582	2.701
Skewness	−1.492	−1.162	−1.152	−0.901
# Most dominating observations	74 (13%)	72 (12%)	80 (14%)	65 (11%)
# Inefficient Observations			297 (50%)	
# Efficient Observations by default			59 (10%)	

\# Number of...
a All observations
b Inefficient observations only
c Generalized Graph

Application of the weak vector dominance analysis on our main data set yields the summary results reported in Table 1[26]. This offers a crude classification of observations on or below the FDH boundary. The results indicate that about 50% of the observations are inefficient. Among the efficient observations some 60 observations are efficient by default. This preliminary analysis is supplemented by computing the four efficiency measures outlined above.

26. Note again that in the empirical results we report the reciprocal of the output Farrell efficiency measure to facilitate comparison with the other measures.

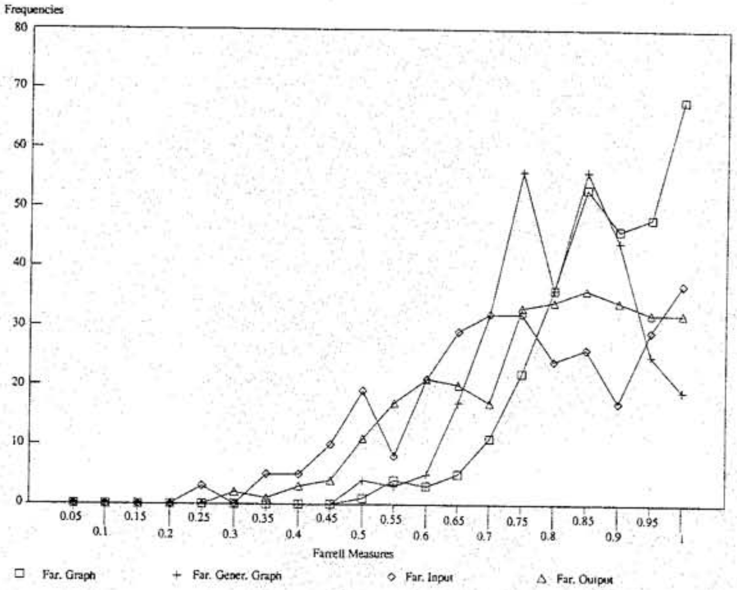
The resulting efficiency measures are also reported in Table 1. Mean efficiency varies between 0.86 and 0.93. It is the lowest for the Farrell input measure and the highest for the Farrell graph measure. The Farrell input and output measures have the largest standard deviations and the lowest minimum. The distribution of the efficiency measures covers a wide range (from 0.23 to 1) and is obviously rather skewed. A histogram of the frequencies is drawn in Figure 3. Since the non-parametric approach provides upper bounds for the estimation of efficiency, the mode at 1 can be interpreted as the discrete part of a censored distribution. The inefficient observations are most dominated by a subset of 11% to 14% of the observations, depending on the efficiency index used.

Although the use of different efficiency measures does not lead to dramatically different mean efficiency levels, one may wonder to what extent they imply different rankings for individual observations. Correlation coefficients, once for all observations and once for the inefficient observations only, are reported in Table 2. The ranking implied by the Farrell input and output measures correlate least. If one considers the inefficient observations only, the correlation coefficient amounts to 0.59. The Farrell generalized graph measure clearly correlates best with the other measures. This is not entirely surprising, as it is the only measure that takes account of differences in all inputs and outputs while at the same time allowing different proportional changes in each of these two orientations. As it is no more difficult to compute than the other efficiency measures, it probably deserves more attention in empirical applications.

Table 2: Correlations between Farrell efficiency measures

	All observations				Inefficient observations only			
	Farrell				Farrell			
	Graph	Input	Output	G.Graph	Graph	Input	Output	G.Graph
Farrell Graph	1.00				1.00			
Farrell Input	.89	1.00			.78	1.00		
Farrell Output	.89	.81	1.00		.78	.59	1.00	
Farrell G. Graph	.94	.94	.93	1.00	.89	.87	.84	1.00

Figure 3: Histogram frequency: Farrell technical efficiency (inefficient observations only)



3. SOME SENSITIVITY RESULTS

In this section we report the results of a sensitivity analysis on the Belgian municipalities. In Section 1.4 we indicated the major strengths and weaknesses of the FDH. It was suggested that the method could be sensitive to the number of inputs and outputs taken into account, to the sample size, and to the existence of outliers. We investigate these claims in some detail for our data set below.

First, we test the effect of the sample size by taking random samples of increasing size from 50 up to 550. For each size we considered five random samples[27]. In each case we report the average results over these five samples in Table 3. The results indicate clearly that increasing the sample size increases not only the absolute, but also the relative number of inefficient observations. The process is apparently highly nonlinear. Although

27. A more satisfactory procedure is to use bootstrapping techniques to approximate the sampling distribution of the efficiency measures. This is an obvious direction for future work.

the differences in the proportion of inefficient observations seem to level off for sample sizes above 400, there is no indication that further increases in sample size have no impact on the percentage of inefficient observations. Interestingly, larger sample sizes seem to have a much less pronounced impact on the absolute number of observations that are efficient by default, except for the smallest samples. Differences between samples of 200 observations and more are almost negligible.

As expected, increasing the sample size decreases the average efficiency measures and increases their standard deviation. However, except for the smallest samples, the differences are trivial. Also, observe that the average Farrell input and output measures vary most and have larger standard deviations than the graph efficiency measures. Finally, for all four measures larger sample sizes increase the number of most dominating observations.

Second, to determine the impact of outliers we eliminated the outliers from our main data set of 589 municipalities using a procedure outlined in BELSLEY, KUH and WELSCH (1980). The method employed constructs a test statistic based on the leverage value of each observation ($h_i = x_i(X'X)^{-1}x_i'$, i.e. the diagonal element of $X(X'X)^{-1}X'$). In our case X is a $kx(m + n)$ datamatrix with k observations and $(m + n)$ input and output dimensions. The leverage value determines the importance of the observations in the data space spanned by all dimensions. Use of the appropriate test statistic resulted in the detection of 35 outliers, including the 5 largest Belgian cities[28]. From these outliers 31 were efficient in the original analysis, and 13 among these were efficient by default.

We recomputed the four efficiency measures based on the data set obtained after deleting the 35 outliers. The result of this exercise is also reported in Table 3. Despite the fact that most of the outliers were efficient, their impact both on the number of inefficient observations and on the distribution of the efficiency scores is very small. Dropping the outliers results in a decrease in both the relative number of efficient observations and most dominating observations. Furthermore, we observe a marginal decrease in the average efficiency measure. These aggregate findings obviously do not necessarily imply the unimportance of correcting for outliers, as the effect on the efficiency scores of some individual observations may well be substantial.

Third, we tried to illustrate the effects of disaggregation and aggregation, i.e., the impact of variations in the number of dimensions. Because of data limitations there was unfortunately no scope at all for increasing the number of inputs and outputs taken into account in the production analysis. Therefore, this part of the sensitivity analysis necessarily remains somewhat unsatisfactory. We proceeded as follows. The main data set has 5 output dimensions and 1 input dimension, a total of 6 dimensions. Aggregation was achieved by dropping output(s) while in each case recalculating the 'residual' output. We calculated efficiency measures for the four combinations to drop one output and for the six combinations to drop two outputs, while in each of these cases the additional

28. The complete list of outliers is available upon request.

output was recalculated. To keep the results tractable we only report in Table 3 the average results for each level of aggregation. These results suggest, consistent with a priori expectations, that increasing the number of dimensions decreases the number of inefficient observations and increases the number of observations efficient by default. Mean efficiency scores increase while their standard deviations decrease. It is somewhat reassuring that the variability in mean efficiency is quite small, despite the large impact on the fraction of efficient municipalities. This implies that, if the analysts' main interest is in computing average efficiency levels, aggregation may not be too harmful[29]. Of course, if one is interested in the precise distribution of efficiency scores over the sample, this statement will probably be incorrect, as the impact of aggregation on individual observations may be nontrivial. Observe, furthermore, that the number of observations on which efficiency measurement depends, i.e., the set of most dominating observations, does not seem to vary systematically with the number of dimensions.

Finally, we attempted to detect the sensitivity of the results to variable selection. Including or excluding critical variables may be interpreted as providing information on the importance of possible misspecification. A variation of the exercise to test for the effect of aggregation was used to investigate the impact of critical variables. Whereas in the case of testing for the impact of aggregation the residual output was systematically recalculated, in the present exercise the residual output was completely ignored. We simply varied the number of outputs in the analysis. The base case for this exercise has five dimensions: one input and four outputs. We calculated efficiency measures for the four combinations to drop one output and for the six combinations to drop two outputs. These sensitivity results are again presented in Table 3 and are similar to the results of the aggregation exercise. Although it is difficult to compare both exercises, it seems that omitting critical variables leads to a stronger reduction in average efficiency and to somewhat more variability in the efficiency measures, as is clear from the increased standard deviation. Thus misspecification leads to a more serious bias in efficiency than aggregation. This is as expected: misspecification can have a significant effect on any estimation procedure. It is however comforting to know that the FDH reference technology is not particularly vulnerable to this problem[30].

29. Tulkens, Thiry and Palm (1988) found similar indications for the FDH reference technology. This is also analogous to the results in DEA reported in Seiford and Thrall (1990).
30. Analogous results in DEA are reported in Seiford and Thrall (1990).

4. SUMMARY AND CONCLUSIONS

The FDH is an alternative deterministic, non-parametric production technology for the evaluation of productive efficiency. The purpose of this paper was threefold. First, we calculated various measures of technical efficiency for a data set of 589 Belgian local governments using the FDH. Second, based on a priori reasoning as well as on the basis of the empirical results obtained we argued in favour of graph efficiency measurement instead of limiting the analysis to either input or output efficiency. Finally, we attempted to illustrate – insofar as possible – the strengths and weaknesses of the FDH using our municipal data set.

First, we presented the methodology for measuring productive efficiency based on the FDH reference technology. Apart from input and output efficiency measures, two graph efficiency indices were defined and the computation of efficiency measures on the FDH was outlined. The advantages and drawbacks of the FDH were systematically discussed. In Section 2 the FDH was used to study the efficiency of local public service provision by Belgian municipalities. The main conclusions were that the FDH has considerable advantages relative to alternative methods from the theoretical, empirical and managerial viewpoints. These have to be traded off against some disadvantages, such as the sensitivity with respect to sample size and the number of inputs and outputs taken into account in the analysis. This sensitivity was illustrated in a third section.

Two final conclusions emerge from this paper. First, the FDH offers a useful reference technology for evaluating technical efficiency. It works best when all aspects of the production process can be captured in a limited number of input and output dimensions, and when a relatively large sample is available. Moreover, it generates a wealth of additional information which is useful for managerial purposes. For example, the set of dominating observations and the identification of a most dominating observation are particularly useful. Second, the empirical results provide evidence in favour of the use of graph efficiency measures. Especially the Generalized Farrell Graph measure is a promising efficiency index deserving more attention in future empirical work.

Table 3: Sensitivity results for samples of different sizes and number of dimensions

			Farrell Graph			Farrell Input			Farrell Output			Gen. Farrell Graph		
Sample Size or # Dimensions	#Inef. Obs.(%)[a]	# Effic. Obs. by def(%)[b]	Mean	StDev[c]	#Most domin. Obs.[d]	Mean	StDev	#Most domin. Obs.	Mean	StDev	#Most domin. Obs.	Mean	StDev	#Most domin. Obs.
Main Data Set:														
589	297(50.4)	59(10.0)	.927	.105	74	.861	.189	72	.878	.166	80	.893	.130	65
Size:														
50	8(15.6)	33(6.6)	.980	.054	5	.968	.087	5	.965	.087	5	.970	.072	5
100	24(23.8)	48(4.8)	.973	.064	16	.946	.124	13	.956	.102	13	.957	.089	12
150	51(34.0)	53(35.1)	.959	.079	22	.917	.150	23	.929	.128	23	.935	.107	21
200	72(36.2)	55(27.7)	.955	.082	32	.911	.153	30	.923	.133	29	.929	.111	26
250	92(37.0)	57(23.0)	.949	.090	33	.905	.160	33	.914	.147	35	.924	.118	29
300	124(41.5)	54(18.1)	.945	.093	43	.897	.166	42	.907	.148	46	.918	.120	38
350	148(42.3)	59(17.0)	.943	.095	49	.892	.170	47	.902	.154	48	.914	.122	40
400	183(45.8)	59(14.6)	.937	.099	53	.881	.177	52	.892	.158	56	.906	.125	46
450	218(48.5)	56(12.5)	.931	.102	61	.869	.185	57	.886	.160	65	.898	.129	52
500	237(47.4)	59(11.9)	.934	.099	67	.872	.182	64	.887	.161	69	.901	.126	56
550	269(49.0)	61(11.2)	.930	.103	72	.865	.186	68	.882	.164	77	.896	.129	62
Outliers:														
554	287(51.8)	49(8.8)	.924	.107	68	.856	.190	66	.874	.167	73	.888	.132	60
Aggregation:														
4	440(74.7)	16(2.8)	.878	.125	78	.764	.204	72	.793	.182	76	.827	.134	58
5	371(63.1)	29(4.9)	.912	.102	81	.825	.189	73	.849	.165	84	.870	.125	63
Misspecification:														
3	519(88.2)	4(0.7)	.703	.188	51	.514	.268	44	.558	.251	49	.621	.192	35
4	447(75.8)	9(1.6)	.813	.167	72	.664	.269	65	.710	.240	75	.742	.190	56
5	358(60.8)	20(3.4)	.876	.147	78	.773	.247	88	.808	.216	87	.823	.179	77

a # Number of Inefficient Observations
b # Number of Efficient Observations by default
c Standard Deviation
d # Number of Most Dominating Observations

REFERENCES

BANKER, R., A. MAINDIRATTA, "Nonparametric Analysis of Technical and Allocative Efficiencies in Production", *Econometrica* 56(6), 1988, 1315-1332.

BELSLEY, D., E. KUH, R. WELSCH, *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*, John Wiley, New York, 1980.

BÖS, D., "Recent Theories on Public Enterprise Economics", *European Economic Review* 32(2-3), 1988, 409-414.

DE BORGER, B., K. KERSTENS, W. MOESEN, J. VANNESTE, "Explaining Differences in Productive Efficiency: An Application to Belgian Municipalities", *Public Choice* 80(2), 1994, 339-358.

DEPRINS, D., L. SIMAR, H. TULKENS, "Measuring Labor-Efficiency in Post Offices", in: M. MARCHAND, e.a. (eds), *The Performance of Public Enterprises: Concepts and Measurement*, Elsevier, Amsterdam, 1984, 243-267.

EPSTEIN, M., J. HENDERSON, "Data Envelopment Analysis for Managerial Control and Analysis", *Decision Sciences* 20(1), 1989, 90-119.

FÄRE, R., S. GROSSKOPF, J. LOGAN, "The Relative Performance of Publicly-Owned and Privately-Owned Electric Utilities", *Journal of Public Economics* 26(1), 1985, 89-106.

FÄRE, R., S. GROSSKOPF, K. LOVELL, *The Measurement of Efficiency of Production*, Kluwer, Boston, 1985.

FÄRE, R., S. GROSSKOPF, K. LOVELL, C. PASURKA, "Multilateral Productivity Comparisons when Some Outputs are Undesirable: A Nonparametric Approach", *Review of Economics and Statistics* 71(1), 1989, 90-98.

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

FØRSUND, F., K. LOVELL, P. SCHMIDT, "A Survey of Frontier Production Functions and of their Relationship to Efficiency Measurement", *Journal of Econometrics* 13(1), 1980, 5-25.

GROSSKOPF, S., "The Role of the Reference Technology in Measuring Productive Efficiency", *Economic Journal* 96(382), 1986, 499-513.

MC FADDEN, D., "Cost, Revenue, and Profit Functions", in: M. FUSS, D. MC FADDEN (eds), *Production Economics: A Dual Approach to Theory and Applications*, North Holland Publishing Company, Amsterdam, 1978, 3-109.

PESTIEAU, P., H. TULKENS, "Assessing the Performance of Public Sector Activities: Some Recent Evidence from the Productive Efficiency Viewpoint", UCL (CORE Discussion Paper 9060), Louvain-la-Neuve, 1990.

PETERSEN, N., "Data Envelopment Analysis on a Relaxed Set of Assumptions", *Management Science* 36(3), 1990, 305-314.

RUSSELL, R., "On the Axiomatic Approach to the Measurement of Technical Efficiency", in: W. EICHHORN (ed), *Measurement in Economics*, Physica-Verlag, Heidelberg, 1988, 207-217.

SEIFORD, L., R. THRALL, "Recent Developments in DEA: The Mathematical Programming Approach to Frontier Analysis", *Journal of Econometrics* 46(1-2), 1990, 7-38.
SCHMIDT, P., "Frontier Production Functions", *Econometric Reviews* 4(2), 1986, 289-328.
THIRY, B., H. TULKENS, "Productivity, Efficiency and Technical Progress: Concepts and Measurement", *Annales de l'économie publique, sociale et cooperative* 60(1), 1989, 9-42.
THIRY, B., H. TULKENS, "Allowing for Inefficiency in Parametric Estimation of Production Functions for Urban Transit Firms", *Journal of Productivity Analysis* 3(1-2), 1992, 45-65.
TULKENS, H., "The performance approach in public sector economics: An introduction and an example", *Annales de l'économie publique, sociale et cooperative* 57(4), 1986, 429-443.
TULKENS, H., B. THIRY, A. PALM, "Mesure de l'Efficacité Productive: Methodologies et Applications aux Sociétés de Transports Intercommunaux de Liège, Charlerloi et Verviers", in: B. THIRY, H. TULKENS (eds), *La Performance Economique des Sociétés Belges de Transports Urbains*, CIRIEC, Charlerloi, 1988, 81-136.
TULKENS, H., "Non-Parametric Efficiency Analyses in Four Service Activities: Retail Banking, Municipalities, Courts and Urban Transit", UCL (CORE Discussion Paper 9050), Louvain-la-Neuve, 1990.
TULKENS, H., "On FDH Efficiency Analysis: Some Methodological Issues and Applications to Retail Banking, Courts, and Urban Transit", *Journal of Productivity Analysis* 4(1-2), 1993, 183-210.
TULKENS, H., P. VANDEN EECKAUT, "Non-Parametric Efficiency, Progress and Regress Measures for Panel Data: Methodological Aspects", UCL (CORE Discussion Paper 9316), Louvain-la-Neuve, 1993.
VANDEN EECKAUT, P., H. TULKENS, "Une mesure de l'efficacité-cout de 253 communes francophones", in: H. TULKENS (ed), *L'Efficacité économique (Huitième Congrès des Economistes Belges de Langue Française: Commission 5: Efficacité et management)*, CIFOP, Charlerloi, 1988, 177-207.
VANDEN EECKAUT, P., H. TULKENS, M. JAMAR, "Cost Efficiency in Belgian Municipalities", in: H. FRIED, K. LOVELL, S. SCHMIDT (eds), *The Measurement of Productive Efficiency: Techniques and Applications*, Oxford University Press, Oxford, 1993, 300-334.
VARIAN, H., "The Nonparametric Approach to Production Analysis", *Econometrica* 52(3), 1984, 579-597.

SUMMARY

The purpose of this paper is to investigate the sensitivity of a recently proposed non-parametric approach to technical efficiency measurement. Using a data set of Belgian municipalities, we analyze the sensitivity of the Free Disposal Hull (FDH) approach with respect to the number of input and output dimensions and with respect to sample size, and we consider the impact of outliers on efficiency scores. We finally investigate the effects of using a variety of alternative (radial) efficiency measures.

RESUME

Notre objectif est d'examiner la sensibilité de l'approche non paramétrique, dite "Free Disposal Hull" (FDH), de la mesure de l'efficacité technique. Nous utilisons pour cela une base de données des communes belges. Nous considérons la sensibilité de la mesure d'efficacité suite à un changement du nombre d'inputs et d'outputs, à une variation de la taille de l'échantillon ainsi qu'à la présence de données aberrantes. Pour conclure, nous examinons l'influence de ce choix en utilisant plusieurs mesures d'efficacité (radiale).

ZUSAMMENFASSUNG

Unser Ziel ist es, die Sensibilität einer neuen nicht parametrischen Methode (Free Disposal Hull [FDH]) zur Messung der technischen Effizienz zu untersuchen. Dazu verwenden wir eine Datenbank der belgischen Gemeinden und untersuchen dabei das Verhalten der Effizienzwerte bei einer Änderung der Input- und Outputanzahl, einer Veränderung der Mustergrösse sowie bei einem Einsatz von abweichenden Daten (*outliers*). Abschliessend untersuchen wir die Wirkung beim Einsatz von verschiedenen Effizienzmessungen.