UNIVERSITY OF OSLO

HEALTH ECONOMICS RESEARCH PROGRAMME

Far out or alone in the crowd:

Classification of self-evaluators in DEA

Dag Fjeld Edvardsen
The Norwegian Building Research Institute

Finn R. Førsund
Department of Economics University of Oslo/The Frisch Centre

Sverre A.C. Kittelsen
The Frisch Centre

Working Paper 2003: 7

Far out or alone in the crowd:

Classification of self-evaluators in DEA*

by

Dag Fjeld Edvardsen[1], Finn R. Førsund[2] and Sverre AC. Kittelsen[3]

April, 2003

Health Economics Research programme at the University of Oslo

HERO 2003

JEL classification: C44, C61, D24, I19, L32

Key words: Self-evaluator, interior and exterior self-evaluator, DEA, efficiency, referencing zone, nursing homes

* The paper is based on results from the NFR-financed Frisch Centre project "Better and Cheaper?" and written within the project "Efficiency analyses of the nursing and home care sector of Norway" at the Health Economics Research Programme at the University of Oslo (HERO) and the Frisch Centre.

[1] The Norwegian Building Research Institute, Norway
[2] Department of Economics University of Oslo/The Frisch Centre, University of Oslo
[3] The Frisch Centre, Norway

Correspondence to:
Finn R. Førsund, Department of Economics, University of Oslo, P O Box 1095 Blindern, N-0317 Oslo, Norway
e-mail: finn.forsund@econ.uio.no

Health Economics Research programme at the University of Oslo
Financial support from The Research Council of Norway is acknowledged.
ISSN 1501-9071, ISBN 82-7756-110-5

Abstract

The units found strongly efficient in DEA studies on efficiency can be divided into self-evaluators and active peers, depending on whether the peers are referencing any inefficient units or not. The contribution of the paper starts with subdividing the self-evaluators into interior and exterior ones. The exterior self-evaluators are efficient “by default”; there is no firm evidence from observations for the classification. These units should therefore not been regarded as efficient, and be removed from the observations on efficiency scores when performing a two-stage analysis of explaining the distribution of the scores. A method for classifying self-evaluators based on the additive DEA model is developed. The application to municipal nursing- and home care services of Norway shows significant effects of removing exterior self-evaluators from the data when doing a two-stage analysis.

1. Introduction

The calculation of efficiency scores for production units based on a non-parametric piecewise linear frontier production function, is well established within the last two decades. Originally introduced by Farrell (1957) the method was further developed in Charnes, Cooper and Rhodes (1978), where the term the *DEA model* was coined. The efficient units span the frontier, but the classification of some of these units as efficient is not based on other observations being similar, but due to the method. We are referring to units, which are classified as being *self-evaluators* in the literature (Cooper, Seiford and Tone, 2000). Self-evaluators may most naturally appear at the "edges" of the technology, but it is also possible that self-evaluators appear in the interior. It may be of importance to distinguish between the self-evaluators being *exterior* or *interior*. Finding the influence of some variables on the level of efficiency by running regressions of efficiency scores on a set of potential explanatory variables, is an approach often followed in actual investigations.[1] Using exterior self-evaluators with efficiency score of 1 may then distort the results, because to assign the value of 1 to these self-evaluators is arbitrary. But regarding interior self-evaluators they may have peers that are fairly similar. They should then not necessarily be dropped when applying the two- stage approach.

The plan of the paper is to review the DEA models in Section 2 and define the new concepts of interior and exterior self-evaluators. In Section 3 the method for classifying the self-evaluators is introduced. Actual data are presented in Section 4 and the method for classifying self-evaluators is applied. The effect of removing exterior self-evaluators is tested. Section 5 concludes.

2. Self-evaluators

[1] The approach was originally introduced in Seitz (1967), inspired by Nerlove (1965), see Førsund and Sarafoglou (2002). Simar and Wilson (2003) review the approach and find it at fault in general due to serial correlation between the efficiency scores, and provides a new statistically sound procedure based on specifying explicitly the data generating process and bootstrapping to obtain confidence intervals.

DEA models

Consider a set, J, of production units transforming multiple inputs into multiple outputs. Let y_{mj} be an output $(m \in M, j \in J)$ and x_{nj} an input $(n \in N, j \in J)$. As the reference for the units in efficiency analyses we want to calculate a piecewise linear frontier based on observations, fitting as closely as possible and obeying some fundamental assumptions, like free disposal, and the technology set being convex and closed as usually entertained (Banker et al., 1984, Färe and Primont, 1995). This frontier can be found by solving the following LP problem, termed the *additive model* in the DEA literature (Charnes et al., 1985):

$$
\begin{aligned}
&Max \quad \left\{ \sum_{m \in M} s_{mi}^{+} + \sum_{n \in N} s_{ni}^{-} \right\} \\
&s.t. \\
&\sum_{j \in J} \lambda_{ij} y_{mj} - y_{mi} - s_{mi}^{+} = 0 \quad , m \in M \\
&x_{ni} - \sum_{j \in J} \lambda_{ij} x_{nj} - s_{ni}^{-} = 0 \quad , \ n \in N \\
&s_{mi}^{+} \ , \ s_{ni}^{-} \geq 0 \\
&\lambda_{ij} \geq 0 \\
&\sum_{j \in J} \lambda_{ij} = 1
\end{aligned}
\tag{1}
$$

The last equality constraint in (1) imposes variable returns to scale (VRS) on the frontier, while dropping this constraint imposes constant returns to scale (CRS). Our analysis will be valid for both scale assumptions. The frontier is found by maximising the sum of the slacks on the output constraints, s_{mi}^{+}, and input constraints, s_{ni}^{-}. The *strongly efficient* units (using the terminology of Charnes et al., 1986) are identified by all the slack variables being zero. All weights, λ_{ij}, must be zero except the weight for itself that will be one (i.e. $\lambda_{ij} = 0 \ for \ i \neq j, \lambda_{ii} = 1$ if i is an efficient unit).[2] The efficient points will appear as vertex points on the frontier function surface, or corner points of facets. The sets of strongly efficient units, P, and the inefficient units, I, are:

$$
\begin{aligned}
&P = \left\{ i \in J : s_{mi}^{+}, s_{ni}^{-} = 0 \, \forall m \in M, \forall n \in N \right\} \\
&I = \left\{ i \in J : s_{mi}^{+}, s_{ni}^{-} \geq 0 \, \forall m \in M, \forall n \in N \right\}, \\
&P \cup I = J
\end{aligned}
\tag{2}
$$

[2] A strongly efficient unit, i, may end up being located exactly on a facet. We may then have multiple solutions for the weights, although the maximal sum of slacks is still zero. One of the solutions will be $\lambda_{ij} = 0$ for $j \neq i$, and $\lambda_{ii} = 1$.

So far we only have slacks as measures of inefficiency. If we want only one measure for each unit, and a measure that is independent of units of measurement, the Farrell (1957) measure of technical inefficiency is the natural choice. The standard DEA model on primal (enveloping) form is set up as a problem of determining the Farrell technical efficiency score, E_{oi}, (o = 1,2), either in the input- (o = 1) or the output (o = 2) direction for an observation, *i*. The following LP model is formulated for each observation in the case of input-orientation:

$$
\begin{aligned}
& E_{1i} \equiv Min \quad \theta_i \\
& s.t. \\
& \sum_{j \in P} \lambda_{ij} y_{mj} - y_{mi} \geq 0 \quad , m \in M \\
& \theta_i x_{ni} - \sum_{j \in P} \lambda_{ij} x_{nj} \geq 0 \quad , \; n \in N \\
& \lambda_{ij} \geq 0 \\
& \sum_{j \in P} \lambda_{ij} = 1
\end{aligned}
\tag{3}
$$

In the case of output orientation we have the following LP program:

$$
\begin{aligned}
& 1/E_{2i} \equiv Max \quad \phi_i \\
& s.t. \\
& \phi_i y_{mi} - \sum_{j \in P} \lambda_{ij} y_{mj} \leq 0 \quad , m \in M \\
& \sum_{j \in P} \lambda_{ij} x_{nj} - x_{ni} \leq 0 \quad , \; n \in N \\
& \lambda_{ij} \geq 0 \\
& \sum_{j \in P} \lambda_{ij} = 1
\end{aligned}
\tag{4}
$$

For notational ease the same symbols have been used for weights in (1), (3) and (4). The proportionality factor, θ_i or ϕ_i, and the weights, λ_{ij} , are the endogenous variables.

Adopting the notation #N and #M for the number of inputs and outputs respectively, the point

$$
(\sum_{j \in P} \lambda_{ij} x_{1j}, .., \sum_{j \in P} \lambda_{ij} x_{\#Nj,} \sum_{j \in P} \lambda_{ij} y_{1j}, .., \sum_{j \in P} \lambda_{ij} y_{\#Mj}) \tag{5}
$$

is per construction on the frontier surface, and is defined as the *reference point* for unit *i*. If there are no slacks on the output- and input constraints in (3) or (4) then the reference points coincide with the radial projection point, using either θ_i or ϕ_i when adjusting an inefficient observation. These points will normally be interior points on facets (but may fall on border lines). With one or more slacks positive the reference point and the radial projection point differ. The reference points will again appear as vertex points on the frontier function surface, or corner points of facets.

It is well known that the radial Farrell efficiency measure E_{oi} may be one, but that the unit may still improve its performance by either using less inputs or producing more outputs. All units with a radial efficiency score of one are by definition located on the frontier, but it is only for the *strongly efficient* units that the reference points coincide with the observation. A unit may have $E_{oi}=1$, but one or more of the constraints in (3) or (4) being non-binding (i.e. one or more slacks positive or zero shadow prices on the constraints in question).

Although the model (3) or (4) can be solved directly by letting the index j run over all observations in J, a two-stage procedure of solving (1) first is often followed. By using the information on strongly efficient units when solving (3) or (4), the LP computations are done more efficiently, and one will only identify reference points by (5) that are in the strongly efficient subset of the frontier.

In the context of the DEA models (3) and (4) the strongly efficient units are termed *peers*. For each inefficient unit, i, a *Peer group set*, P_i, may be formed:

$$P_i = \{p \in P : \lambda_{ip} > 0\}, i \in I \tag{6}$$

If the Peer group sets are empty, then all the units are efficient. The solutions to (1), (3) or (4) do not identify facets systematically, but by using (6) we can identify the corner points of facets where one or more radial projection points of inefficient units are located.

It will also turn out useful to look at the group of inefficient units referenced by a peer. Such a set is defined for each peer, p, as the *Referencing set* in Edvardsen and Førsund (2001) with reference to the solutions of (3) or (4):

$$I_p = \{i \in I : \lambda_{ip} > 0 \}, \quad p \in P \tag{7}$$

The self evaluators

The Referencing set (7) may be empty. This is the definition of a self-evaluator:

Definition 1: A peer $p \in P$*, where the set P is defined in (2), is a self-evaluator if* $I_p = \emptyset$ *, where* I_p *is defined in (7).*

The set of peers may be partitioned into a set of self-evaluators, P^S, and a set, P^A, of *active peers,* i.e. peers with non-empty referencing sets:

$$
\begin{aligned}
&P^S = \{p \in P : I_p = \emptyset\} \\
&P^A = \{p \in P : I_p \neq \emptyset\} \\
&P^S \cup P^A = P
\end{aligned}
\tag{8}
$$

The self-evaluators are vertex points of facets without any reference points defined as the radial projection points of inefficient observations located on these facets. The LP solutions to (3) or (4) do not give us any information as to which efficient units constitute the vertex points of such a facet without reference points. An efficient unit may be a vertex point for many facets. Our definition of a self-evaluator implies that there must not be reference points on any of its facets.

There are two possibilities as to the location of facets formed by self-evaluators on the frontier surface. Such facets may be part of the extreme areas of the frontier, i.e. facets closest to the axes in the case of CRS, or facets, in case of VRS, also most far away from the origin or closest to the origin (the VRS frontier will in general not contain the origin). In the case of CRS only mixes of inputs or outputs may be extreme, while in the case of VRS we in addition have the scale dimension. Such self-evaluators will be termed *exterior* self-evaluators. In the case of CRS facets without any reference points may also be found in the interior of the frontier surface with respect to mixes, while for VRS interior also means interior regarding scale. Such self-evaluators will be termed *interior* self-evaluators.

Figure 1 shows the two different cases in the simplest case of two dimensions. The points *A, B, C, D, F* and *G* are efficient, while O_1 is inefficient. The radial reference or projection point for unit O_1 is *a* in the case of input orientation. The reference point (5) in this simple case coincides with the peer *A*. Considering output-orientation the peers are *D* and *F*, and the reference point is *d*. To illustrate the referencing set of a peer Figure 1 shows the *referencing zone* for the efficient unit *D* in the case of output orientation. All the inefficient units being in unit *D*'s referencing set must be located here (such inefficient units may also appear in referencing sets of other peers; here unit *F*'s). If the referencing zone is empty then the peer is a self-evaluator. Removal of such a unit will not change any of the efficiency scores for other units. We would expect the self-evaluators to be extreme points in one or more of the mix or scale dimensions, but if the referencing zone is narrow a self-evaluator may also be centrally placed within the set of observations. A narrow zone means that other peers are close to the self-evaluator.

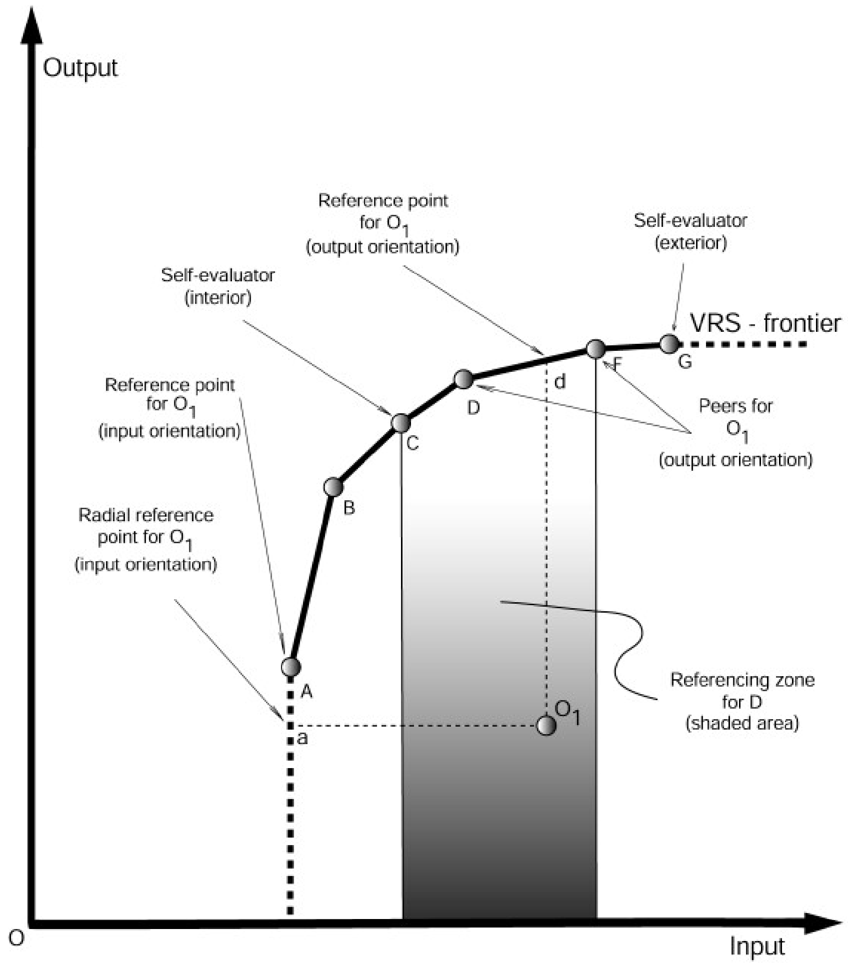

Figure 1: DEA and the two types of self evaluators

Notice that the classification as a self-evaluator is dependent of the orientation of the efficiency measure. Considering output orientation we have that both *B* and *C* are interior self-evaluators, while *A* and *G* are exterior self-evaluators. Considering input orientation we have that *B*, *C*, *D* and *F* are interior self-evaluators, while *G* is an exterior one.

3. The determination of type of self-evaluator

Our purpose is to develop a method for classification into exterior or interior self-evaluators only using the standard DEA format.

Enveloping from below

The production set is by construction convex. If all inefficient units are removed from the data set, and a new run is done with only the efficient units, we will find the exterior peers by reversing the enveloping of the data from “above” to be from “below”. All that needs to be done is to reverse the inequalities in the LP program (1) by adding the slack variables instead of subtracting:

$$
\begin{aligned}
& Max \quad \left\{ \sum_{m\in M} s_{mi}^{+} + \sum_{n\in N} s_{ni}^{-} \right\} \quad (i\in P) \\
& s.t. \\
& \sum_{j\in P} \lambda_{ij} y_{mj} - y_{mi} + s_{mi}^{+} = 0 \quad , m\in M \\
& x_{ni} - \sum_{j\in P} \lambda_{ij} x_{nj} + s_{ni}^{-} = 0 \quad , \; n\in N \\
& s_{mi}^{+} \;,\; s_{ni}^{-} \geq 0 \quad , m\in M \;,\; n\in N \\
& \lambda_{ij} \geq 0 \\
& \sum_{j\in P} \lambda_{ij} = 1
\end{aligned}
\tag{9}
$$

Notice that we are only considering observations belonging to the set of strongly efficient units determined by solving (1). This envelopment of the data is by construction concave.

The units that turn out as “efficient” in solving (9), in the sense that all slacks are zero, must be units belonging to the exterior facets in the solution to the original model (1). We will use this result to define exterior and interior strongly efficient units:

Definition 2: A strongly efficient unit belonging to the set P defined by (2) is exterior if it belongs to the set P^E:

$$P^E = \left\{ p\in P : s_{mp}^{+}, s_{np}^{-} = 0 \right\} \tag{10}$$

where the slack variables, s^+_{mp}, s^-_{np} *, are solutions to the problem (9).*

A strongly efficient unit belonging to the set P defined by (2) is interior if it belongs to the set

$$P^I = \left\{ p \in P : p \notin P^E \right\} \qquad (11)$$
$$(P^E \cup P^I = P)$$

where the set P^E *is defined in (10).*

To determine the nature of a self-evaluator an orientation for the calculation of the Farrell efficiency measures has to be chosen, i.e. either input- or output orientation. The following definition can then be made as to the classification of self-evaluators:

Definition 3: Consider a peer $p \in P$*, where the set P is defined in (2), that is a self-evaluator,* $p \in P^S$*, where the set* P^S *is defined in (8) and found by running either the input-oriented program (3), or the output-oriented program (4). If* $p \in P^E$*, where the set* P^E *is defined in (10), then p is an exterior self-evaluator. If* $p \notin P^E$ *then p is an interior self-evaluator:*

$$P^{SE} = \left\{ p \in P^S : p \in P^E \right\}, \qquad (12)$$
$$P^{SI} = \left\{ p \in P^S : p \notin P^E \right\} (P^{SE} \cup P^{SI} = P^S)$$

where P^{SE} *and* P^{SI} *are the sets of the exterior and interior self-evaluators respectively.*

Illustrating the approach using Figure 1, we have that the new "from below frontier" will be the line from *A* to *G*, thus these units are the only ones on the "from below frontier" and therefore exterior points in P^E. This classification is independent of orientation, and they are both being located on exterior facets in the original problem (1). In the case of output orientation the self-evaluators *B* and *D*, according to solving problem (4), will not appear on the new frontier, and they are therefore interior according to Definition 3. The self-evaluator *G* appears on the new frontier and is therefore exterior. In the case of input orientation solving problem (4) gives *B, C, D, F* and *G* as self-evaluators, and we have that *B, C, D, and F* are interior self-evaluators and *G* an exterior one.

Figure 2 provides another illustration. In a two-dimensional input space an isoquant is shown in the efficient units *A, B, C* and *D*. Consider input orientation and CRS. Assuming inefficient units are only located northeast of the isoquant segment *AB* in the cone delimited

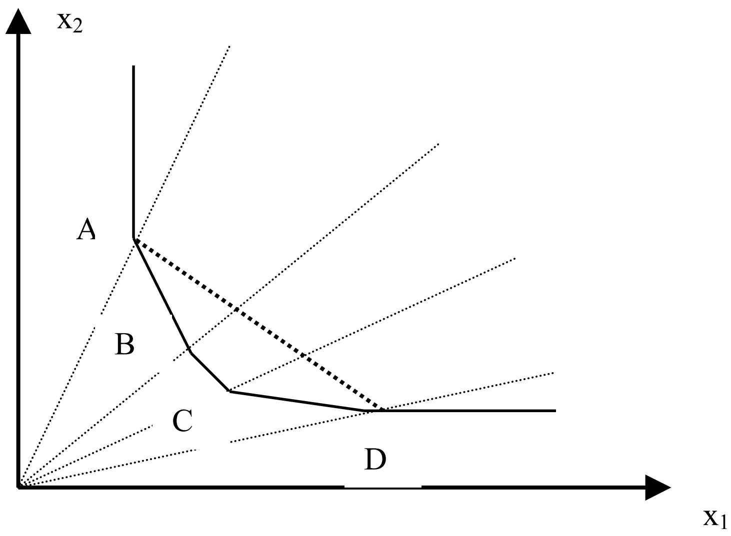

Figure 2. Determining the type of self-evaluator

by the rays going through the points *A* and *B*, we have that *C* is an interior self-evaluator, and *D* is an exterior self-evaluator. Running the “reverse” program (9) we will envelope the four peers from “behind” by the broken line from *A* to *D*. We then know that units *A* and *D* are exterior, and using the information from running the DEA model (1) we then have that unit *C* is an interior self evaluator, and unit *D* an exterior one.

It may also be of interest to classify the active peers according to the type exterior and interior. Building on definition 3 we have.

Definition 4. The active peers defined in (8) belong to the subsets P^{AE} and P^{AI}:

$$P^{AE} = \left\{ p \in P^A : p \in P^E \right\},$$
$$P^{AI} = \left\{ p \in P^A : p \in P^I \right\} \; (P^{AE} \cup P^{AI} = P^A) \qquad (13)$$

where P^E and P^I are defined in (10) and (11) respectively.

The program (9) is not the standard DEA additive formulation, since the sign of the slacks in the restrictions on inputs and outputs have been changed. However, by negating these restrictions, (9) can be rewritten as:

$$
\begin{aligned}
&Max \quad \left\{ \sum_{m\in M} s_{mi}^{-} + \sum_{n\in N} s_{ni}^{+} \right\} \quad (i\in P) \\
&s.t. \\
&\sum_{j\in P} \lambda_{ij} x_{nj} - x_{ni} - s_{ni}^{+} = 0 \quad , n\in N \\
&y_{mi} - \sum_{j\in P} \lambda_{ij} y_{mj} - s_{mi}^{-} = 0 \quad , \; m\in M \\
&s_{mi}^{+} \;,\; s_{ni}^{-} \geq 0 \quad , m\in M \;,\; n\in N \\
&\lambda_{ij} \geq 0 \\
&\sum_{j\in P} \lambda_{ij} = 1
\end{aligned} \tag{14}
$$

Comparing (1) and (14) we see that these are identical except that inputs and outputs are exchanged. Since existing DEA software often will solve the additive model (1), we may as well for convenience find the set of exterior self-evaluators P^{SE} by exchanging inputs and outputs and running (14) on the strongly efficient units, rather than running (9) on these units.

4. An empirical application

The data

We will apply the method for determining interior and exterior self-evaluators on a cross section data set of the nursing and home care sector of Norwegian municipalities. The data is found in Edvardsen et al. (2000). The primary data source is the official yearly statistics for municipal activities published by Statistics Norway. Resource usage is measured by financial data and number of man-years of different categories. Production data contains mainly the number of clients dealt with by institutionalised nursing, home based nursing, and practical assistance. Quality information is lacking, but the clients are split on some age groups that may be of significance for resource use. In cooperation with representatives form the municipalities and the ministries of Finance, Municipal and regional affairs, and Social and health affairs we have chosen to split the clients on two major age groups, 0-66 and above 66 (67+), and use institutions and home care as separate outputs. Within institutions there are also a number of short-stay clients, either coming on a day care basis or on limited stay of convalescence. These usually require fewer resources than the permanent clients. As indicators of quality of institutions we have information of number of single person rooms and on clients staying in closed wards. The separation is regarded both as a quality factor for the clients taken care of (demented cases), and for the other clients. In home-based care mentally disabled may be quite resource demanding. They may also be found in the 0-66 age group within institutions. There is no information on how long time a home visit may last or how often it is received. Such information would obviously have given us some quality indicators. We also run the risk of municipalities cutting down on both length and number of visits showing the same number of clients receiving a more generous support in other municipalities.

To ensure that the data quality was good enough we entered a phase of quality control. We strongly feel that one should not automatically remove outliers, but if possible connect with the municipality in question and ask if the data is correct. This is especially important if the methodology is frontier based (such as DEA) because the units defining the frontier are outliers by definition. This led to many changes in the dataset and required quit a lot of work, but as a result we could be much more confident in the quality of the data (see Aas (2000) for details).

Table 1: Primary variables used in the DEA model, cross-section 1997 of 469 municipalities.

		Average	Standard deviation	Min	Max
Inputs					
Trained Nurses	x_1	31.1	41.4	1.5	410.4
Other Employees	x_2	137.4	169.4	5.3	1821.5
Other expenses	x_3	9066.2	13449.5	190.0	108990.0
Outputs: No. of Clients					
Institutions, age 0-66	y_1	3.4	4.9	0.0	50.0
Institutions, age 67+	y_2	87.7	108.6	0.0	1024.0
Short-term stay	y_3	113.8	163.3	0.0	1614.0
Closed wards	y_4	11.8	19.3	0.0	195.0
Single person room	y_5	65.7	82.2	0.0	747.0
Mentally disabled	y_6	48.7	79.5	0.0	857.0
Practical assistance, 0-66	y_7	51.3	66.3	0.0	597.0
Practical assistance, 67+	y_8	212.7	272.4	1.0	2190.0
Home based nursing, 0-66	y_9	34.1	45.3	0.0	407.0
Home based nursing, 67+	y_{10}	125.8	153.3	1.0	1480.0

Table 1 shows descriptive statistics for the variables used in the DEA model. The first three rows measure the inputs in the model. *Trained nurses* and *Other Employees* shows us that about 18% of the employees (measured in man-years) are trained nurses. *Other expenses* are measured in 1000 NOK (Norwegian currency). The last 9 rows in table 1 measure the outputs. *Inst0-66* and *Inst67+* are the number of institutionalized clients in the age groups 0-66 and above 67 respectively. *Short-stay* shows how many visits the institutions in the municipality have gotten from clients who are not residents, while *Clients in closed ward* shows how many of the residents are in a special ward for dementia clients. *Mentally disabled* shows how many of the clients are mentally disabled (almost all of these clients get home care). *Practical assistance 0-66* and *Practical assistance 67+* counts how many clients get practical assistance (such as cleaning and making food) in the indicated age groups, while *Home based nursing 0-66* and *Home based nursing 67+* count the same for clients getting nursing services in their own homes.

The Farrell output-oriented efficiency scores

Figure 3 shows E_2 (output-increasing efficiency assuming variable returns to scale). Each bar in the diagram represents one of the 469 municipalities, sorted by increasing efficiency. The

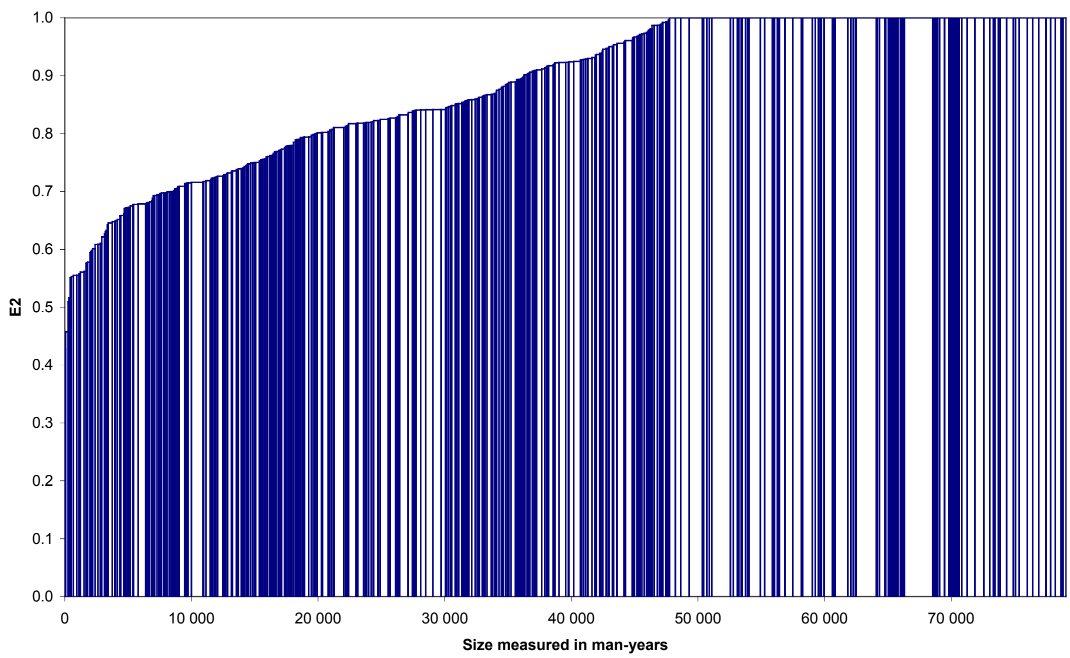

Figure 3: Sorted output-oriented efficiency scores

height of the bars represents the efficiency of the DMU, while the width of the bar shows the size measured by man-years (sum of trained nurses and other employees). Both large and small DMUs can be found in all parts of the diagram, with the exception that no large municipality is located in the (very inefficient) leftmost part of the diagram. The average effiency is 86%, while the efficiency of the average unit is 67% .

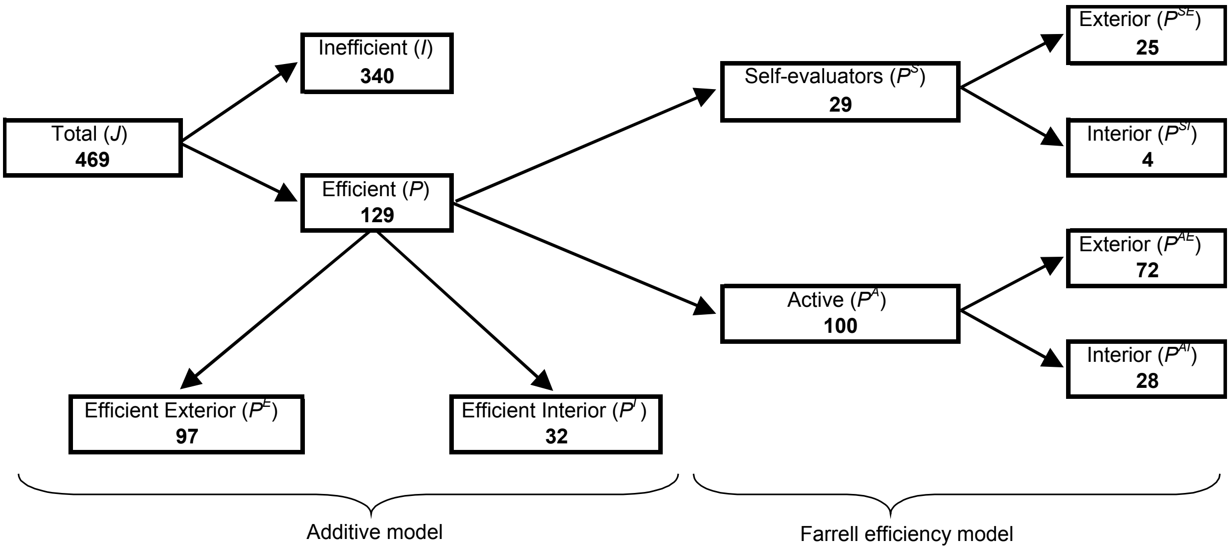

Figure 4: The taxonomy of units in DEA efficiency analyses

An overview of the taxonomy developed in Sections 2 and 3 for classification of units is given in Figure 4, together with the actual decomposition for the data set at hand. In view of the relatively large number of observations it may be surprising that as many as 28 percent of the units are efficient. This may be due to the unusually high number of dimensions, 13 variables in all. Since the efficient units span out the frontier technology it is to be expected that the number of exterior ones is higher than the interior ones, 75 and 25 percent respectively. Turning to the Farrell efficiency model (4) the self-evaluators represent 23 percent of the efficient units. As expected the relative share of exterior peers is larger in the group of self-evaluators than in the group of active peers, 86 versus 72 percent. Among the active peers that share of interior units is higher, 28 percent. This distribution is of importance for the empirical support of the frontier and the associated efficiency distribution.

Far out or alone in the crowd?

In Table 3 the relative distance from the average unit is obtained by measuring each of its variables against the average for the sample (J). The interior units are on both sides of the average, and one of the four units is quite close to the sample average. None is close to either the small or large exterior units. It seems to be appropriate to use the expression “alone in the crowd.”

Table 3. Relative size in percent of interior- and exterior self-evaluators compared with the sample average

Self-evaluators				***Inputs***			***Outputs***								
Interior	***Exterior***	***Municipality***		x_1	x_2	x_3	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9
Indre SE	**Ytre SE**	KomNr	KomNavn	Sykepleiere	Andre årsvei	Andr.□dr.utg.	inst □0-66	inst 67+□(y:	korttidsopph	Skjermet a\	enerom	PU	PB 0-66	PB 67+	HS 0-66
TRUE	FALSE	425	Åsnes	6.17	3.66	-95.22	-0.40	59.34	17.19	7.20	16.35	3.30	-1.27	58.31	1.89
TRUE	FALSE	616	Nes	-17.43	-86.14	-5285.22	-1.40	-38.66	-42.81	-5.80	-16.65	-34.70	-20.27	-75.69	-13.11
TRUE	FALSE	807	Notodden	4.27	49.86	1951.78	0.60	72.34	190.19	6.20	69.35	7.30	-17.27	89.31	5.89
TRUE	FALSE	1567	Rindal	-15.93	-93.34	-5790.22	-2.40	-35.66	-36.81	-4.80	-29.65	-40.70	-37.27	-115.69	-28.11
FALSE	TRUE	101	Halden	39.37	266.76	5644.78	0.60	28.34	-36.81	-11.80	9.35	96.30	85.73	679.31	65.89
FALSE	TRUE	213	Ski	19.47	59.66	6022.78	10.60	-10.66	-43.81	8.20	7.35	26.30	131.73	124.31	-10.11
FALSE	TRUE	217	Oppegård	28.67	8.16	2597.78	2.60	-7.66	25.19	-11.80	18.35	20.30	7.73	92.31	132.89
FALSE	TRUE	219	Bærum	289.27	1011.56	99923.78	28.10	655.84	627.19	131.20	681.35	380.30	337.73	1145.31	196.89
FALSE	TRUE	430	Stor-Elvdal	-21.03	-65.64	-4466.22	-3.40	-45.66	-55.81	-0.80	-32.65	-37.70	-34.27	-79.69	8.89
FALSE	TRUE	615	Flå	-27.93	-100.44	-6653.22	-3.40	-62.66	-87.81	-11.80	-45.65	-43.70	-43.27	-172.69	-33.11
FALSE	TRUE	632	Rollag	-25.33	-99.44	-6738.22	-2.40	-60.66	-100.81	-11.80	-41.65	-31.70	-41.27	-169.69	-8.11
FALSE	TRUE	709	Larvik	107.27	406.16	18544.78	17.60	248.34	345.19	32.20	158.35	191.30	182.73	829.31	57.89
FALSE	TRUE	806	Skien	109.87	488.86	31342.78	8.60	260.34	106.19	2.20	199.35	114.30	178.73	1171.31	109.89
FALSE	TRUE	904	Grimstad	28.57	29.16	-174.22	-0.40	23.34	-25.81	22.20	-4.65	26.30	81.73	93.31	76.89
FALSE	TRUE	941	Bykle	-26.73	-104.84	-7010.22	-3.40	-63.66	-92.81	-6.80	-41.65	-47.70	-48.27	-204.69	-31.11
FALSE	TRUE	1144	Kvitsøy	-26.53	-130.24	-8496.22	-3.40	-83.66	-100.81	-11.80	-59.65	-48.70	-48.27	-200.69	-32.11
FALSE	TRUE	1222	Fitjar	-18.93	-87.74	-6606.22	-2.40	-61.66	-65.81	-3.80	-17.65	-37.70	-38.27	-154.69	-26.11
FALSE	TRUE	1411	Gulen	-21.23	-67.24	-5174.22	-1.40	-20.66	-95.81	-11.80	-37.65	-36.70	-43.27	-124.69	-22.11
FALSE	TRUE	1612	Hemne	-15.73	-78.54	-5231.22	-2.40	-56.66	-17.81	-0.80	-53.65	-42.70	-25.27	-86.69	12.89
FALSE	TRUE	1632	Roan	-27.53	-112.64	-7807.22	-3.40	-71.66	-87.81	-11.80	-46.65	-40.70	-48.27	-167.69	-31.11
FALSE	TRUE	1702	Steinkjer	62.57	159.56	2438.78	4.60	74.34	99.19	42.20	44.35	29.30	80.73	365.31	42.89
FALSE	TRUE	1714	Stjørdal	39.77	132.26	2631.78	16.60	85.34	112.19	4.20	84.35	74.30	64.73	128.31	0.89
FALSE	TRUE	1723	Mosvik	-22.33	-120.04	-8073.22	-3.40	-66.66	-92.81	-11.80	-49.65	-37.70	-45.27	-180.69	-30.11
FALSE	TRUE	1839	Beiarn	-24.13	-103.14	-6688.22	-1.40	-65.66	-85.81	-5.80	-40.65	-43.70	-46.27	-150.69	-26.11
FALSE	TRUE	1868	Øksnes	-14.83	-33.94	-4360.22	-0.40	-32.66	-64.81	-4.80	-12.65	-20.70	-20.27	-119.69	48.89
FALSE	TRUE	1920	Lavangen	-27.13	-112.64	-7495.22	-2.40	-67.66	-93.81	-7.80	-47.65	-38.70	-49.27	-180.69	-33.11
FALSE	TRUE	3001	Bygdøy-Fr(	12.67	14.36	13587.78	-3.40	-32.66	-113.81	-11.80	-10.65	-15.70	7.73	412.31	33.89
FALSE	TRUE	3003	St.Hanshau	81.47	372.26	63035.38	1.60	343.34	190.19	26.20	289.35	16.30	18.73	738.31	-3.11
FALSE	TRUE	3004	Sagene-To	75.17	466.56	73227.48	10.60	267.34	384.19	20.20	165.35	11.30	190.73	1178.31	59.89

The exterior units are distributed with half above and half below the sample average. One unit has maximal sample values for two of the variables. There are several output variables with zero as the lower limit. The variable “Institution 0-66” has seven exterior units with the minimum value of zero; while for “Closed ward” there are eight exterior units with the minimum value of zero. So given that “far out “ means both small and large units the exterior units deserve well this classification. The influence of extreme mixes may also be investigated, but due to all the possible comparison we leave this exercise out.

The idea behind the two-stage approach is based on the distinction between discretionary and non-discretionary variables of the unit in question. The latter group of variables is also called environmental variables. When modeling the production process only variables that are under the control of the management may be considered relevant. However, environmental variables may be relevant for the performance of the units, but their influence may be regarded to be of a nature that is most appropriately revealed by studying the connection between some measure of performance and the environmental variables. Since the crucial point of being concerned with environmental variables is that there must be some influence on either the discretionary inputs or outputs of the environmental variables there is a good case for advocating a single stage approach and incorporating all relevant variables in one single model. One reason for treating environmental variables differently than standard outputs and inputs is that the way the variables interact with the standard production variables may be difficult to model. It may not be so clear-cut whether the variable is an input or an output.

The formulation of the second stage is to establish a connection between the efficiency score and the environmental variables, z_k:

$$E_{oi} = f(z_1,..,z_K) + \varepsilon_i \; , i \in J, o = 1,2 \tag{15}$$

where ε_i is a random variable. There have been several approaches to estimating (15). The first approach was to specify $f(.)$ as a linear function and apply OLS (Seitz 1967, 1971). But there are two special features of the model (15). By definition the efficiency scores are restricted to be between zero and one,

$$0 \le E_{oi} = f(z_1,..,z_K) + \varepsilon_i \le 1 \; , i \in J, o = 1,2 \tag{16}$$

and using the DEA model (3) or (4) to generate the efficiency scores usually leads to a concentration of the values 1. As shown in Figure 4 we have 28 percent of the efficiency scores being at the upper limit of one. This has lead researchers to apply a censored regression like the Tobit model or truncated regressions. These approaches are strongly criticized in Simar and Wilson (2003). The fundamental point is made that the efficiency scores in (15) are estimates of the unknown efficiencies, and that these scores are serially correlated. Therefore, neither applying a Tobit or a truncated regression will solve this problem. A sequence of bootstrapping techniques is proposed that will yield proper confidence intervals of the parameters of $f(.)$.

However, since the purpose of our paper is to demonstrate the importance of the role of exterior peers, we adopt the “easy way out” and use OLS to estimate a linear function using (15). The relation (15) is then interpreted just to represent an investigation of association and not to be a strict causality model. An advantage of OLS is that better diagnostics to characterize the covariations are available, like the multiple regression coefficient.

Table 3: Stage 2 regression results applying OLS to a linear model

	All units included*)		Excluding outer SE		Excluding all SE	
	R2	0.1737	R2	0.2082	R2	0.1713
Variable	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
Climate indicator	-0.007	0.035	-0.006	0.056	-0.007	0.032
Share of private institutions	0.098	0.019	0.099	0.015	0.101	0.016
Free disposable income, 1996	-0.020	0.054	-0.028	0.01	-0.019	0.074
Share of users in home care	-1.019	0	-1.089	0	-1.000	0
Share in home care of age group 0-66	1.823	0.17	1.437	0.282	1.814	0.172
Share in home care of age group 67-79	0.574	0.006	0.665	0.001	0.549	0.009
Share in home care of age group 80-89	0.270	0.004	0.261	0.004	0.271	0.004
Share in home care of age group 90+	0.100	0.019	0.109	0.011	0.094	0.031
Share in inst. care of age group 0-66	24.785	0.019	27.615	0.009	24.957	0.018
Share in inst. care of age group 97-79	-1.072	0.011	-0.926	0.026	-1.060	0.012
Share in inst. care of age group 80-89	-0.101	0.524	-0.152	0.331	-0.097	0.542
Share in inst. care of age group 90+	0.026	0.561	0.053	0.235	0.023	0.613
Constant term	1.527	0	1.562	0	1.516	0

*) Communities within the two major cities Bergen and Oslo are aggregated and one unit is removed from the data set

Table 3 shows the result of an OLS regression using a linear model in (15). The p-values are also given, although they should not be taken at face value due to the inherent statistical problems with the approach, as mentioned above. We perform regressions with the complete data set, excluding the exterior self-evaluators, and also excluding the interior self-evaluators. The last regression is done to compare the effects with the second.

The environmental variables represent non-discretionary background variables that experts have suggested may influence the efficiencies of municipalities. *Climate indicator* is a measure of the average temperature measured over the year in the municipality. It can also be seen as a proxy for amount of snow, altitude and distance from the coast. We note that it does not matter much to remove the interior self-evaluators, but removing the exterior ones changes both the regression coefficient and the p-value, indicating a weaker connection between efficiency scores and this variable.

Share of private institutions is measured by how large share of the total number of institutions are in the private sector (most often NGO's). It would be better to measure this by the number of clients, but such data was not available. Possible interpretations of a positive (and low p values in all three regression models) parameter estimate are that the municipalities own care providers get a learning effect from presence of private service providers, or that private presence reduces inefficiency because they increase the fear of privatization in the municipal nursing sector.

Free disposable Income, 1996 is measure of the relative wealth of the municipality (per inhabitant). It is calculated by finding the difference between the actual income in the municipality, and the "required expenses" in the municipality in other sectors than care for the elderly (i.e. schools, roads etc.). Required expenses are calculated on demographical variables and other factors exogenous to the municipality. (See Aaberge and Langørgen (2003) for the details behind the construction of this indicator.) Data for 1997 (the year all the other data is from) was also available, but we reasoned that the municipality's decision on how it want provide care for the elderly is more strongly based on income in the previous than in the current year. This has some statistical support in that the '96 variable has larger explanatory power measured by R2 of the model and T-value of the parameter estimate. The p-value for the parameter estimate for this variable improves when the exterior self-evaluators are removed from the regression model. One possible explanation of the negative parameter estimate is that a "rich" municipality might use the extra resources on higher quality (not picked up by the DEA model) and/or allowing inefficiency in production of services.

Share of users of home care is a measure of the size of the share of home care clients in relation to all the clients getting nursing services. This coefficient has a negative parameter estimate. This is an indication that the technical efficiency tends to be lower when a larger part of the municipality's clients is in home care. This is interesting, because it is a measure of the product mix in the municipality. The DEA method takes into account the case mix when estimating the frontier. However, the distance between the frontier and the average unit behind the frontier might vary with case mix. It is important to remember that since we have no price information on the products (home care and institutionalized care), we do not know which group has the highest total efficiency. Without price information we can only estimate technical efficiency and scale efficiency, not allocative efficiency, which is also a component of total efficiency. Thus, we can make no recommendation of what is better, only point out that the variation of technical efficiency seem to grow with the share of home based nursing.

Share in home care of users in age group…(four age groups) measures how large share of the total population in an age group gets home based nursing services. With the exception of the lowest age group (0-66) all of the parameter estimates is statistically significant and positive. This supports our hypothesis that the higher the coverage of home based nursing, the lower the required resource usage per client. The reasoning is that the nursing sector behaves as if it ranks its potential clients from the ones that require the most nursing to the ones that requires the least, and that it uses this ranking as a prioritized list of which clients to accept first. If the municipality has a larger share of the population in an age group as its clients, we expect the average required resource usage per client to be lower because the average client is healthier.

Share in inst. care of users in age group ... (four age groups) is similar to the variables described above, but for institutionalized care. The parameter estimate for the youngest age group (0-66) is positive and statistically significant. It is a priori known that some of these clients require a lot of resource usage, but remember that the number of users in this group (inst. 0-66) is included in the DEA model. It might be that the municipalities who has a relatively large share of these users compared to their total population have healthier clients on average. The only other age group in inst. care that gets a statistically significant parameter estimate is 67-79 where the sign is negative. This is an indication that the "youngest of the oldest" require more resource usage in inst. care than the other groups above 67. It might be that it more for difficult for the clients in this relatively young age group to get inst. care, and that the clients who actually get it requires more resources on average than in the older age groups.

Removing the exterior self-evaluators can make a difference. In this case the explained share of the total variance in the model increased as R^2 rose from 17% to 21%. Both coefficient estimates and p-values change, sharpening the estimates of seven coefficients while only three had increased p-values[3]. While numerical changes are small, they are still sizeable considering that only 25 out of 469 observations (5%) were removed. Essentially, we have removed the units that are most likely not to contain any information, i.e. to be pure noise.

[3] In contrast, excluding all self-evaluators, both interior and exterior, would have lowered R^2 and decreased p-values only for three coefficients and increased them for seven.

This is of course not conclusive evidence that one approach is better than the other. The point we want to make is that it can make a difference. We have already argued that it makes theoretical sense to remove the exterior self-evaluators. It may be added that in Simar and Wilson (2003) it is conjectured that the bootstrap works better the denser the data. Since we have removed data points in regions that by definition are as "thin" as possible, the bootstrap should also work better. In sum, we feel that we have made a solid case for the advantages of identifying and removing the exterior self-evaluators when doing a two-stage analysis in a DEA setting.

5. Conclusions

The units found strongly efficient in DEA studies on efficiency can be divided into self-evaluators and active peers, depending on whether the peers are referencing any inefficient units or not. The contribution of the paper starts with subdividing the self-evaluators into *interior* and *exterior* ones. The exterior self-evaluators are efficient "by default"; there is no firm evidence from observations for the classification. Self-evaluators may most naturally appear at the "edges" of the technology, but it is also possible that self-evaluators appear in the interior. It may be of importance to distinguish between the self-evaluators being exterior or interior. Finding the influence of some variables on the level of efficiency by running regressions of efficiency scores on a set of potential explanatory variables is an approach often followed in actual investigations. Using exterior self-evaluators with efficiency score of 1 in such a "two-stage" procedure may then distort the results, because to assign the value of 1 to these self-evaluators is arbitrary. But regarding interior self-evaluators they may have peers that are fairly similar. They should then not be dropped when applying the two- stage approach.

A method for classifying self-evaluators based on the additive DEA model, either CRS or VRS, is developed. The exterior strongly efficient units are found by running the enveloping procedure "from below", i.e. reversing the signs of the slack variables in the additive model (1), after removing all the inefficient units from the data set. Which units of the strongly efficient units from the additive model (1) that turn out to be self-evaluators or active peers, will depend on the orientation of the efficiency analysis, i.e. whether input-or output orientation is adopted. The classification into exterior and interior peers is determined by the strongly efficient units turning out to be exterior ones running the "reversed" additive model (9).

The exterior self-evaluators units should be removed from the observations on efficiency scores when performing a two-stage analysis of explaining the distribution of the scores. The application to municipal nursing- and home care services of Norway shows significant effects of removing exterior self-evaluators from the data when doing a two-stage analysis. Thus the conclusions as to explanations of the efficiency score distribution will be qualified taking our new taxonomy into use.

References

Banker, R.D., A. Charnes and W.W. Cooper (1984) "Some models for estimating technical and scale inefficiencies." *Management Science*, 30, pp. 1078-92.

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

Charnes, A., W.W. Cooper and E. Rhodes (1978): "Measuring the efficiency of decision making units," *European Journal of Operations Research* 2, 429-444.

Cooper, W. W., L. M. Seiford, and K. Tone (2000): *Data Envelopment Analysis. A comprehensive text with models, applications, references and DEA-solver software*, Boston/Dordrecht/London: Kluwer Academic Publishers.

Edvardsen, D. F. and F. R. Førsund (2001): "International benchmarking of electricity distribution utilities, " Memorandum 35/2001, Department of Economics, University of Oslo.

Edvardsen, D. F., F. R. Førsund og E. Aas (2000): " Effektivitet i pleie- og omsorgssektoren" [Efficiency in the nursing- and home care sector], Rapport 2/2000, Oslo: Frischsenteret.

Erlandsen, E. and F. R. Førsund (2002): "Efficiency in the Provision of Municipal Nursing- and Home Care Services: The Norwegian Experience," in K. J. Fox (ed.): *Efficiency in the Public Sector*, Boston/Dordrecht/London: Kluwer Academic Publishers, x-y.

Färe, R. and D. Primont (1995): "Multi output production and duality: Theory and applications," Southern Illinois University at Carbondale.

Farrell, M. J. (1957): "The measurement of productive efficiency," *Journal of the Royal Statistical Society,* Series A, 120 (III), 253-281.

Førsund, F. R. and N. Sarafoglou (2002): "On the origins of data envelopment analysis," *Journal of Productivity Analysis* 17, 23-40.

Nerlove, M. (1965): *Estimation and identification of Cobb – Douglas production functions*, Amsterdam: North-Holland Publishing Company

Seitz, W. D. (1967): “Efficiency measures for steam-electric generating plants”, *Western Farm Economic Association, Proceedings 1966,* Pullman, Washington, 143-151.

Seitz, W. D. (1971): “Productive efficiency in the steam-electric generating industry,” *Journal of Political Economy* 79, 878-886.

Simar, L. and P. W. Wilson (2003): “Estimation and inference in two-stage, semi-parametric models of production processes,” Technical report 0310 IAP statistics network (http://www.stat.ucl.ac.be/Iapdp/tr2003/TR0310.ps).

Aas, E. (2000): “På leting etter målefeil – en studie av pleie- og omsorgssektoren”, Notater 2000:10, Statistics Norway, Oslo.