IHS Economics Series
Working Paper 84
September 2000

New Models for Data Envelopment Analysis: Measuring Efficiency Outwith the VRS Frontier

Iain Paterson

IRIHS Institutional Repository at IHS
Vienna

Institut für Höhere Studien
Institute for Advanced Studies
Vienna

Impressum

Author(s):
Iain Paterson

Title:
New Models for Data Envelopment Analysis: Measuring Efficiency Outwith the VRS Frontier

ISSN: Unspecified

2000 Institut für Höhere Studien - Institute for Advanced Studies (IHS)
Josefstädter Straße 39, A-1080 Wien
E-Mail: office@ihs.ac.at
Web: www.ihs.ac.at

All IHS Working Papers are available online:
http://irihs.ihs.ac.at/view/ihs_series/
This paper is available for download without charge at:
https://irihs.ihs.ac.at/id/eprint/1284/

84

Reihe Ökonomie

Economics Series

New Models for Data Envelopment Analysis

Measuring Efficiency Outwith the VRS Frontier

Iain Paterson

84

Reihe Ökonomie

Economics Series

New Models for Data Envelopment Analysis

Measuring Efficiency Outwith the VRS Frontier

Iain Paterson

September 2000

Institut für Höhere Studien (IHS), Wien
Institute for Advanced Studies, Vienna

Contact:

Iain Paterson
☎: +43/1/599 91-152
email: paterson@ihs.ac.at

Founded in 1963 by two prominent Austrians living in exile – the sociologist Paul F. Lazarsfeld and the economist Oskar Morgenstern – with the financial support from the Ford Foundation, the Austrian Federal Ministry of Education and the City of Vienna, the Institute for Advanced Studies (IHS) is the first institution for postgraduate education and research in economics and the social sciences in Austria. The **Economics Series** presents research done at the Department of Economics and Finance and aims to share "work in progress" in a timely way before formal publication. As usual, authors bear full responsibility for the content of their contributions.

Das Institut für Höhere Studien (IHS) wurde im Jahr 1963 von zwei prominenten Exilösterreichern – dem Soziologen Paul F. Lazarsfeld und dem Ökonomen Oskar Morgenstern – mit Hilfe der Ford-Stiftung, des Österreichischen Bundesministeriums für Unterricht und der Stadt Wien gegründet und ist somit die erste nachuniversitäre Lehr- und Forschungsstätte für die Sozial- und Wirtschaftswissenschaften in Österreich. Die **Reihe Ökonomie** bietet Einblick in die Forschungsarbeit der Abteilung für Ökonomie und Finanzwirtschaft und verfolgt das Ziel, abteilungsinterne Diskussionsbeiträge einer breiteren fachinternen Öffentlichkeit zugänglich zu machen. Die inhaltliche Verantwortung für die veröffentlichten Beiträge liegt bei den Autoren und Autorinnen.

Abstract

Some models are presented in this paper which extend the concept of measuring superefficiency to the useful case of variable returns-to-scales (VRS), thus enabling the ranking of efficient as well as inefficient units. Two models, namely the Universal Radial Model and the Universal Additive Model, are presented that also have strong invariance properties (units and translation invariance). For both of these models a method for normalising the efficiency scores on a (0-1+) scale is presented. These models have been implemented in a software package and applied to the ranking of units in an industrial context.

Keywords

Data envelopment analysis (DEA), superefficiency, universal models, variable returns-to-scales (VRS)

JEL Classifications

C61, C14, C88

Comments

This paper was written in April 1999. Since then both ‘universal’ models presented in this paper have been successfully implemented in C programming language at Daimler Chrysler, Stuttgart and Ulm, by Helmut Berrer. Many thanks are due to Helmut for his excellent support and the fruitful insights which have arisen from this teamwork.

Contents

Introduction

Some Data Envelopment Analysis (DEA) approaches to productivity measurement and ranking of units depend on measuring the ‘distance’ of a DMU (Decision Making Unit) to the convex hull spanned by other DMU’s; calculating Malmquist productivity indices is one case, ‘superefficiency’ measurement (Andersen and Petersen, 1993) is another. Standard input- or output-oriented DEA models exhibit deficiencies when applied to this task. Superefficiency is only well-defined for CRS (constant returns-to-scales) specifications, and such models are units invariant in the radial, but not in the slack components.

Three models are presented in this paper which extend the concept of measuring superefficiency to the useful case of variable returns-to-scales (VRS), thus enabling the ranking of efficient as well as inefficient units. Two of these models also have strong invariance properties.

(i) Universal Radial: This model is units and translation invariant (also for slacks) for the VRS specification: input or output data may thus assume negative or zero values. It is units invariant (also for slacks) for non-negative data in CRS.

(ii) Universal Additive: This model is units and translation invariant (slacks being included in the efficiency score) for the VRS specification: input or output data may thus assume negative or zero values. It is units invariant for non-negative data in CRS. Unlike the universal radial model, it exhibits however discontinuity in the (super)efficiency values along the weak efficient boundary, so the former model may be preferred for ranking purposes.

For both of these models a method for normalising the efficiency scores is presented, so that inefficient units obtain an ‘efficiency’ value less than one, weak efficient units obtain a value of one, and superefficient units obtain a value greater than one.

These models have been implemented in a software package and applied to the ranking of units in an industrial context.

Model 1 for Ranking

Let X and Y be the input and output matrices, respectively, for *all* DMU's. We implement a ranking based on the 'Andersen and Petersen' procedure i.e. each DMU *d* is evaluated with respect to a set including all *other* DMU's, except for itself. So

Y is the $k \times n$ matrix for *k* outputs and *n* DMU's

Y_{-d} is the $k \times (n-1)$ matrix for *k* outputs and *n-1* DMU's, without data for DMU *d*

Y^d is the $k \times 1$ output vector for DMU *d* being evaluated

(for example Y = the augmented matrix $\left[Y^1 \;\; Y_{-1} \right] = \left[Y_{-n} \;\; Y^n \right]$)

X is the $m \times n$ matrix for *m* inputs and *n* DMU's

X_{-d} is the $m \times (n-1)$ matrix for *m* inputs and *n-1* DMU's, without data for DMU *d*

X^d is the $m \times 1$ input vector for DMU *d* being evaluated

$\begin{Bmatrix} j \\ \underline{l} \end{Bmatrix}$ is the $n \times 1$ solution vector, for scalar j and $(n-1) \times 1$ vector $\underline{l}$

$\underline{0}$ is a $(n-1) \times 1$ vector of zeros

and $\bar{1}$ is a $1 \times (n-1)$ row vector of 1's.

Then the problem is to solve the following Linear Programs. Note that Y^d, Y_{-d}, X^d, and X_{-d} are in general different for each DMU *d.*

$$\max_{\underline{l} \geq \underline{0}} \; j \qquad \text{(MODEL 1)}$$

s.t.

$$X^d j + X_{-d} \underline{l} \leq X^d$$

$$Y^d j - Y_{-d} \underline{l} \leq -Y^d$$

$$\bar{1} \underline{l} = 1$$

$$\underline{l} \geq \underline{0}$$

Note that Model 1 may be presented concisely in the form ‘ max $\bar{c}\underline{x}$ s.t. $A\underline{x} \leq \underline{b}$ ’ as follows:

$$\max_{\underline{l} \geq \underline{0}} \; \begin{bmatrix} 1 & \bar{0} \end{bmatrix} \begin{Bmatrix} j \\ \underline{l} \end{Bmatrix} \qquad \text{(MODEL 1a)}$$

s.t.

$$\begin{bmatrix} X^d & X_{-d} \\ Y^d & -Y_{-d} \\ 0 & \bar{1} \\ 0 & -\bar{1} \end{bmatrix} \begin{Bmatrix} j \\ \underline{l} \end{Bmatrix} \leq \begin{Bmatrix} X^d \\ -Y^d \\ 1 \\ 1 \end{Bmatrix}$$

This may be instructive for constructing the matrices needed for the simplex tableau.

The *Efficiency Score* to be returned from the model = $1 - j$.

This model returns efficiency and/or superefficiency values for all real non-zero first-quadrant data i.e. $\forall \underline{x} \in X, \underline{x} > \underline{0}$ and $\forall \underline{y} \in \mathrm{Y}, \underline{y} > \underline{0}$. It may be similar to the radial improvement model (both input/output) of Thanassoulis (see reference), but latter does not allow for superefficiency measurement, and the efficiency score is different.

Universal Models

The aim of 'universal models' is to provide a framework for evaluating the efficiency (or 'superefficiency') of a DMU for the most general of convexity assumptions (VRS), whether or not that DMU is 'inside' or 'outside' the convex hull determined by a set of DMU'S which excludes the DMU itself being evaluated. The ability to carry out this evaluation is critical, for example, for the purposes of ranking, and the calculation of Malmquist productivity indices for panel data.

We are implementing a ranking proposal after the fashion of Andersen and Petersen (1993), i.e. the DMU *d* is not part of the set out of which the convex frontier is obtained, *but we do not know whether it is inside or outside the hull.* Further we utilize the invariance properties of the objective function introduced by Lovell and Pastor (1995)

The basic terminology is as follows:

Y is the $k \times n$ matrix for k outputs and n DMU's

Y_{-d} is the $k \times (n\text{-}1)$ matrix for k outputs and $n\text{-}1$ DMU's

Y^d is the $k \times 1$ output vector for DMU d being evaluated

X is the $m \times n$ matrix for m inputs and n DMU's

X_{-d} is the $m \times (n\text{-}1)$ matrix for m inputs and $n\text{-}1$ DMU's

X^d is the $m \times 1$ input vector for DMU d being evaluated

$\underline{l}$, $\underline{j_i}$, and $\underline{j_o}$, are vectors of length $(n\text{-}1)$, m, and k, respectively; $\overline{s_i^{-1}}$ and $\overline{s_o^{-1}}$ are row vectors of length m, and k, respectively

$\bar{1}$ is a $(n\text{-}1)$ row vector of 1's; $\underline{0}$ denotes zero vectors of appropriate length

$-\langle\rightarrow|+\rangle$ and $+\langle\rightarrow|-\rangle$ are *'algorithmic operators'*, as defined below

and an asterisk denotes pointwise vector multiplication.

Model 2: The Universal Additive Model

The Universal Additive model is formulated (for VRS) in two parts as:

a) The point *d* lies 'inside' the hull

$$I_{UA} = \max_{\underline{l},\underline{j}_i,\underline{j}_o \geq \underline{0}} \overline{(s_i^{-1})}(\underline{j}_i * X^d) + \overline{(s_o^{-1})}(\underline{j}_o * Y^d)$$

s.t.
$$\left(\underline{1 - j_i}\right) * X^d \geq X_{-d}\,\underline{l}$$
(MODEL 2a)
$$\left(\underline{1 + j_o}\right) * Y^d \leq Y_{-d}\,\underline{l}$$
$$\overline{1}\,\underline{l} = 1$$
$$\underline{l}, \underline{j}_i, \underline{j}_o \geq \underline{0}$$

Notice that for each input *i = 1,... , m*, $s_i = f\left(X[i,j]\right)$, *j* = 1, …, *n*, i.e. the standard deviation for each input *i* is calculated over the *entire* set of DMU'S, including the input X_i^d of the DMU *d* being evaluated. I_{UA} may also be written as

$$I_{UA} = \max_{\underline{l},\underline{j}_i,\underline{j}_o \geq \underline{0}} \sum_{i=1}^{m} s_i^{-1} j_i X_i^d + \sum_{o=1}^{k} s_o^{-1} j_o Y_o^d$$

b) The point *d* lies 'outside' the hull

$$I_{UA} = \max_{\underline{l} \geq \underline{0};\underline{j}_i,\underline{j}_o \leq \underline{0}} \overline{(s_i^{-1})}(\underline{j}_i * X^d) + \overline{(s_o^{-1})}(\underline{j}_o * Y^d)$$

s.t.
$$\left(\underline{1 - j_i}\right) * X^d \geq X_{-d}\,\underline{l}$$
(MODEL 2b)
$$\left(\underline{1 + j_o}\right) * Y^d \leq Y_{-d}\,\underline{l}$$
$$\overline{1}\,\underline{l} = 1 \qquad \underline{l} \geq \underline{0}, \quad \underline{j}_i, \underline{j}_o \leq \underline{0}$$

The only difference in Model 2b from Model 2a is that the restrictions on j_i, j_o are non-positive instead of non-negative. However, since we do not know *a priori* which model is applicable, this difference is very significant.

The following algorithm solves the DEA for DMU *d*

Step 1) Run Model 2a for DMU *d*

- if an optimal value for I_{UA} is found, this is the solution

- if no feasible value is found, proceed to step 2

Step 2) Run Model 2b for DMU *d*

- the optimal value for I_{UA} is the solution.

For this reason we can formulate the following universal model:

For DMU *d* inside or outside the hull

$$I_{UA} = \max_{\underline{l}, \underline{j}_i, \underline{j}_o \geq \underline{0}} \quad \langle \rightarrow | - \rangle \overline{(s_i^{-1})} (\underline{j}_i * X^d) + \langle \rightarrow | - \rangle \overline{(s_o^{-1})} (\underline{j}_o * Y^d)$$

s.t.

$$\left(\underline{1 - \langle \rightarrow | + \rangle j_i} \right) * X^d \geq X_{-d} \underline{l} \qquad \text{(MODEL 2c)}$$

$$\left(\underline{1 + \langle \rightarrow | - \rangle j_o} \right) * Y^d \leq Y_{-d} \underline{l}$$

$$\overline{1} \, \underline{l} = 1$$

$$\underline{l}, \underline{j}_i, \underline{j}_o \geq \underline{0}$$

where the symbol $- \langle \rightarrow | + \rangle$, for example, means that the LP solution is initially started with the minus operator: if no feasible initial solution can be found, then the simplex algorithm proceeds with the plus operator. The correct solution is thereby guaranteed.

It can be shown that this formulation is *units invariant and translation invariant in the first quadrant*: this means that the affine transformation must be such that

$$aX_i + b > 0 \quad \forall i = 1,\ldots,m \text{ and } aY_o + b > 0 \quad \forall o = 1,\ldots,k.$$

Two Extensions

First we introduce the *Reference Base* point $r^* = \{X_i^*, Y_o^*\}$ for i = 1,..., m and o =1,..., k as an 'artificial DMU'. It is defined as follows:

$$X_i^* = \max_{j=1,\ldots,n} X_i^j \text{ for each input } i\text{, and}$$

$$Y_o^* = \min_{j=1,\ldots,n} Y_o^j \text{ for each output } o.$$

The 'DMU' r^* is now added to the set of DMU's so that there are n+1 DMU's altogether and X and Y are $m\times(n+1)$ and $k\times(n+1)$ matrices. Likewise X_{-d} and Y_{-d} are $m\times n$ and $k\times n$ matrices now, etc.

Second, making the transformations $j_i X_i^d = s_i \;\; \forall i = 1,\ldots,m$ and $j_o Y_o^d = s_o \;\; \forall o = 1,\ldots,k$, we can rewrite Model 2c as the Universal Additive Model.

For DMU d inside or outside the hull

$$I_{UA} = \max_{\underline{l}, \underline{s_i}, \underline{s_o} \geq \underline{0}} \langle\rightarrow|-\rangle \overline{(s_i^{-1})}\, \underline{s_i} + \langle\rightarrow|-\rangle \overline{(s_o^{-1})}\, \underline{s_o}$$

s.t.

$$X_{-d}\,\underline{l} + \langle\rightarrow|-\rangle \underline{s_i} \leq X^d \qquad \text{(MODEL 2d)}$$

$$Y_{-d}\,\underline{l} - \langle\rightarrow|+\rangle \underline{s_o} \geq Y^d \qquad \textbf{(UNIVERSAL ADDITIVE)}$$

$$\overline{1}\,\underline{l} = 1$$

$$\underline{l},\, \underline{s_i},\, \underline{s_o} \geq \underline{0}$$

The apparent similarity of the Universal Additive Model to the Normalised Additive model of Lovell and Pastor (as described in Holzer. 1999) is deceptive: even although s_i and s_o may be regarded as 'slacks', the *inequalities play an important non-trivial role*. In the simplex solution additional slacks will be added to take account of the inequalities!

The Model 2d is now completely translation and units invariant, which means that the correct solution will be calculated even if the data contains zeros or negative values.

Denote the DEA solution for DMU *d* as $I_{UA}(d)$. The inefficiency measure for the reference point r^* is also calculated by the DEA procedure. Let the result of this calculation be $I_{UA}(r^*)$, i.e. Model 2d is solved with setting *d* = r^*.

We define the ***Normalised Universal Additive Efficiency*** $E^*_{UA} = 1 - \dfrac{I_{UA}}{I_{UA}(r^*)}$, so that

$E^*_{UA} = 1$ for DMU's on the strong *and weak* efficient boundaries, $E^*_{UA} \geq 1$ for DMU's outside the hull , $E^*_{UA}(r^*) = 0$ and in general $E^*_{UA} \geq 0$.

This model exhibits a certain disadvantage: a DMU *d* which is on the weak efficient frontier will in general have $E^*_{UA}(d) < 1$ (and definitely not equal to 1). However if there is, for example, another DMU d^e which merely differs from *d* by a very small $e > 0$ such that $Y_1^{d^e} = Y_1^d + e$ and otherwise $Y_o^{d^e} = Y_o^d$, *o* = 2,..., *k* and $X_i^{d^e} = X_i^d, \forall i$ = 1,... , *m*., then $E^*_{UA}(d^e) = 1$ as e decreases to zero. This represents a certain theoretical discontinuity in the efficiency which is unavoidable in this model, even if its practical implications may be considered as negligible.

Model 3: The Universal Radial Model

The Universal Radial Model is presented here using the same meaning for the notation as above. In particular the reference base-point r^* is included in the set of DMU's.

The model is formulated as follows:

For DMU *d* inside or outside the hull

$$I_{UR} = \max_{z \geq 0,\ \underline{l} \geq \underline{0},} \quad \langle \rightarrow | - \rangle\, z$$

$$\text{s.t.} \qquad X_{-d}\,\underline{l} + \langle \rightarrow | - \rangle\, \underline{s_i}\, z \;\leq\; X^d \qquad \text{(MODEL 3)}$$

$$Y_{-d}\,\underline{l} - \langle \rightarrow | + \rangle\, \underline{s_o}\, z \;\geq\; Y^d \qquad \textbf{(UNIVERSAL RADIAL)}$$

$$\overline{1}\,\underline{l} = 1$$

$$z \geq 0,\ \underline{l} \geq \underline{0},$$

where z is a scalar.

This model is radial because the projection is always along all input and output dimensions, which is not necessarily the case for the additive model, for example. It is 'equi-radial because the 'distance' to the frontier is the same in each input and output, after the normalisation by means of standard deviations is taken into account.

Just as is the case for the Universal Additive model, the Universal Radial model is also completely translation and units invariant. In other words, no matter what zeros or negative values are contained in the data, the solution will be found, and is identical for any affine translation.

There is no discontinuity in efficiency measurement in this model. It is truly a universal model which can be used to measure efficiency from inside or outside the hull for the specification of variable returns-to-scale. Suggestions for implementing constant returns-to-scale, non-increasing returns-to-scale, or non-decreasing returns-to-scale models based on this approach will follow at a later date.

Normalised Universal Radial Efficiency

In just the same manner as above this is calculated for DMU *d* as

$$E_{UA}^{*}(d) = 1 - \frac{I_{UA}(d)}{I_{UA}(r^{*})}$$

and has the same properties mentioned above.

Points to Remember

- The Reference Base-point $\{X_i^*, Y_o^*\}$ must be pre-calculated and added to the data set before constructing the tableau.
- Standard deviations $\boldsymbol{s}_i, \boldsymbol{s}_o$ are calculated *after* adding the Reference Base-point, i.e. $\boldsymbol{s}_i = \boldsymbol{s}\,(X_i^j)$ for j = 1,… , n+1 DMU's, for example.
- In solving the Universal Additive and Universal radial models it is important to construct the initial tableau as indicated by the signs in the formulations.
- Only if no feasible solution can be found, i.e. the artificial variables in the extended tableau can not be removed, do the signs in the tableau get changed according to the symbols $\langle \rightarrow | + \rangle$ and $\langle \rightarrow | - \rangle$.
- In the Universal Additive case, these sign changes will affect the objective function row of the tableau and one variable in each input and output row.
- In the Universal Radial case the sign changes will only affect the one column in the tableau (i.e. that pertaining to variable $\boldsymbol{Z}$ in the objective function and input and output constraint rows).
- It is important that the inequalities as indicated in the constraints are used as given; in particular, although the solution variables $\underline{S_i}$ and $\underline{S_o}$ effectively calculate slack values, extra (here not named) slacks will have to be added to the tableau (even although these will all turn out to have zero value in the solution!).

- If it is chosen to implement subroutines which require the objective function to be minimised instead of maximised then these formulations should be as follows:

$$I_{UA} = \min_{\underline{l}, \underline{s_i}, \underline{s_o} \geq \underline{0}} \; -\langle \rightarrow | + \rangle \overline{(s_i^{-1})}\, \underline{s_i} \; -\langle \rightarrow | + \rangle \overline{(s_o^{-1})}\, \underline{s_o}$$

and

$$I_{UR} = \min_{z \geq 0,\; \underline{l} \geq \underline{0},} \; -\langle \rightarrow | + \rangle\, z$$

- After calculating I_{UA} or I_{UR} for each of n+1 DMU's (including the reference point), the corresponding E^*_{UA} and E^*_{UR} may be calculated.

References

Andersen P., Petersen N. C. (1993): A Procedure for Ranking Efficient Units in Data Envelopment Analysis, *Management Science*, Vol. 39, No. 10, October, pages 1261–1264.

Holzer I. (1999): Einsatz der Data-Envelopment-Analysis zur Optimierung von Drall und Durchflußzahl (in German, English title: Application of Data-Envelopment-Analysis for Swirl and Discharge Optimization), diploma thesis, DaimlerChrysler AG, Stuttgart.

Lovell K., Pastor J. (1995): Units Invariant and Translation Invariant DEA Models, *Operations Research Letters* 18, pages 147–151.

Thanassoulis E.: 'Warwick DEA Software', see http://www.warwick.ac.uk/~bsrlu/ and http://research.abs.aston.ac.uk/mgtsc1/mgtourteam.html

Author: Iain Paterson

Title: New Models for Data Envelopment Analysis. – Measuring Efficiency Outwith the VRS Frontier

Reihe Ökonomie / Economics Series 84

Editor: Robert M. Kunst (Econometrics)
Associate Editors: Walter Fisher (Macroeconomics), Klaus Ritzberger (Microeconomics)

ISSN: 1605-7996
© 2000 by the Department of Economics, Institute for Advanced Studies (IHS),
Stumpergasse 56, A-1060 Vienna • ☎ +43 1 59991-0 • Fax +43 1 5970635 • http://www.ihs.ac.at

ISSN: 1605-7996