Efficient and Close Targets in Data Envelopment Analysis (DEA)

Maria Conceição A. Silva Portela*, *Universidade Católica Portuguesa, Portugal*

Pedro Castro Borges**, *Universidade Católica Portuguesa, Portugal*

Emmanuel Thanassoulis***, *Aston Business School, U.K.*

Abstract

This paper draws attention for the fact that traditional Data Envelopment Analysis (DEA) models do not provide the closest possible targets (or peers) to inefficient units, and presents a procedure to obtain such targets. It focuses on non-oriented efficiency measures both measured in relation to a Free Disposal Hull (FDH) technology and in relation to a convex technology. The approaches developed for finding close targets are applied to a sample of Portuguese bank branches.

1. Introduction

One of the key practical outcomes in an efficiency assessment is the identification of targets. Targets may be identified by any DEA model, radial or non-radial, oriented or non-oriented. Our focus here is on non-oriented (and therefore non-radial) models of efficiency, which assume that production units are able to control, and thus change, inputs and outputs simultaneously. A drawback of the existing non-oriented DEA models [like the hyperbolic model introduced by Färe, Grosskopf, and Lovell (1985) or the additive model due to Charnes et al. (1985)] is that they either impose strong restrictions on the movements towards the efficient frontier, or they aim at maximising slacks. Both these facts contribute to finding targets and peers that may not be the closest possible to the units being assessed. If Pareto-efficiency can be achieved by inefficient units with less effort than that implied by targets derived using traditional DEA efficiency models, then it is at least of practical value to find the closest targets for each inefficient unit we can. Close targets in this sense are in line with the original spirit of DEA of showing each unit in the best possible light.

The idea of finding closest targets and peers has appeared in the literature both associated with oriented models [see for example Coelli (1998), or Cherchye and Puyenbroek (2001)] and non-oriented models [see for example Frei and Harker (1999) or Golany, Phillips, and Rousseau (1993)]. It is our intention to explore this issue for the most general case of non-oriented efficiency measures. In addition, the analysis will be restricted to technical efficiency.

2. Non-Radial-non-oriented Measures of Efficiency

Non-oriented DEA models, like the additive model of Charnes et al. (1985) or its variant the RAM (Range Adjusted Measure) as proposed by Cooper, Park, and Pastor (1999), explicitly maximise slacks in their objective functions. The Russell graph measure of Färe, Grosskopf, and Lovell (1985) (of which the hyperbolic measure of efficiency is a special case), or the directional distance function introduced by Chambers, Chung, and Färe (1996, 1998) also maximise slacks, though this is not explicit in the objective function. Such models, therefore, are not able to provide close and efficient targets to inefficient units they identify.

* Corresponding author, Universidade Católica Portuguesa, Centro Regional do Porto, R. Diogo Botelho, 1327, 4169-005- Porto, Portugal, csilva@porto.ucp.pt

** Universidade Católica Portuguesa, Centro Regional do Porto, Portugal, pmb@porto.ucp.pt

*** Aston Business School, Aston Triangle, B4 7ET Birmingham. U.K., e.thanassoulis@aston.ac.uk

Our objective is on the one hand to find an appropriate measure of efficiency and, on the other hand, to operationalise this measure so targets can be found which are Pareto-efficient and close to some inefficient unit. Two requirements for an appropriate measure of efficiency in a non-oriented context are: (i) it should be capable of incorporating all the sources of inefficiency, while at the same time (ii) retaining the meaning of radial efficiency measures. The above mentioned directional and hyperbolic measures do not satisfy the first requirement, while the RAM, the additive model, and the Russell graph measure do not satisfy the second requirement. A measure that satisfies both requirements is that developed by Brockett et al. (1997), which will be referred to as BRWZ throughout. The BRWZ efficiency measure is shown in (1).

$$BRWZ_o = \frac{1}{m}\left(\sum_{i=1}^{m}\frac{x_{io}-e_i^*}{x_{io}}\right)\times\frac{1}{s}\left(\sum_{r=1}^{s}\frac{y_{ro}}{y_{ro}+s_r^*}\right) \Leftrightarrow BRWZ_o = \frac{\sum_{i=1}^{m}h_{io}\times\sum_{r=1}^{s}1/g_{ro}}{m\times s} \qquad (1)$$

The first expression in (1) assumes that all inefficiencies are captured by additive slack values, e_i^* and s_r^*, where the star denotes an optimal value of the input and output slacks as resulting from the solution of some DEA model which projects units on the Pareto-efficient boundary. The second expression for the BRWZ measure in (1) makes it possible to show that the BRWZ measure is closer to the meaning of radial efficiency measure. In addition, if we assume that all inputs change equiproportionately (each $h_i = \theta$) and that outputs are not allowed to change (each $g_r = 1$), then the BRWZ measure reduces to θ, which coincides with the Farrell measure of input efficiency. The BRWZ measure is also units invariant which is a considerable advantage.

3. Closer Targets and Efficiency

Let us first define the notion of closeness of targets. In general, we say that unit B is closer to A than to C, if moving from A to B requires smaller changes in inputs and outputs than those required in moving from A to C. Such changes can be expressed in terms of ratios of input and output levels at the two different points concerned, where the larger the ratios the closer the points will be. Obviously in a non-oriented space with multiple inputs and outputs one needs to choose a form of aggregating such ratios. In our case, the BRWZ efficiency measure was chosen for this aggregation. Thus, the closer the target point to an observed point the higher the BRWZ efficiency as a measure of the distance between the two points.

The closeness between two points can also be measured using an L_p metric. Such metrics are not expressed in ratio form but in difference form. Therefore they have the disadvantage of not being units invariant. The L_p distance between two points (A and B) is given by $\left[\sum_{i=1}^{n}|A_i - B_i|^p\right]^{1/p}$. We can illustrate concepts of closeness between points using a simple one input - one output example as shown in Figure 1. Unit F is FDH and BCC inefficient, where in the first case it is dominated by units B and C. Unit C is closer to F than is unit B. This can be seen in
Table 1 where the BRWZ measure and some metric distances between points F, C, and B are presented. Clearly point B is the point that maximises the sum of slacks (see L_1 metric), meaning that the non-oriented models we saw previously - additive, RAM, and Russell Graph measure - identify point B as the target of unit F rather than point C. This happens both for the case of FDH and convex technology. In the convex context the closest point in terms of the BRWZ measure is point (5, 5.33) – a convex combination between points B and D.
Table 1 shows that this point is closer to F than the target point B in terms of the BRWZ measure and in terms of the L_1 norm, but not in terms of the two other norms. We favour

comparisons based on the BRWZ measure because it is units invariant, - a characteristic that is important when units of measurement are subjective.

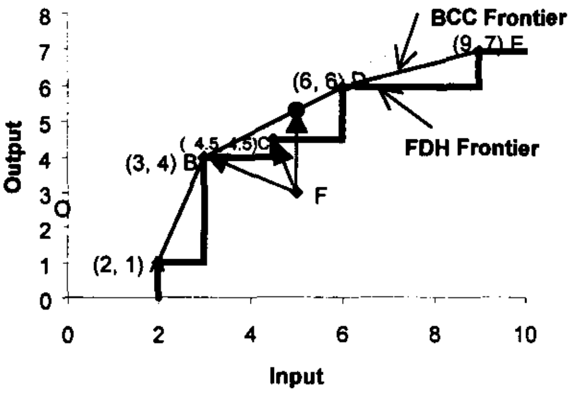

Figure 1: One input/output example

Point	BRWZ	L_1	L_2	L_∞
B	45%	3	$\sqrt{5}$	2
C	60%	2	$\sqrt{2.5}$	1.5
(5, 5.33)	56.25%	2.33	$\sqrt{5.44}$	2.33

Table 1: Distance of F from points C, B, and **(5, 5.33)**

4. Identifying Closer Targets in a FDH Technology

The approach developed in this paper takes advantage of the fact that in FDH targets correspond to a single observed unit (peer), which simplifies their identification and the calculation of efficiency. Calculating efficiency requires first the knowledge of the set of dominating units for each dominated unit, and then the selection of the one (the closest) that should be used as the peer.

Our approach follows three steps:

Step I Determine the set of non-dominated units (100% FDH efficient);
Step II Determine a peer unit for each dominated unit;
Step III Calculate the efficiency score.

Step I partitions the set of observed units into two sets: *ND* and *D*. *ND* is the set of non-dominated (or dominating) units (units in relation to which no other unit exists presenting all lower or equal inputs and all higher or equal outputs) and *D* is the set of the remaining units, called dominated. Although this operation can be performed for each unit by comparing it with all the other units or with the current non-dominated set, there are more efficient implementations used, especially in multiple objective combinatorial optimization. In our case we used an algorithm presented in Borges (2000), together with other well known quad tree algorithms to identify non-dominated units.

Step II finds a peer unit for each inefficient or dominated unit. In order to find this unit, we consider a subset

of *ND*, named K_o, consisting of the units that dominate the unit o being assessed. For each inefficient unit, its closest peer is determined through the BRWZ efficiency measure, that is, calculating (2) for every unit $k \in K_o$, where the subscript o identifies the inefficient unit being assessed.

$$Peer\ of\ unit\ o = \max_k \left(\frac{\sum_{i=1}^{m} x_{ik} / x_{io}}{m} \times \frac{\sum_{r=1}^{s} y_{ro} / y_{rk}}{s} \right) \quad \textbf{(2)}$$

Step III generates the efficiency of the unit being assessed in reference to the peer unit identified in the previous step. The measure of efficiency is given directly from the value obtained in (2). Therefore, steps II and III take place simultaneously.

5. Identifying Closer Targets in a Convex Technology

Extending the above procedure to convex technologies is not straightforward because in this case target points are not to be restricted to observed units but to convex combinations of Pareto-efficient units. As a result, an enumeration oriented procedure, which calculates the BRWZ measure for a set of potential target points can no longer be applied. The approach to follow in the case of convex frontiers is to use a DEA model where the BRWZ is maximised. However, in order to assure that the maximum BRWZ projection corresponds to a Pareto-efficient point one needs to impose some restrictions on the reference set [see for example Golany, Phillips, and Rousseau (1993) and Frei and Harker (1999)]. For this purpose we identify efficient facets through Qhull as proposed by Olesen and Petersen (2001). This software identifies all full dimension efficient facets (FDEF) in a DEA model, and provides a supporting hyperplane equation for each facet. The procedure can also be modified to identify non-full dimensional efficient facets.

Our procedure for finding the closest targets in convex technologies consists of three steps:

Step I Determine the set of Pareto-efficient units (E) by solving the additive model;
Step II Determine all Pareto-efficient facets (F_k) using QHull;
Step III For each F_k $k = 1,\ldots, K$ solve model (3) to find the closest targets for inefficient unit *o*.

$$Max \left\{ \begin{array}{l} BRWZ_o = \dfrac{1}{m \times s}\left(\sum_{i=1}^{m} h_{io} \times \sum_{r=1}^{s} 1/g_{ro}\right) \mid \sum_{j \in F_k} \lambda_j y_{rj} = g_{ro} y_{ro} \,, \sum_{j \in F_k} \lambda_j x_{ij} = h_{io} x_{io}, \\ \sum_{j \in F_k} \lambda_j = 1, \;\; \lambda_j \geq 0, \;\; g_{ro} \geq 1, \;\; 0 \leq h_{io} \leq 1 \end{array} \right\} \qquad (3)$$

In order to assure projection to the Pareto-efficient frontier, only points on F_k are considered as potential projections of unit *o* in (3). The final BRWZ efficiency measure of unit *o* is the maximum value found for the measure after model (3) is solved for all K facets. Step III is repeated for each inefficient unit for which we wish to identify the closest targets. We used GAMS and its non-linear programming solver (CONOPT) to solve (3).

6. An Illustrative Application to Bank Branches

The above procedures were applied to a sample of 24 Portuguese bank branches which are located in mid sized cities (as classified by the bank) in the northern region of Portugal. An intermediation approach of banking activities will be used, as this requires in principle non-oriented models. In this sense on the input side cost related variables are used (staff costs and other operating costs), and on the output side revenue related variables are used (value of current accounts, value of credit, and interest revenues). We assume that all inputs and outputs are discretionary. The data correspond to the month of July 2001 and values are expressed in thousands of Euros. Here only a few results of some inefficient units are exemplified.

For the FDH case, the application of the additive units invariant model, the RAM model, and the Russell graph model result in the same peers for inefficient units in all the cases. This is illustrated in Table 2, which shows the BRWZ measure calculated *a posteriori* in relation to the targets identified by these models. It also shows the BRWZ efficiency measure obtained under our closer target (CT) FDH procedure.

Ineff. Unit	B3	B5	B9	B13	B15	B19	B21	B22	B59
Peer Unit	B10	B10	B16	B10	B10	B10	B10	B10	B10
BRWZ Effcy	67.02%	77.26%	64.7%	74.85%	53.57%	68.15%	71.87%	52.76%	74%
BRWZ CT Effcy	67.02%	77.26%	64.7%	74.85%	53.57%	81.30% (B20)	71.87%	78.00% (B52)	74%

Table 2: Results from additive-FDH, RAM-FDH, FGL-FDH and CT procedure

The BRWZ measure has the same value under all the procedures for identifying targets, except in two cases. The reason for this is simple: unit B10 dominates most of the units in the sample and most of them are solely dominated by this unit. As the set of potential referents consists of a single unit there is not much for the alternative procedures to choose. Only in two cases is there a genuine choice of targets to be made: the case of inefficient units B19 and B22. The first unit is dominated by B10 and also by B20, and the second unit is dominated by B10, B26, B50 and B52. The application of our CT procedure clearly identifies closer targets to units B19 and B22 (respectively B20 and B52) as testifies a higher value of the CT BRWZ efficiency score in Table 2. These higher efficiency scores also correspond to lower metric distances.

In the convex VRS technology case, the application of the CT procedure to the bank branches example results (in its first step) in a set of efficient units, which are used in QHull to identify the set of efficient facets. These are: F_1 = {B10, B16, B20, B29, B50}; F_2 = {B20, B27, B29, B50, B57}; F_3 = {B10, B20, B27, B29, B50}; F_4 = {B10, B27, B56, B57}; F_5 = {B10, B11, B16, B29}; F_6 = {B10, B11, B26, B29}; F_7 = {B10, B26, B27, B29}, where the first three facets are full dimensional and the last four are not. In the third step, model (3) was applied to each inefficient unit in relation to each efficient facet. The facet chosen for projection in each case was the one maximising the objective function of model (3).

We applied the additive, the RAM and the CT procedure to our data set. Results in terms of the BRWZ show for all units better results of the CT procedure than the two other models. This confirms that our model shows each inefficient unit in a much better light than the other two models not only in terms of the BRWZ measure but also in terms of L_p metric measures. Take for example units B15 and B59 shown in Table 3. Results for these units show closer targets identified by the CT procedure for convex technologies than those identified by the additive model (the same being true for the RAM model). This fact is expressed in higher BRWZ efficiency scores and smaller L_p metrics, as illustrated for the two cases above (this fact can however be generalised to the entire sample of units). Interestingly the additive model tends to identify most of the inefficiencies associated with outputs, while the CT procedure for convex technologies identifies most of the inefficiencies associated with inputs. For the additive model the average BRWZ-input efficiency is 98.27% and the average BRWZ-output efficiency is 73.36%, while the corresponding values for the CT procedure are 90.72% and 92.03%, respectively. This clearly indicates that our procedure and the additive model identify different directions for improvement of inefficient units. The choice of the model to use should not, thus, be taken lightly.

	B15			B59		
	Observed	**Targets Additive**	**Targets CT convex**	**Observed**	**Targets Additive**	**Targets CT convex**
x_1	11.717	11.717	11.487	13.338	13.338	12.606
x_2	29.314	24.726	16.122	24.820	24.820	19.030
y_1	4070.630	5682.936	4070.630	4354.301	6073.258	4475.281
y_2	6418.995	14409.226	6418.995	10889.840	14368.013	10889.840
y_3	40.328	69.268	45.086	57.033	74.865	57.033
L_1		9636.066	18.181		5214.962	127.502
L_2		8151.330	14.027		3879.796	121.121
L_∞		7990.231	13.193		3478.173	120.980
BRWZ		53.58%	73.83%		74.56%	84.82%

Table 3: Distance to targets from inefficient units in the VRS case

Conclusion

The analysis of non-oriented measures of efficiency and their use to identify the closest targets for inefficient units was performed both considering FDH and convex technologies. The chosen criterion of closeness is based on the maximisation of the BRWZ efficiency measure, which has the advantage over other efficiency measures of capturing all the sources of inefficiency and retaining a meaning that is close to that associated with radial oriented efficiency measures. In order to use this measure multi-stage procedures are required both in the FDH and in the convex case to find the closest targets. The application of our procedure to a real bank branch example shows that it provides closer and easier-to-achieve targets in both, the FDH and convex, cases.

References

Borges, P. C. 2000, 'CHESS: Changing Horizon Efficient Set Search. A simple principle for multiobjective optimisation' *Journal of Heuristics*, vol. 6, pp. 405-418.

Brockett P. L., Rousseau J. J., Wang Y., and Zhow L. 1997, 'Implementation of DEA Models Using GAMS' Research Report 765, University of Texas, Austin.

Chambers, R. G. and Chung, Y. and Färe, R. 1996, 'Benefit and distance functions' *Journal of Economic Theory*, vol 70, pp. 407-419.

Chambers, R. G. and Chung, Y. and Färe, R. 1998, 'Profit, directional distance functions, and Nerlovian efficiency' *Journal of Optimization Theory and Applications*, vol. 98, no.2, pp. 351-364.

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

Cherchye, L. and Puyenbroeck, T.V. 2001, 'A Comment on Multi-stage DEA methodology', *Operations Research Letters,* vol. 28, pp. 93-98.

Coelli, Tim 1998, 'A multi-stage methodology for the solution of orientated DEA models', *Operations Research Letters*, vol. 23, pp. 143-149

Cooper W.W., Park K.S., Pastor J.T. 1999, 'RAM: A Range Measure of Inefficiency for Use with Additive Models, and Relations to other Models and Measures in DEA' *Journal of Productivity Analysis*, vol. 11, pp. 5-42.

Färe, R. and Grosskopf, S. and Lovell, C. A. K. 1985, *The Measurement of Efficiency of Production*. Kluwer-Nijhoff Publishing, Boston..

Frei, F. X. and Harker, P. T. 1999, 'Projections onto efficient frontiers: theoretical and computational extensions to DEA' *Journal of Productivity Analysis*, vol. 11, no. 3, pp. 275-300.

Golany, B. and Phillips, F.Y. and Rousseau, J.J. 1993, 'Models for improved effectiveness based on DEA efficiency results' *IIE Transactions*, vol. 25, no. 6, pp. 2-10.

Olesen, O. B. and Petersen, N. C. 2001, 'Identification and use of efficient facets in DEA' *Forthcoming in Journal of Productivity Analysis.*