A Layer DEA Model for Measuring and Improving the Efficiency in the Presence of Special Decision Making Units

M.R. Alirezaee,   and   S.A Mir-Hassani

Abstract: In the evaluation of non-efficient units by Data Envelopment Analysis (DEA) referenced Decision Making Units (DMU's) have an important role. Unfortunately DMU's with extra ordinary output can lead to a monopoly in a reference set, the fact called abnormality due to the outliers' data. In this paper, we introduce a DEA model for evaluating DMU's under this circumstance. The layer model can result in a ranking for DMU's and obtain an improving strategy leading to a better layer.

Keywords: Data Envelopment Analysis, Layer Model, Special Decision Making Units.

1. Introduction

Data Envelopment Analysis (DEA) is a well-known technique for measuring the relative efficiency of Decision Making Units (DMU's) with multiple inputs and outputs. Traditional approaches to efficiency have focused on averages of parameters, utilizing one optimized regression equation assumed to be appropriate for every DMU, but in DEA, focus is on the individual observation. The efficiency measure of each DMU is optimized thereby giving an understanding of each DMU, not a description of the average.

Also this method does not make assumption about functional forms; it makes a piecewise frontier (Efficient Frontier) with calculation of a maximal efficiency measure for each DMU relative to all other observed measures. While, the drawback of this approach is its weakness in detecting the measurement error the underestimation of which can lead to the derived efficient frontier that contains some units without a wide spread acceptance. Because of this shortcoming, a classification of the observed input-output vectors is necessary.

Therefore, the reminder of this paper is organized into 6 sections. Section 2 presents a general view of Data Envelopment Analysis. Section 3 illustrates measurement pitfalls of DEA results. Section 4 describes the proposed methodology. Section 5 presents efficiency improvement algorithm. Section 6 discusses on the computational aspects of the algorithm and after that, concluding remarks appear in section 7.

2. Data Envelopment Analysis

Data Envelopment Analysis (DEA), is a linear programming based method which evaluates the relative efficiency of Decision Making Units (DMUs), with multiple inputs and outputs, using a linear programming based model.

A major advantage cited in support of use of DEA in measuring efficiency, is that, this method do not require any price data. This is a distinct advantage, because in general, input price data are seldom available.

Therefore, this method does not make assumption about functional forms; it makes a piecewise frontier (Efficient Frontier) with calculation of a maximal efficiency measure for each DMU relative to all other observed measures.

Also, it identifies a subset of efficient "best-practice" DMUs and for the remaining DMUs, the magnitude of their non-productive is measured by compare to a frontier constructed from the efficient DMUs.

Charnes et al. (1978) first proposed DEA as an evaluation tool to measure and compare a DMU's the relative efficiency.

Their model which is commonly refereed to as a CCR model, assumed Constant Returns to Scale. It was developed for Variable Returns to Scale, by Banker et al. (1984). That is commonly refereed to as a BCC model.

Definition 1: Production Possibility Set (PPS)
A PPS or Production Technology is a set of points which represents all output vector Y, which can be

produced using the input vector X. (see figure 1). So it is:
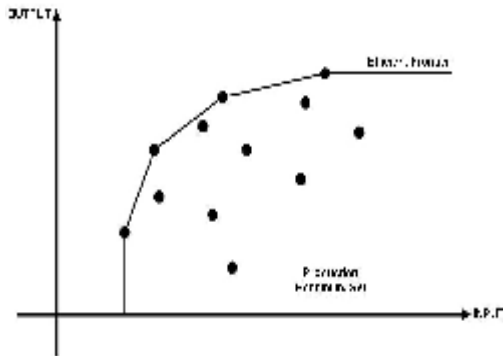
PPS = { (X,Y) : X can produce Y }

Fig. 1. Definition 2: Distance Function

As noted in previous definition, the Output Distance Function is:

$$d(X,Y) = \min\{d : (X, Y/d) \in PPS\}. \tag{1}$$

Suppose, there are k decision making units with s outputs and m inputs, where

X , is a $n \times m$ matrix of input quantities for all n, DMUs.

Y , is a $n \times s$ matrix of output quantities for all n, DMUs.

x_p, is a $m \times 1$ vector of input quantities for the p-th, DMU.

y_p, is a $s \times 1$ vector of output quantities for the p-th, DMU.

z , is a scaler.

The general DEA model of relative efficiency for the *p*-th DMU, is calculated by below formula:

$$[d(x,y)]^{-1} = Max \ \ z = \frac{U^T y_p + \delta_1 \tau}{W^T x_p}$$

subject to :

$$U^T Y_j - W^T X_j - \delta_1 \tau \le 0 \tag{2}$$

$$W^T X_j = 1$$

$$W \ge \varepsilon, U \ge \varepsilon, \ \ \delta_1 . \delta_2 . (-1)^{\delta_3} . \tau \ge 0$$

Where, W and U are weights vector for inputs and outputs, respectively. Also, the non-Archimedean infinitesimal Epsilon is used in the model for some computational considerations, for more details see [6]. It can be easily verified that

When $(\delta_1, \delta_2, \delta_3) = (0, \nabla, \nabla)$, the model is based on constant returns to scale assumption.

When $(\delta_1, \delta_2, \delta_3) = (1, 0, \nabla)$, the model is based on variable returns to scale assumption.

When $(\delta_1, \delta_2, \delta_3) = (1, 0, 0)$, the model is based on decreasing returns to scale assumption.

When $(\delta_1, \delta_2, \delta_3) = (1, 1, 1)$, the model is based on Increasing returns to scale assumption.

So that, ∇, can be either 0 or 1.

The optimal $z = z_p^*$, is called the quantity of the efficiency for *p*th DMU, under corresponding DEA model, $z_p^* = 1$, we say DMU-p is efficient, other wise, it is Inefficient and it's the efficiency is a quantity of z_p^* .

3. Measurement Pitfalls of DEA Results

Now-a-days, DEA is an efficient tool for evaluating the performance of DMUs, But running it without enough knowledge may cause some obvious errors.

There are several applications and case studies implying many incorrect estimates in evaluation due to the inappropriate use of DEA.

[6,7] refer to another source of error from the computational point of view regarding the selection of a value for Epsilon in the model that can lead to incorrect evaluations. [8,9] point out a source of error regarding the number of selected DMU's and the number of inputs and outputs that can result in an overestimation.

But in this paper, we tackle the problem from a different point of view, which assumes that the computational considerations and modeling are applied properly, and that there is no misusing errors regarding to DEA models. As noted in previous description, we separate the important difficulties in evaluation to two classes:

The coverage of the production possibility set is constructed by only one DMU which dramatically effects the evaluation of all other DMU's.

This may cause a drastic tendency in the assigned weights of one or more factor(s) toward upper or lower bounds for all units.

The tendency can be so tough that removing one factor won't have any influence on the evaluated results.

This problem can be expressed in another way. Providing a compensatory environment is the most important characteristic of the DEA models that increases the level of competition between DMU's.

Since the frontier is made by just one unit (or even several special units) the competition changes to a monopoly.

As a result, the competition level comes down to zero and the units on the frontier influence the weights of all inputs and outputs more or less in the same way.

Another problem shows up when the enhancement suggestions are presented for inefficient DMU's. Practically it is impossible to ask an inefficient unit to increase its outputs 50 times, in order to reach to an efficient level, such as the units on the frontier. These solutions are neither applicable nor valid. Impractical

recommendations for the improvement of the inefficient units may lead to the reduction of the competition and finally giving up of the trend, which is in direct opposition with the aims of the evaluation.

One may think that these difficulties are due to the data collection procedure.

To clarify the fact that the measurement error may not be the source of the problems, we should mention that the nature of the most studies, such as bank branches, is caused these difficulties and that there is no error in the data sets. So, the remedy would be to revise the structure of the measurement model.

4. Methodology

In this section, we will concentrate mainly on the structure of the measurement model and will consequently propose a model based on this point of view.

In fact, in the DEA methodology the units are supposed to be homogenous and comparable. This assumption obviously holds not true when there exist some extraordinary DMU's. This situation imposes zero or a small constant as a weight for inputs and outputs of almost all DMU's.

The layer measurement model is trying to implement a policy in order to solve such conflicts, and reinforce the competition between the units.

The general scheme of the layer measurement model is to find the first efficient frontier via conventional DEA models, then remove all the unit(s) on this frontier and then run again.

This way, we can find the second, third and consequent efficient frontiers.

Through this process, we would be able to partition all the units into some finite disjoint sets corresponding to different frontiers (call it efficient layers).

All the calculations in this example and throughout the paper are done by GAMS/MINOS [11]. Then using a predetermined pattern, the layers will be classified into some categories.

For instance, we classify all the layers to: the best level category, the first level category, the second level category and so on. After categorizing the layers, we try to find the improvement plan for all DMU's in order to identify

(i) The best layer (top layer) of the category or
(ii) The worst layer (down layer) of the upper category.

As shown in the following table, all DMU's have been classified into different layers and, the layers set out in k categories.

Categories	Layers	Decision Making Units		
C_1 :	L_1 :	$DMU_{1,1}$	⋯	DMU_{1,l_1}
	M	**M**		**M**
	L_{k_1} :	$DMU_{k_1,1}$	⋯	$DMU_{k_1,l_{k_1}}$
C_2 :	L_{k_1+1} :	$DMU_{k_1+1,1}$	⋯	$DMU_{k_1+1,l_{k_1+1}}$
	M	**M**		**M**
	L_{k_2} :	$DMU_{k_2,1}$	⋯	$DMU_{k_2,l_{k_2}}$
M	**M**			
C_k :	$L_{k_{k-1}+1}$:	$DMU_{k_{k-1}+1,1}$	⋯	$DMU_{k_{k-1}+1,l_{k_{k-1}+1}}$
	M	**M**		**M**
	L_{k_k} :	$DMU_{k_k,1}$	⋯	$DMU_{k_k,l_{k_k}}$

In order to obtain the optimal weights for $DMU_p = (X_{i+1,p}, Y_{i+1,p})$ of C_{i+1}, we consider the following multipliers model. The model shows how the $DMU_p = (X_{i+1,p}, Y_{i+1,p})$ of C_{i+1} may reach to the best layer of its category.

$$Max \qquad UY_{i+1,p}$$
$$s.t. \tag{3}$$
$$VX_{i+1,p} = 1,$$
$$U Y_{ij} - VX_{ij} \leq 0, \qquad (j; DMU_j \in L_{k_i+1})$$
$$U \geq 0 \ \& \ V \geq 0$$

Where X_{ij} and Y_{ij} are the data vectors corresponding to inputs and inputs of DMU_j respectively.

The DMU's on the $(k_i + 1)$th layer imply the DMU's on the top layer in category $i+1$ and *U, V* are the variable vectors that are related to outputs and inputs respectively.

The dual model or the envelopment side is as follows:

Min q_p

s.t.   .

$$\sum_{j;DMU_{j}\in L_{k_{i}+1}}^{C_{i}} l_{j}X_{ij} + S_{I} = q_{p}X_{i+1,p} \qquad (4)$$

$$\sum_{j;DMU_{j}\in L_{k_{i}+1}}^{C_{i}} l_{j}Y_{ij} - S_{O} = Y_{i+1,p}$$

$$l_{j} \geq 0 \qquad j;DMU_{j} \in L_{k_{i}+1_{i}}$$

Running model 4 l_{j}^{*}'s, q_{p}^{*}, the slack and surplus variables S_{I}^{*} and S_{O}^{*} are obtained as the model optimal solutions.

Therefore the improvement approach to the top layer of $i+1$ category is obtained through the following formula:

$$(q_{p}^{*}X_{i+1,p} - S_{I}^{*}, Y_{i+1,p} + S_{O}^{p}) = (\sum_{j;DMU_{j}\in L_{k_{i}+1}}^{C_{i}} l_{j}^{*}X_{ij}, \sum_{j;DMU_{j}\in L_{k_{i}+1}}^{C_{i}} l_{j}^{*}Y_{ij})$$

These recommendations are to arrive at the upper layer of the current category. Now we suggest a model that would help $DMU_{p} = (X_{i+1,p}, Y_{i+1,p})$ of C_{i+1} to reach to C_{i}. In this case it is sufficient to choose the down layer of C_{i} as the final target.

The following model presents the multiplier side of the model for this proposes:

Max $UY_{i+1,p}$

s.t.

$$VX_{i+1,p} = 1, \qquad (5)$$

$$UY_{ij} - VX_{ij} \leq 0, \quad (j; \quad DMU_{j} \in L_{k_{i}})$$

$$U \geq 0 \ \& \ V \geq 0$$

Where X_{ij} and Y_{ij} are the vectors of inputs and outputs of DMU$_j$ respectively.

The DMU's in category k_{i} means all the units on the down layer are included in the category.

The vectors U and V are the vectors of the weight variables related to inputs and outputs respectively.

The corresponding envelopment model would thus be as follows:

Min q_p

s.t.   .

$$\sum_{j;DMU_{j}\in L_{k_{i}}}^{C_{i}} l_{j}X_{ij} + S_{I} = q_{p}X_{i+1,p} \qquad (6)$$

$$\sum_{j;DMU_{j}\in L_{k_{i}}}^{C_{i}} l_{j}Y_{ij} - S_{O} = Y_{i+1,p}$$

$$l_{j} \geq 0 \qquad j;DMU_{j} \in L_{k_{i}}$$

Running model 6 we can calculate the optimal value for l_{j}^{*}'s, q_{p}^{*} and also the slack and surplus variables S_{I}^{*} and S_{O}^{*}. So the improvement solution to reach to the down layer of ith category is given by the following formula:

$$(q_{p}^{*}X_{i+1,p} - S_{I}^{*}, Y_{i+1,p} + S_{O}^{p}) = (\sum_{j;DMU_{j}\in L_{k_{i}}}^{C_{i}} l_{j}^{*}X_{ij}, \sum_{j;DMU_{j}\in L_{k_{i}}}^{C_{i}} l_{j}^{*}Y_{ij})$$

Now to implement the evaluation process and have a keep going improvements we use the following algorithm:

5. Efficiency Improvement Algorithm

1. Let S be the set of all DMU's and $i \leftarrow 1$
2. Run the evaluation model for all units in S and form SE as the set of all efficient units.
3. $i \leftarrow i + 1, c_{i} \leftarrow Card(SE), L_{i} \leftarrow SE$.
4. $S \leftarrow S - L_{i}$.
5. If S is nonempty, go to stage 2 otherwise put $l \leftarrow i$ and continue.
6. $C_{1} \leftarrow L_{1} \ \mathbf{U} ... \mathbf{U} \ L_{k_{1}}$
7. For $p = 2,...,k$, $C_{p} \leftarrow L_{k_{p-1}+1} \ \mathbf{U} ... \mathbf{U} \ L_{k_{p}}$.
8. $p \leftarrow 1$.
9. $i \leftarrow 1$
10. $j \leftarrow 1$
11. If $p = 1$ solve model 4 and find the improvement solution to DMU_{ij}

    If $p > 1$ solve models 4 and 6 and find the improvement solution to DMU_{ij}
12. $j \leftarrow j + 1$ if $j < l_{i}$ go to 11 otherwise continue.
13. $i \leftarrow i + 1$ if $i < k_{p}$ go to 10 otherwise continue.
14. $p \leftarrow p + 1$ if $p < l_{i}$ go to 9 otherwise continue.
15. In this stage, the improvement solutions for all DMU's are produced. In the next run, the algorithm will be repeated.
16. If the system's life is finished this process would be stopped, otherwise go to stage 1.

6. Computational Aspects of the Algorithm

Due to the fact that the number of the DMU's is finite and each time the new nonempty layer (because using DEA models) will be definite, the algorithm convergence is guaranteed.

From the computational point of view, this algorithm is divided into two basic phases.

The first phase partitions the DMU's to efficient layers and the second accounts for the process of performance analysis and stands for improvement solutions.

The main computation efforts would thus be in the first phase when we run the model for all decision making units individually in presence of different number of units such as n, $n - Card(L_1)$, $n - \sum_{i=1}^{2} Card(L_i)$,

$n - \sum_{i=1}^{3} Card(L_i)$, and ....

In the second phase of the algorithm, we deal with small models and the computation is not time consuming.

7. Conclusion

In this paper, we examined the problems caused by some extraordinary DMU's and their influence in replacing the competition with a monopoly in the process of evaluation. To this end, an algorithm was presented screening DMU's into efficient layers, through a computational process. Then the performance improvement solutions leading to a better competition were introduced. The fact that in the case of bank branches, there are differing in the wide range and there are always some branches with particular specifications, makes the necessity of our classification clear. This way there is no monopoly in the evaluating process, and the way out for improvement will be practical. Hence the presented model can wieldy be used.

References

[1] Charnes, A., Cooper, W.W., Rhodes, E., *Measuring the efficiency of decision making units*, European Journal of Operation Research, No. 2, 1978, pp.429-444.

[2] Banker, R.D., Charnes, A., Cooper, W.W., *Some models for estimating technical and scale inefficiencies in data envelopment analysis*, Management Science, No. 10, 1984, pp.78-092.

[3] Charnes, A., Cooper, W.W., Lewin, A.Y., Seiford L. M., *Data Envelopment Analysis: Theory, Methodology, and Applications*, Boston: Kluwer Academic Publishers.1994.

[4] Cooper, W.W., Tone, K., Seiford, L.M., *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software,* Kluwer Academic Publishers. 2000.

[5] www.DEAzone.com

[6] Alirezaee, M.R., Howland, M., van de Panne, C., *Sampling size and efficiency bias in Data Envelopment Analysis*, Journal of Applied Mathematics & Decision Sciences, No. 2, 1998, pp. 51-64.

[7] Alirezaee, M.R., Howland, M., van de Panne, C., *Efficiency bias in DEA: A simulation study on a large scale bank branches,* Journal of Economics and Management, No. 37, 1998, pp. 89-105.

[8] Mehrabian, S., Jahanshahloo, G.R., Alirezaee, M.R., Amin, G.R., *An assurance interval of the non-Archimedean Epsilon in DEA models*, Operations Research, No. 48, 2000, pp. 344-347.

[9] Alirezaee, M.R., *The overall assurance interval for the non-Archimedean Epsilon in DEA models; a partition base algorithm,* Applied Mathematics and Computations, No. 164, 2005, pp. 667-674.

[10] Farrell, M.J. *The measurement of productive efficiency*, Journal of Royal Statistical Society, Ser. A, III:253-290, 1957.

[11] Brooke, A.D., Kendrick A., Meeraus A., Raman R., *GAMS a user guide*, GAMS Development Corporation, 1998.