Measurement 65 (2015) 29–40

Contents lists available at ScienceDirect

Measurement

journal homepage: www.elsevier.com/locate/measurement

Two alternative approaches for selecting performance measures in data envelopment analysis

Mehdi Toloo *, Tomáš Tichý

Faculty of Economics, Technical University of Ostrava, Ostrava, Czech Republic

ARTICLE INFO

Article history:
Received 10 August 2014
Received in revised form 24 November 2014
Accepted 23 December 2014
Available online 8 January 2015

Keywords:
Decision making
Data envelopment analysis
Multiplier and envelopment forms
The rule of thumb
Selective measures
Banking industry

ABSTRACT

Data envelopment analysis seeks a frontier to envelop all data with data acting in a critical role in the process and in such a way measures the relative efficiency of each decision making unit in comparison with other units. There is a statistical and empirical rule that if the number of performance measures is high in comparison with the number of units, then a large percentage of the units will be determined as efficient, which is obviously a questionable result. It also implies that the selection of performance measures is very crucial for successful applications. In this paper, we extend both multiplier and envelopment forms of data envelopment analysis models and propose two alternative approaches for selecting performance measures under variable returns to scale. The multiplier form of selecting model leads to the maximum efficiency scores and the maximum discrimination between efficient units is achieved by applying the envelopment form. Also individual unit and aggregate models are formulated separately to develop the idea of selective measures. Finally, in order to illustrate the potential of the proposed approaches a case study using a data from a banking industry in the Czech Republic is utilized.

1. Introduction

Data envelopment analysis (DEA), as a well-known non-parametric mathematical approach, evaluates the relative efficiency score of a set of homogeneous decision making units (DMUs), which utilize the multiple inputs to produce multiple outputs. In managerial applications, DMUs may include banks, department stores and supermarkets, and extend to car makers, hospitals, schools, public libraries and even portfolios. In engineering, DMUs may take such forms as airplanes or their components such as jet engines. Basic DEA models measure the efficiency of a DMU by maximizing the ratio of the weighted sum of its outputs to the weighted sum of its inputs, based on the condition that this ratio is less than or equal to one for all DMUs. The large number of studies has been accomplished in DEA which show that this is an outstanding and straightforward methodology for modeling operational process in performance evaluations. Nowadays, DEA is becoming a very important analysis tool and research method in management science, operational research, system engineering, decision analysis, etc.

Cooper et al. [14] considered a rough rule of thumb, which expresses the relation between the number of DMUs and the number of performance measures to have a reliable results. In other words, when the number of performance measures is high in comparison with the number of DMUs, then most DMUs are classified as efficient, which is not logical. Therefore in this paper we propose an approach to deal with this issue. In particular, we extend standard models so that important measures (i.e., inputs and outputs) can be identified, their total number decreased and complex result optimized.

* Corresponding author at: Sokolska trida 33, 701 21 Ostrava 1, Ostrava, Czech Republic. Tel.: +420 597 322 009.

E-mail addresses: mehdi.toloo@vsb.cz, m_toloo@yahoo.com (M. Toloo).

http://dx.doi.org/10.1016/j.measurement.2014.12.043

The rest of the paper is organized as follows: Section 2 summarizes the rule of thumb in DEA. In Section 3, the multiplier and envelopment forms of BCC model is presented. Section 4 introduces a new approach to choose selective measures. Section 5 illustrates the capabilities of our proposed approach by using a real data set of banking industry in the Czech Republic. Conclusions and further remarks are provided in the last section.

2. The rule of thumb in DEA

Suppose there are n DMUs that consume m inputs to produce s outputs. If a performance measure (input/output) is added or deleted from consideration, it will influence the relative efficiencies. Empirically, when the number of performance measures is high in comparison with the number of DMUs, then most DMUs are evaluated efficient so that such results cannot be regarded as reliable. A rough rule of thumb expresses the relation between the number of DMUs and the number of performance measures (see Cooper et al. [14] for more details):

$$n \geqslant \max\{3(m+s), m \times s\}$$

Table C.1 in Appendix C practically represents the number of DMUs and the number of performance measures applied in some studies (for more details see [1,4–7,10–13,15–53]).

There are some cases that the number of performance measures and DMUs does not satisfy the rule of thumb. To rectify this issue, in many scenarios we select some performance measures in a manner which complies with the rule of thumb and imposes progressive effect on the efficiency scores. These selected inputs and outputs are known as "selective measures" implies

Establishing criteria among a set of suggested standards is an important issue for the decision maker. For example, consider the problem of evaluating 50 branches of a bank; in such a setting, the manager may practically run into more than 25 inputs and 30 outputs. Apparently, the total number of measures, i.e. 55, and DMUs do not satisfy the rule of thumb and subsequently to have an acceptable efficiency scores some performance measures must be omitted. To make the problem easier, suppose that the manager pre-selected three inputs, e.g. employees, expenses and space, and three outputs, e.g. loans, profits and deposits. With this assumptions, if the manager wants to select 2 out of 22 remaining inputs and 1 out of 27 remaining outputs and also consider all possible combinations of performance measures, then an optimization problem must be solved at most $196350\left(=50 \times \binom{22}{2} \times \binom{27}{1}\right)$ times, which shows that this approach is illogical.

In this paper, we propose an approach to deal with the selective measures under variable returns to scale (VRS) assumption. However, in some situations, some special measures must appear in evaluation model, so that the manager's concern comes true. These measures are named "fixed measures" and we also consider them in developing our approach.

3. The BCC model

Charnes et al. [8] proposed an innovative mathematical approach for evaluating the relative efficiencies among DMUs with multi-input and multi-output as a linear programming model. This approach, which is referred to as CCR (**C**harnes, **C**ooper and **R**hodes), assumed constant returns to scale (CRS) technology. Banker et al. [3] suggested a new model, which is known as BCC (**B**anker, **C**harnes and **C**ooper), to deal with VRS situation.

Consider a set of n decision making units (DMUs), each consuming various amounts of m inputs to produce s outputs. Let $\mathbf{x}_j = (x_{1j},\ldots,x_{mj})^T$ and $\mathbf{y}_j = (y_{1j},\ldots,y_{sj})^T$ represent the input and output vectors for DMU_j $(j = 1,\ldots,n)$, respectively. There are two BCC forms which are mutually dual: multiplier and envelopment.

Multiplier form of BCC

$$\begin{aligned} \max z_1 &= \sum_{r=1}^{s} u_r y_{ro} + u_o \\ \text{s.t.}\quad & \sum_{i=1}^{m} v_i x_{io} = 1 \\ & \sum_{r=1}^{s} u_r y_{rj} + u_0 - \sum_{i=1}^{m} v_i x_{ij} \leqslant 0 \quad \forall j \\ & v_i \geqslant \varepsilon \quad \forall i \\ & u_r \geqslant \varepsilon \quad \forall r \end{aligned} \tag{1}$$

Envelopment form of BCC

$$\begin{aligned} \max z_2 &= \theta - \varepsilon\left(\sum_{i=1}^{m} s_i^x + \sum_{r=1}^{s} s_r^y\right) \\ \text{s.t.}\quad & \sum_{j=1}^{n} \lambda_j x_{ij} + s_i^x = \theta x_{io} \quad \forall i \\ & \sum_{j=1}^{n} \lambda_j y_{rj} - s_r^y = y_{ro} \quad \forall r \\ & \sum_{j=1}^{n} \lambda_j = 1 \\ & \lambda_j \geqslant 0\,\forall j,\, s_i^x \geqslant 0\,\forall i,\, s_r^y \geqslant 0\,\forall r \end{aligned} \tag{2}$$

where $\text{DMU}_o = (\mathbf{x}_o, \mathbf{y}_o)$ is being evaluated $(o \in \{1, 2, \ldots, n\})$.

In the multiplier form of BCC model (1), $\mathbf{u} = (u_1, \ldots, u_s)$ and $\mathbf{v} = (v_1, \ldots, v_m)$ are the set of output and input weights (or multipliers, dual variables or shadow prices), respectively, and u_0 is a variable with free in sign, i.e. $u_0 \in (-\infty,+\infty)$. Charnes et al. [9] added ε, which is called the non-Archimedean epsilon, to the CCR model to have positive weights. It is important to set a suitable value for the non-Archimedean epsilon ε. Amin and Toloo [2] designed a polynomial-time algorithm to obtain an appropriate value for epsilon. This model maximizes the relative efficiency score of DMU_o subject to the condition that the similar score for all DMUs must be less than or equal to one. Suppose that model (1) is solved and the optimal solution $(\mathbf{u}^*,\mathbf{v}^*,u_0)$ is at hand, DMU_o is efficient when $z_1^* = 1$ and otherwise is inefficient. As a result, DEA models can categorize all DMUs into two different groups: efficient and inefficient. Toloo et al. [47] suggested an efficient approach to find an initial basic feasible solution for the multiplier form of BCC model and practically showed it decreases required computations by at least one half.

On the other hand, the envelopment form of BCC model (2) seeks a (virtual) unit $\left(\sum_{j=i}^{n}\lambda_j^*\mathbf{x}_j, \sum_{j=1}^{n}\lambda_j^*\mathbf{y}_j\right)$ that guarantees at least the output level $\mathbf{y}_o$ of DMU_o in all components $\left(\sum_{j=1}^{n}\lambda_j^* y_{rj} \geqslant y_{ro},\ r = 1,\ldots,m\right)$, while reducing the input vector $\mathbf{x}_o$ proportionally to a value as small as possible $\left(\sum_{j=1}^{n}\lambda_j^* x_{ij} \leqslant \theta^* x_{io},\ i = 1,\ldots,m\right)$. If $z_2^* = 1$ (or equivalently $\theta_o^* = 1$, $\forall i\ s_i^{x^*} = 0$ and $\forall r\ s_r^{y^*} = 0$), then DMU_o is efficient.