Computers & Operations Research 37 (2010) 99 – 107

Contents lists available at ScienceDirect

Computers & Operations Research

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

Improving the discrimination power and weights dispersion in the data envelopment analysis

Hasan Bal*, H. Hasan Örkcü, Salih Çelebioğlu

Gazi University, Arts and Science Faculty, Department of Statistics, 06500, Ankara, Turkey

A R T I C L E   I N F O

Available online 9 April 2009

Keywords:
Data envelopment analysis
Goal programming
Discrimination power
Unit invariance

A B S T R A C T

Data envelopment analysis (DEA) has been a very popular method for measuring and benchmarking relative efficiency of peer decision making units (DMUs) with multiple input and outputs. Beside of its popularity, DEA has some drawbacks such as unrealistic input–output weights and lack of discrimination among efficient DMUs. In this study, two new models based on a multi-criteria data envelopment analysis (MCDEA) are developed to moderate the homogeneity of weights distribution by using goal programming (GP). These goal programming data envelopment analysis models, GPDEA-CCR and GPDEA-BCC, also improve the discrimination power of DEA.

1. Introduction

Data envelopment analysis (DEA) is a very popular mathematical programming technique which is used to evaluate relative efficiency of decision making units (DMUs) and has been developed by Charnes et al. [1]. Nowadays, DEA has been one of the fastest growing areas of operational research and management science. DEA has become an increasingly powerful approach to analyze the efficiency of public and private sector organizations. However, some problems have also appeared as the applications of DEA advance. Two inter-related problem that have long been known are the lack of discrimination power and the unrealistic weight dispersion. The lack of discriminating power problem occurs when the number of DMUs under evaluation is not large enough compared to the total number of inputs–outputs. In this situation, classical DEA models often yield solutions that identify too many DMUs as efficient. To better the lack of disrimination power of DEA, some DEA approaches; such as super efficiency, multiple criteria DEA and cross efficiency, have been proposed in DEA literature. However, in some cases it is also possible to meet the infeasibility problems in super efficiency models and complexity of multiple objectives for multiple criteria DEA models [2–5]. The cross efficiency approach is a useful technique developed by Sexton et al. [6] so as to rate the DMUs by using the cross evaluation scores computed as related to all DMUs and hence identify the best DMUs [7]. Although the cross efficiency method has a widespread usage, it has also some deficiencies arising from the classical DEA. As stated by Doyle and Green [8], the non-uniqueness, i.e., having multiple solutions to optimal weights in DEA decreases the usefulness of cross efficiency method. The problem of unrealistic weight dispersion for DEA occurs when some DMUs are rated as efficient because of input and output weights have the extreme or zero values. To overcome the unrealistic weight dispersion problem, weight restriction techniques such as cone ratio envelopment, assurance region and value efficiency have been proposed in DEA literature [9–19]. However, these techniques are dependent on the measurement units of inputs–outputs and may also give infeasible solutions for weights. On the other hand, these methods incorporate additional constraints to the model, and hence make harder to solve the problem. It has been denoted that, in presence of additional homogeneous weight restrictions, which include absolute weight bounds, some DEA models may identify the maximum relative efficiency of DMUs incorrectly [14,20]. Recently, Bal et al. [21] proposed a CVDEA model which incorporates the coefficient of variation (CV) for input and output weights reducing the number of efficient DMUs and produces more homogeneous weight dispersion for inputs–outputs. However, the CVDEA model should be used providently since it does not preserve the unit-invariance and linearity properties.

In this study, new and easy-to-use models (GPDEA-CCR and GPDEA-BCC) are presented and generally more balanced input–output weight dispersions are obtained with respect to the basic DEA models (CCR and BCC) and also reduced the number of efficient DMUs without any additional constraints on weights. In addition, these GPDEA models have also unit-invariance property. In order to show the improvement of the dispersal of input–output weights and the increasing discrimination power for our suggested models, we have used a hypothetical, a literature and a real world

* Corresponding author. Tel.: +90 312 202 14 65; fax: +90 312 212 2279.
E-mail address: hasanbal@gazi.edu.tr (H. Bal).

0305-0548/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.cor.2009.03.028

data. In addition, the applicability of the new models is tested by a simulation study.

The study is organized as follows. In Section 2, the basic DEA models are briefly presented. In Section 3, the multi-criteria data envelopment analysis (MCDEA) model is presented and explained. In Section 4, MCDEA model formulations as a weighted goal programming (GPDEA-CCR and GPDEA-BCC) are explained. In Section 5, the basic CCR, GPDEA-CCR, BCC, and GPDEA-BCC are applied to a hypothetical and a literature data and their solutions are compared. In Section 6, the applicability of the new models is tested by a simulation study. In Section 7, the basic DEA and GPDEA models are applied to an OECD data. Lastly, in Section 8, a summary of the research and its results are provided.

2. Data envelopment analysis

DEA appraises the relative efficiency of homogeneous units by considering multiple inputs and outputs. Inputs can be resources used by a DMU and outputs can be products produced or performance measures of DMU. The efficiency is indicated as a ratio of the weighted sum of outputs to the weighted sum of inputs. DEA has been extensively used to compare the efficiencies of non-profit and profit organizations such as schools, hospitals, shops, banks and other environments where there are relatively homogeneous DMUs [22].

Assuming that there are n DMUs each with m inputs and s outputs, the relative efficiency of a particular DMU_0 is obtained by solving the following fractional programming problem:

$$
\begin{aligned}
w_o = \mathrm{Max} \quad & \frac{\sum_{r=1}^{s} u_r y_{ro}}{\sum_{i=1}^{m} v_i x_{io}} \\
& \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \leqslant 1 \quad j = 1,2,\ldots,n \\
& u_r \geqslant 0, \quad r = 1,2,\ldots,s \\
& v_i \geqslant 0, \quad i = 1,2,\ldots,m
\end{aligned} \tag{1}
$$

where j is the DMU index, $j=1,\ldots,n$; r is the output index, $r=1,\ldots,s$; i is the input index, $i=1,\ldots,m$; y_{rj} is the value of the rth output for the jth DMU, x_{ij} is the value of the ith input for the jth DMU, u_r is the weight given to the rth output; v_i is the weight given to the ith input, and w_o is the relative efficiency of DMU_0, the DMU under evaluation.

In this model, the measure of efficiency of any DMU is obtained as the maximum of a ratio of weighted output to weighted input subject to the condition that similar ratios for every DMU are less than or equal to unity. According to this model, $w_o = 1$ means that DMU_0 is efficient relative to other units. If $w_o < 1$, then this unit is inefficient.

This fractional program, well known as CCR model, can be converted into a linear programming problem where the optimal value of the objective function indicates the relative efficiency of DMU_0. Hence the reformulated linear programming problem is as follows:

$$
\begin{aligned}
w_o = \mathrm{Max} \quad & \sum_{r=1}^{s} u_r y_{ro} \\
& \sum_{i=1}^{m} v_i x_{io} = 1 \\
& \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leqslant 0 \quad j = 1,2,\ldots,n \\
& u_r \geqslant 0, \quad r = 1,2,\ldots,s \\
& v_i \geqslant 0, \quad i = 1,2,\ldots,m
\end{aligned} \tag{2}
$$

In this model, the weighted sum of the inputs for DMU_0 is forced to 1, thus allowing for the conversion of fractional programming problem into a linear programming problem which can be solved using any linear programming software. Similarly, the model of returns to scale for the DEA, namely BCC, can be given as follows [22,23]:

$$
\begin{aligned}
w_o = \mathrm{Max} \quad & \sum_{r=1}^{s} u_r y_{ro} + c_o \\
& \sum_{i=1}^{m} v_i x_{io} = 1 \\
& \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} + c_o \leqslant 0 \quad j = 1,2,\ldots,n \\
& u_r \geqslant 0, \quad r = 1,2,\ldots,s \\
& v_i \geqslant 0, \quad i = 1,2,\ldots,m \\
& c_o \text{ free in sign}
\end{aligned} \tag{3}
$$

where c_o indicates returns to scale. w_o, u_r $(r = 1,2,\ldots,s)$ and v_i $(i = 1,2,\ldots,m)$ are defined as in the CCR model.

The solution to these models assigns the value 1 to all efficient DMUs. The super efficiency concept is proposed for all efficient DMUs when there are more than one efficient DMU. One of the super efficiency models for ranking efficient DMUs in DEA was introduced by Andersen et al. [2]. This method enables an extreme efficient unit o to achieve an efficiency score greater than one by removing the oth constraint in the envelopment linear programming formulation.

3. Multiple criteria DEA model

Model (2) can be expressed equivalently in the form given by Li and Reeves [4].

$$
\begin{aligned}
\min \quad & d_o \left(\text{or } \max \sum_{r=1}^{s} u_r y_{ro} \right) \\
& \sum_{i=1}^{m} v_i x_{io} = 1 \\
& \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} + d_j = 0 \quad j = 1,2,\ldots,n \\
& u_r \geqslant 0, \quad r = 1,2,\ldots,s \\
& v_i \geqslant 0, \quad i = 1,2,\ldots,m \\
& d_j \geqslant 0, \quad j = 1,2,\ldots,n
\end{aligned} \tag{4}
$$

where d_o is the deviation variable for DMU_0 and d_j is the deviation variable of DMU_j. The quantity d_o, which is bounded by the interval $(0,1]$, can be regarded as a measure of inefficiency. Under this model, DMU_0 is efficient if and only if $d_o = 0$ or $\sum_{r=1}^{s} u_r y_{r_o} = 1$. If DMU_0 is not efficient, its efficiency score is $1 - d_o$. We shall call model (2) the basic DEA model. We specify that the basic DEA method minimizes DMU_0's inefficiency, as measured by d_o, under the constraint that the weighted sum of the outputs is less than or equal to weighted sum of the inputs for each DMU.

The lack of discriminating power of the basic DEA can be overcomed by using a single objective function in place of multiple and more discriminating objective functions, as proposed by Li and Reeves [4]. A multiple criteria data envelopment analysis model formulation with the minmax and minsum criteria, which minimizes a deviation variable, d_o, rather than maximizing the efficiency score,