Expert Systems with Applications 42 (2015) 7759–7767

Contents lists available at ScienceDirect

Expert Systems with Applications

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

Multi-period data envelopment analysis based on Chebyshev inequality bounds

Seyed Hossein Razavi Hajiagha [a,*], Shide Sadat Hashemi [b], Hannan Amoozad Mahdiraji [b], Jamshid Azaddel [c]

[a] Department of Management, Khatam Institute of Higher Education, Tehran, Iran
[b] Department of Management, Kashan Branch, Islamic Azad University, Kashan, Iran
[c] Payame Noor University, Khalkhal, Iran

ARTICLE INFO

Article history:
Available online 9 June 2015

Keywords:
Data envelopment analysis
Multi-period efficiency
Chebyshev inequality bounds
Two-stage efficiency approximations

ABSTRACT

Data envelopment analysis is a cross-sectional approach to evaluate the relative efficiency of a set of homogeneous units in a single time point; nonetheless, organizational units have been performing continuously over a period of time; hence, their performances are considered within this period. Cumulating inputs and outputs over the time periods provide an unnecessary compensating impact, making the efficiency appraisal unrealistic. To avoid this negative impact of data accumulation, a two-stage approach on the basis of Chebyshev inequality bounds is proposed to find interval efficiency of decision making units (henceforth DMUs). The proposed method is applied in a real case encompassing 115 bank branches over 6 periods of time. This application indicated the significant cautious approach of the proposed method in multi-period data envelopment analysis (hereafter DEA).

1. Introduction

Data envelopment analysis, initially introduced by Charnes, Cooper, and Rhodes (1978) is an accepted and widely employed framework to analyze the relative efficiency of a set of DMUs, using m inputs to produce a set of s outputs. DEA is extended based on economic foundations of production possibility sets and seeks a production frontier to measure the relative efficiency of DMUs (Charnes, Cooper, Lewin, & Seiford, 1994; Førsund, Kittelsen, & Krivonozhko, 2009). Recent investigations by Emrouznejad, Parker, and Tavares (2008) and Liu, Lu, Lu, and Lin (2013) have shown a large variety of applications using DEA for measuring and improving the efficiency.

Classical DEA models can be considered as cross-sectional analysis. Admittedly, the performance of DMUs is compared with a particular point of time. Contrariwise, comparing the performance of DMUs over several periods of time is considerable, knowing as longitude or time series analysis (Charnes et al., 1994; Ramanathan, 2003). This problem is generally called multi-period DEA. Some implications of multi-period DEA can be found at banks (Kao & Liu, 2014) and insurance companies (Kao & Hwang, 2014).

In a nutshell, multi-period DEA deals with inputs and outputs fluctuation among DMUs. Classically, stochastic DEA models can be applied for handling this fluctuation, where inputs and outputs are assumed to follow certain statistical distributions. Cooper, Huang, and Susan (2011), Wei, Chen, and Wang (2014), and Branda (2015) are some of the latest researches regarding stochastic DEA; nevertheless, a different perspective is followed in multi-period DEA models in which, data are observed in different time points and are captured in the form of time series. A conventional approach for dealing with multiple periods is to aggregate the data of different periods in a single data point and to ignore the specific situation of each period (Charnes, Clark, Cooper, & Golany, 1985).

To avoid this simplification mode, several methods are proposed for time series DEA problems. One of the first approaches in DEA analysis of multiple time periods is window analysis; where a moving average pattern of analysis is applied (Caves, Christensen, & Diewert, 1982). Actually, the performance of a DMU is compared with its performance in other periods, and with other DMUs' performance in the same period (Ramanathan, 2003). However, as mentioned by Charnes et al. (1994), choosing the number of time periods in the window is really a controversial issue. Alongside with, another classic approach is Malmquist-type indexes of productivity (Färe & Grosskopf, 1996). Beyond their usefulness, Kao and Liu (2014) pointed that these methods "do not take into account an aggregated measure of efficiency for multiple-period production systems".

* Corresponding author.
E-mail addresses: s.hossein.r@gmail.com (S.H. Razavi Hajiagha), Shide_hashemi@yahoo.com (S.S. Hashemi), h.amoozad@st.atu.ac.ir (H. Amoozad Mahdiraji), j.azad63@yahoo.com (J. Azaddel).

http://dx.doi.org/10.1016/j.eswa.2015.06.008
0957-4174/

Classes of dynamic DEA models are also extended for multi-period problems. Their main advantage is the ability to account the effect of carry-over activities between two consecutive terms. Dynamic DEA models were initially introduced by Färe and Grosskopf (1996), subsequently were developed by Nemoto and Goto (1999, 2003) and Sueyoshi and Sekitani (2005), as well as Bogetoft et al. (2008). On the other hand, it is worth noting here that, as a weakness these models need a perfect foresight regard to input costs, while Thompson, Langemeier, Lee, Lee, and Thrall (1990), besides Thompson, Dharmapala, and Thrall (1995) believed that exact input cost is not determined even in a given period.

Sengupta (1995, 1999) developed different types of dynamic DEA models, via which various possible scenarios of aggregating input costs were considered over the time. In these models, an optimal level is determined for inputs and the overall efficiency is defined as the ratio of actual used inputs over optimal expected inputs. Sengupta (1995) model assumed that inputs' future prices are determined exactly, while Sengupta (1999) extended his initial model to incorporate the uncertainty of inputs' future prices for measuring overall efficiency. As previously declared, the main restriction of dynamic DEA models is their dependency on knowledge about input prices, especially in the future, imposing an additional uncertainty to the models.

The multi-period DEA problems can be imagined in the context of network DEA models, whilst classic network DEA models (Kao, 2008, 2014a, 2014b) considered DMUs internal structures and the relations among the subunits of DMUs, the multi-period DEA model can be considered as a network of time frames where a DMU performed continually in a time horizon whereas the aim of the model is to evaluate the relative efficiency of DMUs in this time-based network. A similar conceptualization of multi-period DEA in the form of network DEA is considered by Kao and Liu (2014).

Park and Park (1995) presented a multi-period data envelopment analysis (MDEA) model upon the concept of Debreu–Farrell's technical efficiency. The MDEA model relies on finding the efficiency of DMUs in different periods whereas a DMU is called full efficient if it gains full efficiency in all periods.

Amirteimoori and Kordrostami (2010) defined the aggregated efficiency of a DMU as a convex combination of individual period efficiencies and developed a model to find total and period efficiencies of DMUs. Kao and Liu (2014) proposed a model for multi-period efficiency evaluation, through which the overall and period efficiencies of a DMU are calculated simultaneously, while the overall efficiency is defined as the weighted average of the period efficiencies. Kao and Hwang (2014) applied the idea of overall efficiency as the weighted average of period efficiencies in two-stage network production systems.

Whilst different methods presented valuable views toward finding multi-period efficiency of a set of DMUs over a period of time, the main drawback to these models is ignoring the individual input and output variances during the time. Since the performance of DMUs is evaluated in multiple periods, the inputs and outputs of these DMUs are treated differently over the period of time, remark that these marked differences should be considered in their efficiency evaluation. Different methods usually accumulate inputs and outputs of considered periods to evaluate relative efficiency of units with cumulative data. This accumulation results in an unnecessary compensating effect ignoring the impact of inputs and outputs fluctuation on efficiency. As a case in point, if one of the outputs of a given DMU is increased in a specific time period, while this output is decreased dramatically in another time period, without considering the variance of this output engenders an unrealistic approximation of efficiency. The aim of this paper is to extend a DEA model incorporating the variability of input and output measures directly in the model.

The reminder of paper is organized as follows. Section 2 describes the perceived problem of evaluating bank branches efficiency during a period of 12 months. The proposed algorithm is explained in Section 3. Numerical results are presented in Section 4. Finally, the paper is concluded in Section 5.

2. Problem description

Banks play an important role in the economic system of countries. This importance makes them an interesting subject of DEA applications. As Emrouznejad et al. (2008) and Liu et al. (2013) surveys' illustrated, financial institutions including banks are the most implicational field of DEA.

The noted problem in this paper deals with evaluating the efficiencies of 115 branches of HNI, a corporate bank in Iran. A set of 5 inputs and 4 outputs are identified to appraise the branches efficiency. These inputs and outputs are specified upon Berger and Humphrey (1997) and Luo, Bi, and Liang (2012). Table 1 depicts the input and output measures. Among the inputs, the ratio of non-current to total receivables is an undesirable output. Different methods are proposed encompassing undesirable outputs. Färe, Grosskopf, Lovell, and Pasurka (1989) and Färe, Grosskopf, and Tyteca (1996), as well as Tyteca (1997) modeled and developed the concept of hyperbolic output efficiency measure to deal with undesirable outputs in terms of an equiproportionate increase in desirable outputs and decrease in undesirable ones. Considering undesirable outputs as inputs is applied in different DEA studies (Matthews, 2013). The main strength of this approach is its easiness of use. In conjunction with the main advantage this approach does not rely on any assumption about data structure and shape of linear transformation.

In this study, data are gathered from the second half of the financial year 2013, from 1 July 2013 to 1 January 2014. Accordingly, there are a set of 6 input/output matrices of the size of 115×9. According to these matrices, a set of 9 time series data, each for one of the inputs or outputs is constructed. The aim of the problem is to assess the efficiency of bank branches in aforementioned time horizon.

The attended data in this problem are observed values of inputs and outputs in discrete time points forming different time series for every input and output of each DMU.

3. Problem formulation

Generally, suppose that there are a set of n DMUs which received an m-dimensional input vector X to produce an s-dimensional output vector Y. The efficiency of these DMUs will be evaluated in a time horizon of T periods. Let

- $x_j^t = \left(x_{1j}^t, x_{2j}^t, \ldots, x_{mj}^t\right)$ be the m-dimensional input vector of $DMU_j,\ j = 1, 2, \ldots, n$;
- $y_j^t = \left(y_{1j}^t, y_{2j}^t, \ldots, y_{sj}^t\right)$ be the s-dimensional output vector $DMU_j,\ j = 1, 2, \ldots, n$;

Table 1
Input and output measures.

Inputs	Outputs
Personnel costs	Sum of deposits
Current and administrative costs	Loans
Cost accounts	Securities
Renting cost	Branch income
The ratio of non-current to total receivables	