Expert Systems with Applications 38 (2011) 9147–9156

Contents lists available at ScienceDirect

Expert Systems with Applications

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

Using fuzzy super-efficiency slack-based measure data envelopment analysis to evaluate Taiwan's commercial bank efficiency

Bo Hsiao [a,*], Ching-Chin Chern [a,1], Yung-Ho Chiu [b], Ching-Ren Chiu [c]

[a] Department of Information Management, National Taiwan University, Taipei 10617, Taiwan
[b] Department of Economics, Soochow University, Taipei 10048, Taiwan
[c] Department of Business Administration, National Taiwan University of Science and Technology, Taipei 10607, Taiwan

ARTICLE INFO

Keywords:
DEA
Fuzzy set
Fuzzy SBM
Fuzzy Super SBM
Banking performance evaluation

ABSTRACT

Data envelopment analysis (DEA) mainly utilizes envelopment technology to replace production function in microeconomics. The input and output of decision making units (DMUs) are projected into the attributes to evaluate or measure their performance. However, if the inputs and outputs are linguistically termed or are fuzzy-numbered, conventional DEA can not easily measure the performance. Therefore, we propose the use of a fuzzy super-efficiency slack-based measure DEA to analyze the operational performance of 24 commercial banks facing problems on loan and investment parameters with vague characteristics. After our analysis, we find that the fuzzy slack-based measure of efficiency (Fuzzy SBM)/fuzzy super-efficiency slack-based measure of efficiency (Fuzzy Super SBM) can not only effectively characterize uncertainty, but also have a higher capability to evaluate bank efficiency than the conventional Fuzzy DEA approach.

1. Introduction

Most studies on bank efficiency to date have mainly focused on the economies of scale and scope (Berger, Hanweck, & Humphrey, 1987; Berger & Humphrey, 1991; Hunter & Timme, 1986; McAllister & McManus, 1993), total productivity (Aly, Grabowski, Pasurka, & Rangan, 1990; Favero & Papi, 1995; Fukuyama, Guerra, & Weber, 1999; Grabowski, Rangan, & Rezvanian, 1993; Schaffnit, Rosen, & Paradi, 1997; Zaim, 1995), comparisons on parametric and non-parametric programming approaches (Bhattacharyya, Lovell, & Sahay, 1997; Chen & Yeh, 2000; Ferrier & Lovell, 1990; Huang and Wang, 2002; Resti, 1997), the efficiency effect (Barr, Seiford, & Siems, 1994; Casu & Molyneux, 2003; Cebenoyan, Cooperman, & Register, 1993; Chang, 1999; DeYoung & Hasan, 1998; Elyasiani, Mehdian, & Rezvanian, 1994). There are two issues concerning banks efficiency and risk. One treats risk as exogenous, which allows for the analysis of the efficiency effects (Barr et al., 1994; Cebenoyan et al., 1993; Elyasiani et al., 1994). The works cited above each describe how the level of efficiency is significantly correlated with various risk indicators. The other issue treats risk as endogenous, which allows for the analysis of bank efficiency (e.g., Chang, 1999). These studies show that when risk factors are taken into account, the risk factors have the greatest influence on efficiency estimates and rankings. The authors also find that increasing risk (such as with problem loans) tends to decrease efficiency.

In these studies, the great majority of input and output variables focus on crisp data, and they cannot easily measure linguistic terms. Kao and Liu (2004) argue that we cannot gather crisp data because respondents cannot easily decide on the values by means of intuition. Sengupta (1992) is considered the pioneer in solving these kinds of issues; he proposed a fuzzy objective function and constraints using results from Zimmermann (1976, 1996). Though fuzzy concept[2] has been applied to many fields and easy to adopt, as such, some research works attempted to solve the problems with the limitation of conventional data envelopment analysis (DEA) by introducing the fuzzy concepts (Cooper, Park, & Yu, 1999; Despotis & Smirlis, 2002; Guo & Tanaka, 2001; Jahanshahloo, Soleimani-damaneh, & Nasrabadi, 2004). Cooper et al. (1999) addressed the problem of imprecise data in DEA in its general form. Furthermore, Despotis and Smirlis (2002) calculated upper and lower bounds for the radial efficiency scores of DMUs with interval data. These studies successful adopt fuzzy concepts in DEA, they contributed on not only relax the assumptions of conventional DEA on the outputs and inputs are known exactly, but also assume these inputs and outputs

* Corresponding author.
E-mail address: d96725002@ntu.edu.tw (B. Hsiao).
[1] Tel.: +886 2 33661190; fax: +886 2 33661199.

[2] From Wikipedia, fuzzy logic is a form of multi-valued logic derived from fuzzy set theory to deal with reasoning that is robust and approximate rather than brittle and exact. In contrast with "crisp logic", where binary sets have two-valued logic, fuzzy logic variables may have a truth value that ranges in degree between 0 and 1.

0957-4174/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2011.01.075

could be represented as interval data. Furthermore, in contrast to the approach of Sengupta (1992), Kao and Liu (2000a), believed that only certain variables (versus *all* variables) with either linguistic issues or missing data (Kao & Liu, 2000b) could be applied using fuzzy concepts. Kao and Liu (2000a, 2004) based their analysis on the α-cut and extension principle, which was earlier proposed by Zadeh (1965) as a way to measure performances of linguist variables. Later, Saati, Memariani, and Jahanshahloo (2002) and Lertworasirikul, Fang, Nuttle, and Joines (2003) proposed a triangle fuzzy measure to replace the α-cut of the Kao and Liu (2000a, 2000b) model, and then formed it into a linear programming model. Moreover, Entani, Maeda, and Tanaka (2002) proposed another approach to replace the triangle fuzzy measure and use the three-point measure, which referred to "optimistic," "most possible," and "pessimistic".

In addition, the recently DEA literatures related to fuzzy measure (e.g., Cooper et al., 1999), are focused on radial-based measurements, however these measurements are hardly to differenced on the same efficiency score of DMUs ranking, because radial-based measurements have weaker discrimination power than slack-based measurement. Therefore, we based on the findings of Tone (2001), resulting in the use of combined fuzzy concepts with a non-radial lack-based measure of efficiency (SBM) model. We use this model to evaluate the performance of 24 commercial banks in Taiwan. Section 2 briefly explains the fuzzy slack-based measure DEA (hereinafter referred as Fuzzy SBM), followed by the introduction of fuzzy super-efficiency slack-based measure (hereinafter referred as Fuzzy Super SBM) in Section 3. In Section 4, the 24 commercial banks with fuzzy data are examined with Fuzzy Super SBM/Fuzzy SBM and compared with conventional Fuzzy DEA. In Section 5, a discussion is provided to illustrate ranking issues from Fuzzy Super SBM, Fuzzy SBM, and conventional Fuzzy DEA. Section 6 offers our final thoughts and conclusion.

2. Fuzzy slacks-based measure of efficiency (Fuzzy SBM)

The following are the concepts drawn by Kao and Liu (2000a) in Fuzzy DEA. First, we assume there are n DMUs, each DMU (DMU_j, where $j \in R_+^n$) with m inputs $(x_j \in R_+^m)$, and s outputs $(y_j \in R_+^s)$. Based on this description, slack-based measure DEA could be expressed through Model (1)

$$
\begin{aligned}
\min \quad & \rho_k = \frac{1-\frac{1}{m}\sum_{i=1}^{m}\frac{s_i^-}{x_{ik}}}{1+\frac{1}{s}\sum_{r=1}^{s}\frac{s_r^+}{y_{rk}}} \\
\textbf{subject to} \quad & x_{ik} = \sum_{j=1}^{n} x_{ij}\lambda_j + s_i^-, \quad i = 1,\ldots,m, \\
& y_{rk} = \sum_{j=1}^{n} y_{rj}\lambda_j - s_r^+, \quad r = 1,\ldots,s, \\
& \sum_{j=1}^{n}\lambda_j = 1, \\
& \lambda_j \geqslant 0, \quad j = 1,\ldots,n, \\
& s_i^- \geqslant 0, \quad s_r^+ \geqslant 0,
\end{aligned} \tag{1}
$$

where (x_{ik}, y_{rk}) represents the kth DMU, the ith input, and rth outputs. The λ_j is represented by the kth DMU weighting for evaluating efficiency. The s_i^- is represented by input excess, while s_r^+ is represented by output shortfall. In order to represent fuzzy input and output variables, we used $\mu_{\tilde{x}_{ij}}, \mu_{\tilde{y}_{rj}}$ as their membership functions, such that $\tilde{x}_{ij}, \tilde{y}_{rj}$ are the fuzzy number of input and output, respectively. Since Model (1) is based on fractional linear programming, it might produce an infinite number of solutions. Tone (2001) solved this problem by multiplying non-negative scalar variable q to the denominator and numerator, and by assuming that the objective function denominator is equal to 1. As such, Model (1) forms a simple linear programming similar to Model (2)

$$
\begin{aligned}
\min \quad & \tilde{\rho}_k = q - \frac{1}{m}\sum_{i=1}^{m}\frac{S_i^-}{\tilde{x}_{ik}} \\
\textbf{subject to} \quad & q + \frac{1}{s}\sum_{r=1}^{s}\frac{S_r^+}{\tilde{y}_{rk}} = 1, \\
& q\tilde{x}_{ik} = \sum_{j=1}^{n}\tilde{x}_{ij}\lambda'_j + S_i^-, \quad i = 1,\ldots,m, \\
& q\tilde{y}_{rk} = \sum_{j=1}^{n}\tilde{y}_{rj}\lambda'_j - S_r^+, \quad r = 1,\ldots,s, \\
& \sum_{j=1}^{n}\lambda'_j = q, \quad q > 0, \\
& \lambda'_j \geqslant 0, \quad j = 1,\ldots,n, \\
& S_i^- \geqslant 0, \quad S_i^- = qs_i^-, \\
& S_r^+ \geqslant 0, \quad S_r^+ = qs_r^+, \quad \lambda'_j = \frac{\lambda_j}{q}
\end{aligned} \tag{2}
$$

where $\tilde{\rho}_k = \rho_k$. In Model (2), all inputs and outputs are assumed to be linguist data. If some inputs or outputs are crisp data, then we can use degenerated member functions to describe them, such that this range will only have one value. We assume $\mathbf{S}(\tilde{x}_{ij})$, $\mathbf{S}(\tilde{y}_{rj})$ are $\tilde{x}_{ij}$, $\tilde{y}_{rj}$'s support sets that could achieve a definition with an elements set value larger than zero. The $\tilde{x}_{ij}$ and $\tilde{y}_{rj}$ by α-cut is therefore defined as (Eq. (3a) and (3b))

$$
(\mathbf{X}_{ij})_\alpha = \{x_{ij} \in \mathbf{S}(\tilde{x}_{ij}) | \mu_{\tilde{x}_{ij}}(x_{ij}) \geqslant \alpha\}, \quad \forall j, i, \tag{3a}
$$

$$
(\mathbf{Y}_{rj})_\alpha = \{y_{rj} \in \mathbf{S}(\tilde{y}_{rj}) | \mu_{\tilde{y}_{rj}}(y_{rj}) \geqslant \alpha\}, \quad \forall j, r, \tag{3b}
$$

where both $(\mathbf{X}_{ij})_\alpha$ and $(\mathbf{Y}_{rj})_\alpha$ are crisp set. Thus, inputs and outputs can be represented in intervals at different α-cut levels. Based on the above, we can express (Eq. (3a) and (3b)) as Model (4) and Model (5):

$$
\begin{aligned}
(\mathbf{X}_{ij})_\alpha &= \{x_{ij} \in S(\tilde{x}_{ij}) | \mu_{\tilde{x}_{ij}}(x_{ij}) \geqslant \alpha\} = [(\mathbf{X}_{ij})_\alpha^L, (\mathbf{X}_{ij})_\alpha^U] \\
&= \left[\min_{x_{ij}}\{x_{ij} \in S(\tilde{x}_{ij}) | \mu_{\tilde{x}_{ij}}(x_{ij}) \geqslant \alpha\}, \max_{x_{ij}}\{x_{ij} \in S(\tilde{x}_{ij}) | \mu_{\tilde{x}_{ij}}(x_{ij}) \geqslant \alpha\}\right],
\end{aligned} \tag{4}
$$

$$
\begin{aligned}
(\mathbf{Y}_{rj})_\alpha &= \{y_{rj} \in S(\tilde{y}_{rj}) | \mu_{\tilde{y}_{rj}}(y_{rj}) \geqslant \alpha\} = [(\mathbf{Y}_{rj})_\alpha^L, (\mathbf{Y}_{rj})_\alpha^U] \\
&= \left[\min_{y_{rj}}\{y_{rj} \in S(\tilde{y}_{rj}) | \mu_{\tilde{y}_{rj}}(y_{rj}) \geqslant \alpha\}, \max_{y_{rj}}\{y_{rj} \in S(\tilde{y}_{rj}) | \mu_{\tilde{y}_{rj}}(y_{rj}) \geqslant \alpha\}\right].
\end{aligned} \tag{5}
$$

If the α of Model (4) and Model (5) lies on a value between zero to one value (i.e., $\{(\mathbf{X}_{ij})_\alpha | 0 < \alpha \leqslant 1\}$, $\{(\mathbf{Y}_{rj})_\alpha | 0 < \alpha \leqslant 1\}$), the Fuzzy SBM model can easily be transformed into a crisp DEA model. Moreover, we can define the kth DMU membership functions using Zadeh's (1978) extension principle, as seen in Model (6)

$$
\mu_{\tilde{E}_k}(z) = \sup_{x,y} \min\{\mu_{\tilde{x}_{ij}}(x_{ij}), \mu_{\tilde{y}_{rj}}(y_{rj}), \ \forall j, r, i | z = E_k(x, y)\} \tag{6}
$$

where $E_k(x, y)$ is represented by the kth DMU efficiency score at a set of (x, y) by conventional SBM model and the $\mu_{\tilde{E}_k}$ is the minimum of $(\mu_{\tilde{x}_{ij}}(x_{ij}), \mu_{\tilde{y}_{rj}}(y_{rj}))$. According to Model (6), we need $\mu_{\tilde{x}_{ij}}(x_{ij}) \geqslant \alpha$, $\mu_{\tilde{y}_{rj}}(y_{rj}) \geqslant \alpha$ and at least one $\mu_{\tilde{x}_{ij}}(x_{ij})$ or $\mu_{\tilde{y}_{rj}}(y_{rj})$ equal to α, such that $z = E_k(x_{ij}, y_{rj})$ satisfies $\mu_{\tilde{E}_k}(z) = \alpha$. Therefore, given the $\alpha_2 < \alpha_1$ condition, the $\mathbf{X}_{ij}$ interval lies on (Eq. (7a)) while $\mathbf{Y}_{rj}$ lies on (Eq. (7b))

$$
\left[(\mathbf{X}_{ij})_{\alpha_1}^L, (\mathbf{X}_{ij})_{\alpha_1}^U\right] \subseteq \left[(\mathbf{X}_{ij})_{\alpha_2}^L, (\mathbf{X}_{ij})_{\alpha_2}^U\right], \tag{7a}
$$

$$
\left[(\mathbf{Y}_{rj})_{\alpha_1}^L, (\mathbf{Y}_{rj})_{\alpha_1}^U\right] \subseteq \left[(\mathbf{Y}_{rj})_{\alpha_2}^L, (\mathbf{Y}_{rj})_{\alpha_2}^U\right]. \tag{7b}
$$

Based on the above formulas, we can find that $\mu_{\tilde{E}_k}$ has a lower and upper bound by *a-cut*, as seen in Model (8) and Model (9)