CrossMark

ORIGINAL RESEARCH

A mixed integer bi-level DEA model for bank branch performance evaluation by Stackelberg approach

Morteza Shafiee[1] · Farhad Hosseinzadeh Lotfi[2] · Hilda Saleh[3] · Mehdi Ghaderi[1]

Abstract One of the most complicated decision making problems for managers is the evaluation of bank performance, which involves various criteria. There are many studies about bank efficiency evaluation by network DEA in the literature review. These studies do not focus on multi-level network. Wu (Eur J Oper Res 207:856–864, 2010) proposed a bi-level structure for cost efficiency at the first time. In this model, multi-level programming and cost efficiency were used. He used a nonlinear programming to solve the model. In this paper, we have focused on multi-level structure and proposed a bi-level DEA model. We then used a liner programming to solve our model. In other hand, we significantly improved the way to achieve the optimum solution in comparison with the work by Wu (2010) by converting the NP-hard nonlinear programing into a mixed integer linear programming. This study uses a bi-level programming data envelopment analysis model that embodies internal structure with Stackelberg-game relationships to evaluate the performance of banking chain. The perspective of decentralized decisions is taken in this paper to cope with complex interactions in banking chain. The results derived from bi-level programming DEA can provide valuable insights and detailed information for managers to help them evaluate the performance of the banking chain as a whole using Stackelberg-game relationships. Finally, this model was applied in the Iranian bank to evaluate cost efficiency.

✉ Morteza Shafiee
m_shafiee277@yahoo.com

[1] Department of Industrial Management, Shiraz Branch, Islamic Azad University, Shiraz, Iran

[2] Department of Mathematics, Science and Research Branch, Islamic Azad University, Tehran, Iran

[3] Department of Mathematics, Central Tehran Branch, Islamic Azad University, Tehran, Iran

Introduction

Banks, as one of the most complex industries in this rapidly changing high-tech world of computers and telecommunications, need to be flexible enough to respond rapidly to change and also to keep up with stiff competition. In such competitive environment, continuous improvement is critical for any successful organization. Therefore, improving performance is widely recognized as essential to gaining an extra competitive edge. Due to working in a competitive business environment, banking chain has a multi-dimensional and complex structure. Banking chain plays an important role in the economic cycle of each country, so to gain/remain on sustainable competitive edge performance evaluation has become a critical role which management of every financial institution may play. In this regard, bank branch performance, as the focus of empirical application in this paper, has become more difficult due to size variety, offering different services to different costumers, and multi-dimensional structure. Generally two methods are used to measure bank branch operational efficiency; parametric and non-parametric. The drawback of parametric techniques is a number of inherent limitations which make them unsuitable for fully reflecting the increasingly complex nature of banking chains. For example, regression analysis as one of the best parametric techniques is a central tendency method and is only suitable for modeling single input-multiple outputs or multiple inputs-single output systems. Data envelopment analysis (DEA), as a

🍀 Springer

non-parametric method, is an excellent efficiency analysis tool that creates efficient production frontier to compare the DMU under evaluation relative to the best decision making units (DMUs) if they are operating under the same conditions. Dealing with multiple inputs and multiple outputs is an appealing feature which gives DEA an edge over other analytical tools. Standard DEA evaluates the relative efficiency of DMUs in presence of multiple inputs and multiple outputs, but it does not provide sufficient details for managerial decisions. In real world scenarios, companies (DMUs) are mainly comprised of various divisions/levels that are linked together and have a great deal of interactions between them and follow multi-level Stackelberg relationships. Complex hierarchical structure of banking chains makes their managers deal with many challenges. The most important challenges that managers should cope with are: variable costs and demand under the multi-level Stackelberg relationships with interactions among levels, and coordinating these relationships to provide high quality services for customers. Bank managers can improve their cost efficiency through remaining or improving the quality of their services in order to create competitive advantage. Banking chain cost efficiency evaluates the ability of banks in producing current outputs at minimal level of cost which provides managers some insights behind the total cost in operations. The problems in the way of evaluation of banking branch performance naturally exhibit a multi-level decision making models which are connected in a hierarchical way. Since in the multi-level decentralized decision companies the individual set of decision variables is often controlled by each level which have their own, often mutually conflicting, objectives the evaluation of banking chain performance is based on multi-perspective. Therefore, performance evaluation in such particular multi-level decentralized decision structures can be modeled by bi-level programming DEA approach. Bi-level programming data envelopment analysis (BLPDEA) approach goes to the black box and embodies internal structure and interior interactions of system when it has a hierarchical structure. BLPDEA also can provide a valuable insight and detailed information to manager when evaluating the performance of a system with Stackelberg-game relationships (Wu 2010). The main purpose of this paper is to illustrate a seldom utilized non-parametric analysis technique called BLPDEA which addresses the problem of cost efficiency evaluation under the Stackelberg leader–follower relationships in the context of banking. Bi-level DEA model is a NP-hard nonlinear programming. Since nonlinear models are reducing the validity of the model, in this paper, we recommended the Mixed integer programming method which converts the bi-level mathematical programming to a linear mathematical programming and significantly improves the process to achieve the optimum solution.

The remainder of this paper is organized as follows: In the next section, "Literature review" section, literature is reviewed. Section "Fundamentals of DEA, DEA cost efficiency and bi-level programming" introduces the concepts of cost efficiency DEA and bi-level programming. In "Proposal model" section bi-level programming DEA model proposed by Wu (2010) is presented. In "The framework of the efficiency evaluation of the banks" section, we demonstrate empirical use of bi-level programming DEA model in the form of a case study (Iranian bank). Conclusions are presented in "Empirical study" section.

Literature review

In comparison with techniques of assessing organization performance, the method of DEA proposed by Charnes et al. (1978) is a better way to organize and analyze data because it allows efficiency to change over time and requires no prior assumption on the specification of the efficient frontier. Thus, DEA is an excellent approach for the performance analysis of banking industry. In many real world scenarios, DMUs have a two-stage network process and due to this reason, DEA has been extended to examine the efficiency of two-stage processes, where all of the outputs from the first stage are intermediate measures that make up the inputs of the second stage. Wang (1997) present a two-stage process in the banking industry where the banks use inputs (of the first stage) including fixed assets, labor, and information technology (IT) investments to generate deposits. The banks then use the deposits (intermediate measure) to generate loans and profits (as the outputs). Momen et al. (2012) measured the operational risk of Iranian banks based on Loss Distribution Approach. Bhattacharya et al. (1997) used a two-stage DEA approach to examine the impact of liberalization on the efficiency of the Indian banking industry. In the first stage, a technical efficiency score was calculated, whereas in the second stage a stochastic frontier analysis was used to attribute variation in efficiency scores of three sources: temporal, ownership and noise component. Seiford and Zhu (1999) examined the performance of the top 55 US banks using a two-stage DEA approach. Results indicated that relatively large banks exhibit better performance on profitability, whereas smaller banks tend to perform better with respect to marketability. Sexton and Lewis (2003) proposed a two-stage process for evaluating Major League Baseball performance. Khalili-Damghani and Taghavifard (2012) proposed a generic process in which just-in-time (JIT) practices are changed into agility indices, and agility indices are converted into performance measurement in supply chain in form of a conceptual model. Then, a three-stage data envelopment analysis (TSDEA) model proposed

to measure the relative efficiency of aforementioned process and sub-processes. Babazadeh et al. (2012) designed a network and applied a mixed integer linear programming to evaluate supply chain efficiency. Khalili-Damghani and Tavana (2013) proposed a new network DEA (NDEA) model for measuring the performance of agility in supply chains. Their proposed fuzzy NDEA model was linear and independent of the α-cut variables. Sanieemonfared and Safi (2013) applied a novel DEA network structure to measure relative efficiency of academic colleges in both teaching quality and research productivity. Kao and Hwang (2008) developed a different approach where the entire two-stage process can be decomposed into the product of the efficiencies of the two sub-processes. As a result, both the overall efficiency and each stage's efficiency are obtained. Tone and Tsutsui (2009) extended the SBM model into a network framework to deal with intermediate products formally. Avkiran (2009), as the first empirical study of NSBM, relies on actual aggregate data of domestic commercial banks in the UAE, and applied the non-oriented network slacks-based measure for evaluating the profit efficiency. Fukuyama and Weber (2010) extended the slacks-based inefficiency measure for evaluating a two-stage system with bad outputs in a Japanese bank. Paradi et al. (2011) developed a two-stage DEA approach for simultaneously benchmarking the performance of operating units. Li et al. (2012) extended the findings of Liang et al. (2008) and proposed a centralized and non-cooperative model to evaluate the efficiency of two-stage process to further decompose the overall efficiency for complex network structure. Khalili-Damghani and Hosseinzadeh Lotfi (2012) developed a fuzzy two-stage data envelopment analysis (FTSDEA). In this paper, each decision making unit supposed to make up of two serially connected sub-DMUs. Tavana and Khalili-Damghani (2014), proposed an efficient two-stage fuzzy DEA model to calculate the efficiency scores for a DMU and its sub-DMUs. They used the Stackelberg (leader–follower) game theory approach to prioritize and sequentially decompose the efficiency score of the DMU into a set of efficiency scores for its sub-DMUs.

Bi-level decision making or bi-level programming techniques, first introduced by Von Stackelberg (1952), have been developed for mainly solving decentralized decision process with decision makers in a hierarchical organization. Decision maker at the upper level is termed as the leader, and in the lower level, the followers which have their own and perhaps mutually conflicting objective. Bi-level programming has been applied in a great deal of fields. Ryu et al. (2004) presented a bi-level programming framework to capture conflicting interests of multiple elements in the context of supply chain planning problems. Huijun et al. (2008) proposed a bi-level programming model to describe and solve the location problem in which

both the benefits of costumers and logistics planers are taken into account. Sakawa et al. (2002) dealt with a transportation problem in a housing material manufacturer and derived a satisfactory solution to the problem. Roghanian et al. (2007) considered a probabilistic bi-level linear multi-objective programming problem and its application in enterprise-wide supply chain planning problem. Arora and Gupta (2009) presented an interactive fuzzy goal programming approach for bi-level programming problem with the characteristics of dynamic programming. Hongjie et al. (2011) established two inventory control models of deteriorating items respectively according to time-based and the quantity-based integrated delivery strategies for suppliers under VMI model based on bi-level programming. Lachhwani and Poonia (2012) developed a procedure for solving multilevel fractional programming problems in a large hierarchical decentralized organization by fuzzy goal programming approach. In this paper, fuzzy goal programming approach is used for achieving the highest degree of each of the membership goal by minimizing negative deviational variables. Alimardani et al. (2013) developed a continuous review policy for inventory control in a three-echelon supply chain including retailers, a central warehouse with limited storage space, and two independent manufacturing plants which offer two kinds of product to the customer. Arianezhad et al. (2013) presented a new two-echelon model to control the inventory of perishable goods. The main purpose of the model is to minimize the maintenance cost of the entire chain. Due to the complexity of the model, they used genetic algorithm under MATLAB to solve and confirm the accuracy of the model's performance.

Cost efficiency model is used to show the ability of DMUs to produce current outputs at minimal level of cost and how DEA can be used to identify types of inefficiency, which can emerge for treatment when information on costs is known exactly (Cooper et al. 2007). Tohidi and Khodadadi (2013) introduced a new model to evaluate cost efficiency of DMUs with negative data. They also demonstrated that proposed cost efficiency is a product of a locative and range directional measure efficiencies. Bahri and Tarokh (2012) focused on "seller–buyer" supply chain model with exponential distribution lead time and showed that their method can minimize the costs compared with systems that ignore the relation between seller and buyer.

Wu (2010) developed an innovative quantitative approach to evaluate the performance of multi-level decision network structure by integrating cost DEA into the bi-level programming framework and create bi-level programming DEA model. To show applicability of bi-level programming DEA model, Wu (2010) have demonstrated applications of the model in two practical examples: a banking chain and a manufacturing supply chain.

Fundamentals of DEA, DEA cost efficiency and bi-level programming

Data envelopment analysis

DEA is a linear programming based methodology which can calculate multiple inputs and outputs and can also evaluate DMUs both qualitatively and quantitatively. DMU, which can be related to different firms or the condition of the same firm over time, stands for decision making unit.

DEA was first proposed by Charnes et al. (1978). The evolutionary form of CCR model was suggested in 1984 by Banker et al. In subsequent years, several models were developed by a large number of researchers. Orientation, disposability, diversification, and return to scale are different aspects that can be seen in these models.

There are numerous studies on efficiency evaluation with the DEA model in several filed. For example, in banking such as Elyasiani and Mehdian (1990), Kao and Liu (2004), Camanho and Dyson (1999, 2005), and Cook and Hababou (2001). In supply chain efficiency evaluation, such as Khalili-Damghani and Sadi-Nezhad (2014), Khalili-Damghani and Hosseinzadel Lofit (2012), Khalili-Damghani and Taghavifard (2012, 2013), Abtahi and Khalili-Damghani (2011), Khalili-Damghani and Tavana (2013), and Khalili-Damghani et al. (2011, 2012). In other companies such as Tavana et al. (2014), Khalili-Damghani et al. (2015), Khalili-Damghani and Taghavifard (2012) and Tavana et al. (2013).

DEA cost efficiency

Many different types of DEA models with different aims have been developed. The aim of the majority of DEA models is focused on the technical-physical aspects of production for use in situations with unknown unit price and cost information, or where their uses are limited because of variability in the prices and costs that might need to be considered. Cost efficiency model is used to show the ability of DMUs to produce current outputs at minimal level of cost and how DEA can be used to identify types of inefficiency which can emerge for treatment when information on costs is known exactly (Cooper et al. 2007).

Suppose there are n DMUs under evaluation, each indexed by $(j = 1, \ldots, n)$, and $X = (x_1, \ldots, x_m)^T$ is the input vector which produces the output vector $Y = (y_1, \ldots, y_s)^T$ under the production possibility set. Then, the DEA cost efficiency model of 0th DMU, $(0 \in \{1, \ldots, n\})$, can be formulated as following linear programming:

$$cx^* = min \sum_{i=1}^{m} c_i x_i$$

$$s. to$$

$$\sum_{j=1}^{n} x_{ij} \lambda_j \leq x_i, \quad i = 1, \ldots, m;$$

$$\sum_{j=1}^{n} y_{rj} \lambda_j \geq y_{r0}, \quad r = 1, \ldots, s; \quad (1)$$

$$\lambda_j \geq 0, \ x_i \geq 0.$$

where (x_i, λ_j) are decision variables and (c_i) is the unit cost of input i which may vary from one DMU to another. This model allows substitutions in inputs. The objective function of model is to minimize the total cost of 0th DMU.

Based on an optimal solution (x^*, λ^*) of the above LP, the cost efficiency of DMU$_0$ is defined as:

$$E_c = \frac{CX^*}{CX_0} \quad (2)$$

where X_0 is the existing input vector of DMU$_0$.

Bi-level programming

Bi-level programming which is motivated by Von Stackelberg's game theory (1952) refers to situations where there are two decision makers in an organization which are inter-connected in a hierarchical structure. In such situations, the decision maker who first makes decision is termed as the leader and the other who knows the decision of opponent then makes a decision is termed as the follower. These two decision makers have independent, perhaps mutually conflicting, objectives. In the context of bi-level programming, the leader first specifies a decision and then the follower with the full knowledge of the leader's decision determines a decision so as to optimize his/her objective function. Accordingly, the leader also makes a decision so as to optimize his/her objective function. The obtained solution of the above mentioned procedure is a Stackelberg equilibrium solution (Sakawa and Nishizaki 2009). A bi-level linear programming problem for obtaining the Stackelberg solution is formulated as follows:

$$\underset{x}{minimize}\, z_1(x, y) = c_1 x + d_1 y$$

$$where\ y\ solves$$

$$\underset{y}{minimize}\, z_2(x, y) = c_2 x + d_2 y \quad (3)$$

$$subject\ to \quad Ax + By \leq b;$$

$$x \geq 0, y \geq 0.$$

where $c_i\ i = 1, 2$ are n_1-dimensional row coefficient vector, $d_i,\ i = 1, 2$ are n_2-dimensional row coefficient vector, A is an $m \times n_1$ coefficient matrix, B is an $m \times n_2$ coefficient

matrix, b is an m-dimensional column constant vector. $z_1(x,y)$ and $z_2(x,y)$ are the objective function of the leader and the follower, respectively. x and y are a set of decision variables which are controlled by the leader and follower, respectively (Sakawa and Nishizaki 2009).

Sakawa and Nishizaki (2009) gave the following definitions on basis of bi-level programming:

Definition 1 (Sakawa and Nishizaki 2009) S is the feasible region of the bi-level linear programming problem:

$$S = \{(x,y)|Ax + By \le b, x \ge 0, y \ge 0\}.$$

Definition 2 (Sakawa and Nishizaki 2009) $S(x)$ is the decision space of the follower after x is specified by the leader:

$$S(x) = \{y \ge 0|By \le b - Ax, x \ge 0\}.$$

Definition 3 (Sakawa and Nishizaki 2009) S_X is the decision space of the leader:

$$S_X = \{x \ge 0|\text{there is a } y \text{ such that } Ax + By \le b, y \ge 0\}.$$

Definition 4 (Sakawa and Nishizaki 2009) $R(x)$ is the set of rational responses of the follower for x specified by the leader:

$$R(x) = \left\{ y \ge 0 \middle| y \in \arg \min_{y \in S(x)} z_2(x,y) \right\}.$$

Definition 5 (Sakawa and Nishizaki 2009) Inducible region:

$$\text{IR} = \{(x,y)|(x,y) \in S, y \in R(x)\}.$$

Definition 6 (Sakawa and Nishizaki 2009) Stackelberg solution:

$$\left\{ (x,y)|(x,y) \in \arg \min_{(x,y) \in \text{IR}} z_1(x,y) \right\}.$$

In the bi-level programming, follower optimization problem is considered as one of the constraints of a bi-level optimization problem. So, by applying the Kuhn–Tucker approach as a popular way in solving BLP, the follower's problem can be replaced by the Kuhn–Tucker conditions of the follower's problem. Then the leader's problem with constraints involving the optimality conditions of the follower's problem is solved (Sakawa and Nishizaki 2009). Using Kuhn–Tucker conditions, the bi-level linear programming problem (3) can be rewritten as the following equivalent single-level nonlinear programming problem:

$$\begin{aligned}
&minimize \quad z_1(x,y) = c_1 x + d_1 y \\
&subject\ to \quad uB - v = -d_2; \\
&u(Ax + By - b) - vy = 0; \\
&Ax + By \le b; \\
&x \ge 0, y \ge 0, u^T \ge 0, v^T \ge 0,
\end{aligned} \tag{4}$$

where u is an m-dimensional row vector and v is an n_2-dimensional row vector. u and v are the dual variable associated with constraints $Ax + By \le b$ and $y \ge 0$, respectively.

By introducing zero-one vectors $w_1 = (w_{11}, \ldots, w_{1m})$ and $w_2 = (w_{21}, \ldots, w_{2n_2})$, the non-linear programming problem (4) is transformed into a following linear mixed zero-one programming problem, and it can be solved by a zero-one mixed integer solver:

$$\begin{aligned}
&minimize \quad z_1(x,y) = c_1 x + d_1 y \\
&subject\ to \quad 0 \le u^T \le Mw_1^T; \\
&0 \le b - Ax - By \le M(e - w_1^T); \\
&0 \le (uB + d_2)^T \le Mw_2^T; \\
&0 \le y \le M(e - w_2^T); \\
&x \ge 0,
\end{aligned} \tag{5}$$

where e is an m-dimensional vector of ones, and M is a large positive constant. The readers who are interested in bi-level programming may refer to Sakawa and Nishizaki (2009).

Proposal model

Banking branch performance evaluation problems naturally exhibit a two-level decision modeling which are connected in a hierarchical way. In a bank chain structure, the funds from costumers in the form of deposits are collected within the first level and in the second level the deposits from previous level are taken to make profit. Since the funds collected from the first level determines on the investment decision in the second level, the banking branch performance evaluation problem can be modeled as a leader-follower Stackelberg problem (Wu 2010). The conceptual bi-level DEA model with shared resource is depicted in Fig. 1.

Consider the n banking branch, which was indexed by $(j = 1, \ldots, n)$, is involving two levels L_1, and L_2. This L_1–L_2 chain has been addressed using bi-level programming structure, where the first level is termed as a leader and the second level as a follower. These two bank branch chains performance evaluation problem for specific DMU_0 can be
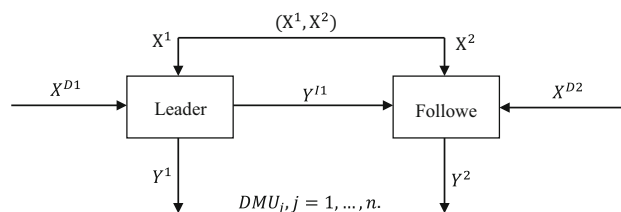
Fig. 1 Bi-level programming DEA model with shared resource (Wu 2010)

mathematically modeled in the bi-level programming DEA model, which considers the hierarchical structure of the bank branch including decision maker in each level who makes a decision to control a set of decision variables independently, as follows:

$$\underset{\overline{X}^1, \overline{X}^{D1}, \lambda}{Min} \left(C^{1^T}\overline{X}^1 + C^{2^T}\overline{X}^{D1} \right) + \left(C^{1^T}\overline{X}^2 + D^{1^T}\overline{X}^{D2} + D^{2^T}\overline{Y}^{I1} \right)$$

$s.\,to$

$$\sum_{j=1}^{n} X_j^1 \lambda_j \leq \overline{X}^1;$$

$$\sum_{j=1}^{n} X_j^{D1} \lambda_j \leq \overline{X}^{D1};$$

$$\sum_{j=1}^{n} Y_j^1 \lambda_j \geq Y_0^1;$$

$$\sum_{j=1}^{n} Y_j^{I1} \lambda_j \geq Y_0^{I1};$$

$$\overline{X}^1 + \overline{X}^2 \leq E(const.); \tag{6}$$

$$\underset{\overline{X}^2, \overline{X}^{D2}, \overline{Y}^{I1}, \pi}{Min} \left(C^{1^T}\overline{X}^2 + D^{1^T}\overline{X}^{D2} + D^{2^T}\overline{Y}^{I1} \right)$$

$s.\,to$

$$\sum_{j=1}^{n} X_j^2 \pi_j \leq \overline{X}^2;$$

$$\sum_{j=1}^{n} X_j^{D2} \pi_j \leq \overline{X}^{D2};$$

$$\sum_{j=1}^{n} Y_j^{I1} \pi_j \leq \overline{Y}^{I1};$$

$$\sum_{j=1}^{n} Y_j^2 \pi_j \geq Y_0^2;$$

$$\overline{X}^1, \overline{X}^2, \overline{X}^{D1}, \overline{X}^{D2}, \overline{Y}^{I1},\ \lambda, \pi \geq 0,$$

where the X^1 and X^2 are m_1-dimensional row vectors of the shared input of the leader and the follower, respectively, X^{D1} is an m_2-dimensional row vector of the direct input of the leader, X^{D2} is the m_3-dimensional row vector of the direct input to the follower, Y^{I1} is an m_4-dimensional row vector which is the intermediate output to the leader and the intermediate input to the follower, Y^1 is an m_5-dimensional row vector of the direct output of the leader, and Y^2 is an m_6-dimensional row vector of the direct output to the follower. $C^{1T}, C^{2T}, D^{1T}, D^{2T}$ are the input unit cost vectors associated with the shared input, the direct input to the leader, the direct input to the follower, and the intermediate input to the follower, respectively. λ and π are the nonnegative multiplier used to aggregate existing leader and follower activities, respectively (Wu 2010).

To solve the bi-level model problem Wu (2010) converts bi-level mathematical programming to a non-linear mathematical programming and then uses trial and error process to achieve the optimum solution, which makes the process lengthy and time-consuming. Since nonlinear models are reducing the validity of the model, in this paper, we recommended the Mixed integer programming method, which converts the bi-level mathematical programming to a linear mathematical programming and significantly improves the process to achieve the optimum solution.

The bi-level programming DEA model (6) is transformed into the mixed integer single-level linear programming DEA model as follows:

$$min\left(C^{1^T}\overline{X}^{1^T} + C^{2^T}\overline{X}^{D1^T} \right) + \left(C^{1^T}\overline{X}^2 + D^{1^T}\overline{X}^{D2} + D^{2^T}\overline{Y}^{I1^T} \right)$$

$s.\,to$

$$\sum_{j=1}^{n} X_j^1 \lambda_j \leq \overline{X}^1;$$

$$\sum_{j=1}^{n} X_j^{D1} \lambda_j \leq \overline{X}^{D1};$$

$$\sum_{j=1}^{n} Y_j^1 \lambda_j \geq Y_0^1;$$

$$0 \leq \overline{X}^2 - \left(\sum_{j=1}^{n} X_j^2 \pi_j \right) \leq M w_1^T;$$

$$0 \leq U^1 \leq M(e - w_1^T);$$

$$0 \leq \overline{X}^{D2} - \left(\sum_{j=1}^{n} X_j^{D2} \pi_j \right) \leq M w_2^T;$$

$$0 \leq U^2 \leq M(e - w_2^T);$$

$$0 \leq \overline{Y}^{I1} - \left(\sum_{j=1}^{n} Y_j^{I1} \pi_j \right) \leq M w_3^T;$$

$$0 \leq U^3 \leq M(e - w_3^T);$$

$$0 \leq \sum_{j=1}^{n} Y_j^2 \pi_j - Y_0^2 \leq M w_4^T; \tag{7}$$

$$0 \leq U^4 \leq M(e - w_4^T);$$

$$0 \leq E - \overline{X}^1 - \overline{X}^2 \leq M w_5^T;$$

$$0 \leq U^5 \leq M(e - w_5^T);$$

$$U^1 - U^5 + V^1 = C^1;$$

$$U^2 + V^2 = D^1;$$

$$U^3 + V^3 = D^2;$$

$$- X_j^2 U^1 - X_j^{D2} U^2 - Y_j^{I1} U^3 + Y_j^2 U^4 + V^4 = 0;$$

$$0 \leq \overline{X}^2 \leq Mw_6^T;$$
$$0 \leq V^1 \leq M(e - w_6^T);$$
$$0 \leq \overline{X}^{D2} \leq Mw_7^T;$$
$$0 \leq V^2 \leq M(e - w_7^T);$$
$$0 \leq \overline{Y}^{I1} \leq Mw_8^T;$$
$$0 \leq V^3 \leq M(e - w_8^T);$$
$$0 \leq \pi_j \leq Mw_9^T;$$
$$0 \leq V^4 \leq M(e - w_9^T);$$
$$\overline{X}^1, \overline{X}^{D1}, \lambda \geq 0,$$

where the U^1 and V^1 are the m_1-dimensional dual vectors correspond to shared input constraints and variables to the follower, respectively, U^2 and V^2 are the m_3-dimensional dual vectors associated with direct input constraints and variables to the follower, respectively, U^3 and V^3 are the m_4-dimensional dual vectors correspond to the intermediate input constraints and variables of the follower, respectively, U^4 and V^4 are the n-dimensional dual vectors, and U^5 is an m_1-dimensional dual vectors correspond to the constrained resources. Correspondingly, W_i^T, $i = 1, \ldots, 9$, are the zero-one vectors. e and M are the vector of ones and the large positive constant, respectively.

By solving the bi-level programming DEA model, the optimal solution of $\left(\overline{X}^{1*}, \overline{X}^{2*}, \overline{X}^{D1*}, \overline{X}^{D2*}, \overline{Y}^{I1*}, \lambda_j^*, \pi_j^* \right)$ are obtained. Based on optimal solution, the cost efficiency of the leader of DMU$_0$ is defined as:

$$CE_0^L = \frac{C^{1^T}\overline{X}^{1*} + C^{2^T}\overline{X}^{D1*}}{C^{1^T}X_0^1 + C^{2^T}X_0^{D1}} \tag{8}$$

and the cost efficiency of the follower of DMU$_0$ is defined as:

$$CE_0^F = \frac{C^{1^T}\overline{X}^{2*} + D^{1^T}\overline{X}^{D2*} + D^{2^T}\overline{Y}^{I1*}}{C^{1^T}X_0^2 + D^{1^T}X_0^{D2} + D^{2^T}Y_0^{I1}} \tag{9}$$

and the total cost efficiency of DMU$_0$ is defined as:

$$CE_0^S = \frac{\left(C^{1^T}\overline{X}^{1*} + C^{2^T}\overline{X}^{D1*} \right) + \left(C^{1^T}\overline{X}^{2*} + D^{1^T}\overline{X}^{D2*} + D^{2^T}\overline{Y}^{I1*} \right)}{\left(C^{1^T}X_0^1 + C^{2^T}X_0^{D1} \right) + \left(C^{1^T}X_0^2 + D^{1^T}X_0^{D2} + D^{2^T}Y_0^{I1} \right)}. \tag{10}$$

The framework of the efficiency evaluation of the banks

After analyzing the previous literature review, the efficiency of the banks which was proposed throughout this research was shown in Fig. 2. The analytical processes
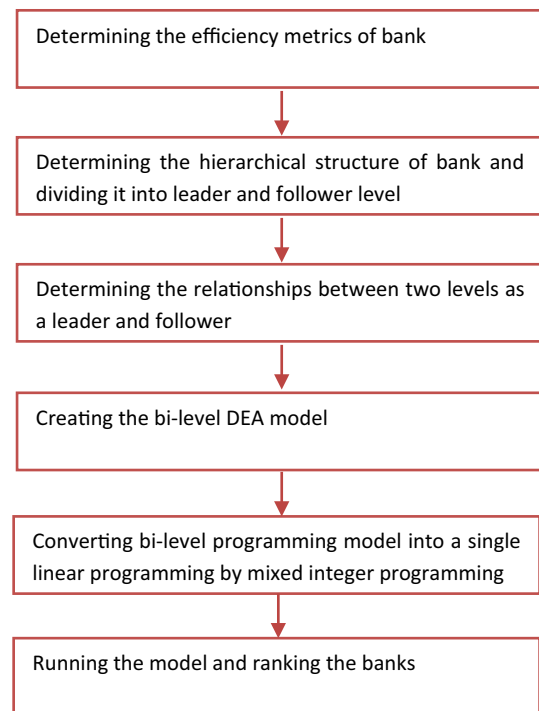

Fig. 2 Proposed performance evaluation model of bank

were divided and carried out in six steps: (1) in the first step the efficiency measurements of the bank were determined by reviewing literature and expert ideals; (2) in the second step the hierarchical structure of the bank was determined and the measurements were divided into two levels namely leader and follower levels; (3) in the third step the relationships between the leader and the follower were determined; (4) in the fourth step the main bi-level DEA model of the bank was created; (5) in the fifth step the created bi-level model was converted into a single linear model using mixed integer programming; (6) and finally in the sixth step the efficiency of the bank was obtained and the its ranking was recognized.

Empirical study

Banks are financial institutions that gather their assets from different resources and they made available for the sections that need liquidity. Therefore, banks are considered critical currents for each nation. Along with the emergence of private banks in the financial markets, Demand has dramatically increased for variety of banking services. Banks are looking for different procedures of functional improvement to attract customers, since they overtake one another to increase their contribution in market and Profitability; performance evaluation of banks is significantly considered among them and it is become one of the main bank manager activities.

Growth and economic development of every country requires moving additional funds of saver to the investors in a proper way. Availability of an extensive and efficient financial market which in financial resources is directed toward the best investment opportunities is critical. On the other hand, the maximum turnover in Iran is achieved through banking system. In addition, desired functioning of the banking system plays an important role to improve the economic actions.

However, it can be used the most traditional factors of an enterprise performance to evaluate a bank performance with modern procedures. One of the institutions investigating function of banking system is Iran Banking education Institute. Most performance evaluation criteria of it is quantitative, in addition it is considered financial standards. It investigates banks based on Liabilities, Assets, Number of bank branches, international and exchange Activities, The combination of human resources, profits and losses, facilities, overdue demands and the benefit of electronic banking technology. In this study, first it is determined the indicators through bank experts interview and library studies and criteria are extracted. These include: fixed assets, space, non-invest deposits, IT cost, profit, Deposit, Marketability employees and profitability employees. Usually the bank performance evaluation process involves the measurement of the performance of bank through its profitability levels. There are some other factors, in addition to profitable activities, that play an important role in bank performance evaluation process, including: bank location (ratio of residential to commercial region of the banks), cultural context of the region (whatever people tend to deposit their money as long term or current deposit), the services bank provides to compete with other banks and attract customers. These factors affect bank's performance indirectly, and usually are ignored in performance evaluation. So, in addition to the bank's profitability, intangible factors that indirectly affect bank's performance should also be taken into account while measuring the performance of banks. Thus, to measure all aspects of bank performance, the banks were evaluated in two levels so that all activities that affect the performance of banks would be taken into consideration.

Each bank studied in this research considers certain strategies according to the needs of people and market traction due to the geographic location of each region in terms of area, culture, population density, the ratio of commercial to residential area, etc. The banks offer different services to their clients by hiring the required number of employers, and with the deployment of fixed assets such as office supplies, computers, desks, chairs, banking software, IT costs, storage systems, and data protection. Banks collect their resources by attracting deposits from clients, and they begin profiting by investing those resources on various projects and giving loans.

As mentioned, banks can be viewed as an entity in which two decision makers in a hierarchical structure make decisions in turn so as to optimize their performance. Based on the proposal model in the previous section, the bi-level programming DEA model was applied to evaluate the performance of 15 branches of Iranian banks in 2011. Each branch had a certain amount of marketability and profitability level, where the marketability played as a leader and the profitability level served as a follower. Marketability level indicates the ability of the branch in marketing to collect funds from costumer in form of deposit by consuming the bank's resources. Profitability level indicates the ability of the branch to make profit by investing the deposits in other activities. The evaluation index system of bank branch performance evaluation problem is shown in Table 1.

The performance of each branch is characterized by seven variables where fixed assets (X^{D11}), space (X^{D12}), and non-invest deposits (Y^1) are associated with leader level, while IT cost (X^{D2}) and profit (Y^2) are associated with the follower level. Deposit (Y^{I1}) from leader to follower level is intermediate variable. Marketability employees (X^1) and profitability employees (X^2) are resource shared variables. The data related to these variables are presented in Table 2. Shared employees and the space costs are also given in Table 2 and due to cost nature of fixed assets, deposits, IT costs, and the correspond costs are assumed to be unit.

The first column of Table 3 depicts the results of performance evaluation using the CCR model, where each unit is as a black box with only a few input and output while the relationships between the elements are not considered. To calculate the performance of the banks in this way, four

Table 1 Evaluation index system

Factors	Name of index	Unit of index
Leader		
Shared input	Employees	Person
Direct inputs	Fixed assets	1,000,000,000 Riyal
	Space	m^2
Intermediate output	Deposit	10,000,000,000 Riyal
Output	Non-invest deposit	100,000,000 Riyal
Input costs	Employees	1,000,000 Riyal
	Space	1,000,000 Riyal
Follower		
Shared input	Employees	Person
Direct input	IT cost	100,000,000 Riyal
Intermediate input	Deposit	10,000,000,000 Riyal
Output	Profit	10,000,000,000 Riyal
Input cost	Employees	1,000,000 Riyal

Table 2 Data

	Leader employee	Fixed asset	Space	Non-invest deposit	Deposit	Follower employee	IT cost	Profit	Employees cost	Space cost
DMU 1	23	4.93	110	10.57	15.78	14	4.93	3.15	5.97	5.97
DMU 2	34	3.64	167.50	14.82	43.90	28	3.64	3.31	6.25	4.41
DMU 3	14	2.87	150	7.60	7.74	25	2.87	3.53	7.86	6.02
DMU 4	28	1.78	366.6	14.80	45.95	31	1.78	1.98	5.70	10.64
DMU 5	33	3.61	555	13.39	38.70	35	3.61	2.69	9.65	11.78
DMU 6	34	2.24	690	14.15	36.81	28	2.24	1.93	7.59	8.25
DMU 7	25	1.41	750.92	13.08	61.11	28	1.41	2.12	8.53	16.02
DMU 8	35	2.64	368.3	20.31	26.33	18	2.64	2.55	12.37	9.92
DMU 9	27	2.97	205	16.16	11.74	40	2.97	1.62	6.87	12.96
DMU 10	35	2.65	195	16.21	70.42	11	2.65	2.00	7.90	7.21
DMU 11	33	3.76	210	38.24	34.25	15	3.76	2.90	11.24	12.97
DMU 12	28	2.47	481	17.94	37.03	26	2.47	2.25	5.61	7.33
DMU 13	20	2.29	200	13.73	11.66	24	2.29	2.45	8.51	12.05
DMU 14	28	3.55	506	14.17	10.31	36	3.55	4.19	100.03	11.75
DMU 15	35	1.97	814	14.41	31.98	30	1.97	1.96	7.53	10.45

Table 3 Cost efficiency and reference set unit

	Classic DEA efficiency	Bank cost efficiency	Leader cost efficiency	Follower cost efficiency	Reference set for the leader	Reference set for the follower
DMU 1	1	0.66	0.61	1.00	10, 11	1
DMU 2	0.82	0.50	0.90	0.70	10, 11	3
DMU 3	1	0.48	0.37	1.00	10, 11	3
DMU 4	0.84	0.14	0.41	0.28	10, 11	1
DMU 5	0.58	0.08	0.24	0.35	10, 11	1
DMU 6	0.63	0.10	0.20	0.31	10, 11	1
DMU 7	1	0.05	0.24	0.31	10	1
DMU 8	0.75	0.14	0.41	0.92	10, 11	3
DMU 9	0.44	0.19	0.54	0.29	11	3
DMU 10	0.61	0.32	1.00	0.52	10	1
DMU 11	0.67	0.18	1.00	0.79	11	1
DMU 12	0.69	0.15	0.33	0.52	10, 11	3
DMU 13	0.86	0.21	0.44	0.71	11	3
DMU 14	0.92	0.09	0.20	0.83	11	3
DMU 15	0.69	0.07	0.15	0.30	10, 11	1

inputs (Employee, Fixed asset, Space, and IT Cost) are used to obtain one output (Profit). (these inputs and outputs belong to leader and follower levels.)

At this stage we ignored the intermediate variables that exist between the leader and follower. The results showed that three out of the 15 banks were efficient, and the other 12 banks had high performance compared to the performance of the bi-level models. This experiment shows how weak the DEA's classic model is in regards to separation ability. Compared to classic DEA models, bi-level DEA model has higher separation ability, mainly because it provides a tool to

reveal the internal activities and relationships between activities within the black box and provides a detailed assessment of the existing subsystem performance.

Based on mixed integer single-level linear programming DEA model (7), cost efficiency scores for the bank branches and the followers and the leaders are obtained. Cost efficiency scores and the reference units correspond to the leader and the follower are given in following Table 3.

According to Table 3, there are no cost efficient banks, because banks are not performing efficiently at both decision levels. The tenth and eleventh banks are cost efficient at the

leader, and the first and third banks are cost efficient at the follower level, and because the other player in these banks is inefficient, these banks systematically are termed as inefficient. Table 3 also indicates that the second bank is more efficient than the third, tenth, and eleventh banks which are only efficient in one level, so that, it can be the advantages of coordination between players. In addition, the classic DEA efficiency score and reference set for the leader and follower are shown in the Table 3. DMU 1, 3 and 7 are efficient in classic DEA model, DMU 10 and 11 are reference for the leader and DMU 1 and 3 are reference for the follower.

Conclusion

In the real world, banks have a decentralized structure in which multi decision makers in a hierarchical structure makes decision in turn or at the same time to optimize their objective function. In this rapidly changing world, responding to change needs the ability of management to identify the location of inefficiency. Thus, efficiency analysis can be a source of competitive advantages (Avkiran 2009). DEA is a better way to measure efficiency since it requires no prior assumption on the specification of the efficient frontier. In this paper, we applied bi-level programming DEA model with two inter-related decision makers in a decentralized decision structure to evaluate the performance of 15 Iranian bank branches with one level correspond to a leader while the other a follower. Bi-level programming DEA model proposed by (Wu 2010) can provide insight and detailed information to bank managers when measuring the efficiency of a bank with Stackelberg-game relationships. The results obtained from bi-level programming DEA model have a strong discriminating power due to considering internal operations in the banking chain.

Further researchers can develop our model in the three-level or multi-level. Since the multi-level DEA models are NP-hard problem, we proposed to use heuristic model to solve them. Finally, we propose to further researcher to create a benchmark unit in bi-level DEA model and determine VRS in efficiency model.

References

Abtahi AR, Khalili-Damghani K (2011) Fuzzy data envelopment analysis for measuring agility performance of supply chains. Int J Model Oper Manag 1(3):263–288

Alimardani M, Jolai F, Rafiei H (2013) Bi-product inventory planning in a three-echelon supply chain with backordering, Poisson demand, and Limited warehouse apace. J Ind Eng Int 9:22

AriaNezhad MG, Makuie A, Khayatmoghadam S (2013) Developing and solving two-echelon inventory system for perishable items in a supply chain: case study (Mashhad Behrouz Company). J Ind Eng Int 9:39

Arora SR, Gupta R (2009) Interactive fuzzy goal programming approach for bi-level programming problem. Eur J Oper Res 194:368–376

Avkiran NK (2009) Opening the black-box of efficiency analysis: an illustration with UAE banks. Omega 37:930–941

Babazadeh R, Razmi J, Ghodsi R (2012) Supply chain network design problem for a new market opportunity in an agile manufacturing system. J Ind Eng Int 8:19

Bahri M, Tarokh MJ (2012) A seller-buyer supply chain model with exponential distribution lead time. J Ind Eng Int 8:13

Bhattacharya A, Lovell CAK, Sahay P (1997) The impact of liberalization on the productive efficiency of indian commercial banks. Eur J Oper Res 98(2):332–345

Camanho AS, Dyson RG (1999) Efficiency, size, benchmarks and targets for bank branches: an application of data envelopment analysis. J Oper Res Soc 50:903–915

Camanho AS, Dyson RG (2005) Cost efficiency measurement with price uncertainty: a DEA application to bank branch assessments. Eur J Oper Res 161(2):432–446

Charnes A, Cooper WW, Rhodes E (1978) Measuring the efficiency of decision making units. Eur J Oper Res 2:429–444

Cook WD, Hababou M (2001) Sales performance measurement in bank branches. Omega. Int J Manag Sci 29(4):299–307

Cooper WW, Seiford LM, Tone K (2007) Data envelopment analysis: a comprehensive text with models, applications, references and DEA-solver software. Springer, New York

Elyasiani E, Mehdian SM (1990) A nonparametric approach to measurement of efficiency and technological change: the case of large US banks. J Financ Serv Res 4(2):157–168

Fukuyama H, Weber WL (2010) A slacks-based inefficiency measure for a two-stage system with bad outputs. Omega 38:398–409

Hongjie L, Ruxian L, Zhigao L, Ruijiang W (2011) Study on the inventory control of deteriorating items under VMI model based on bi-level programming. Expert Syst Appl 38:9287–9295

Huijun S, Ziyou G, Jianjum W (2008) A bi-level programming model and solution algorithm for the location of logistics distribution centre. Appl Math Model 32:610–616

Kao C, Hwang SN (2008) Efficiency decomposition in two-stage data envelopment analysis: an application to non-life insurance companies in Taiwan. Eur J Oper Res 185(1):418–429

Kao C, Liu ST (2004) Predicting bank performance with financial forecasts: a case of Taiwan commercial banks. J Bank Finance 28(10):2353–2368

Khalili-Damghani K, Hosseinzadeh Lotfi F (2012) Performance measurement of police traffic centers using fuzzy DEA-based Malmquist productivity index. Int J Multicrit Decis Mak 2(1):94–110

Khalili-Damghani K, Sadi-Nezhad S, Hosseinzadeh-Lotfi F (2014) Imprecise DEA models to assess the agility of supply chains. In: Kahraman C, Öztayşi B (eds) Supply chain management under fuzziness, Studies in fuzziness and soft computing. Springer, Berlin, Heidelberg, vol 313, pp 167–198

Khalili-Damghani K, Taghavifard M (2012) A three-stage fuzzy DEA approach to measure performance of a serial process including JIT practices, agility indices, and goals in supply chains. Int J Serv Oper 13(2):147–188

Khalili-Damghani K, Taghavifard B (2013) Sensitivity and stability in two-stage models with fuzzy data. Int J Oper Res 17(1):1–37

Khalili-Damghani K, Tavana M (2013) A new fuzzy network data envelopment analysis model for measuring the performance of agility in supply chains. Int J Adv Manuf Technol 69(1–4):291–318

Khalili-Damghani K, Taghavifard M, Olfat L, Feizi K (2011) A hybrid approach based on fuzzy DEA and simulation to measure the efficiency of agility in supply chain: real case of dairy industry. Int J Manag Sci Eng Manag 6(3):163–172

Khalili-Damghani K, Taghavi-Fard M, Abtahi AR (2012) A fuzzy two-stage DEA approach for performance measurement: real case of agility performance in dairy supply chains. J Appl Decis Sci 5(4):293–317

Khalili-Damghani K, Tavana M, Haji-Saami E (2015) A data envelopment analysis model with interval data and undesirable output for combined cycle power plant performance assessment. Expert Syst Appl 42(2):760–773

Lachhwani K, Poonia MP (2012) Mathematical solution of multilevel fractional programming problem with fuzzy goal programming approach. J Ind Eng Int 8:16

Li Y, Chen Y, Liang L, Xie J (2012) DEA models for extended two-stage network structure. Omega 40:611–618

Liang L, Cook WD, Zhu J (2008) DEA models for two-stage processes: game approach and efficiency decomposition. Nav Res Logist 55:643–653

Momen O, Kimiagari A, Noorbakhsh E (2012) Modeling the operational risk in Iranian commercial banks: case study of a private bank. J Ind Eng Int 8:15

Paradi CJ, Rouatt S, Zhu H (2011) two stage evaluation of bank branch efficiency using data envelopment analysis. Omega 39:99–109

Roghanian E, Sajjadi SJ, Aryanezhad MB (2007) A probabilistic bi-level linear multi-objective programming problem to supply chain planning. Appl Math Comput 187:786–800

Ryu J-H, Dua V, Pistikopoulos EN (2004) A bi-level programming framework for enterprise-wide process networks under uncertainty. Comput Chem Eng 28:1121–1129

Sakawa M, Nishizaki I (2009) Cooperative and non-cooperative multi-level programming, concepts, system, algorithms & case studies. Springer, New York

Sakawa M, Nishizaki I, Uemura Y (2002) A decentralized two-level transportation problem in a housing material manufacturer: interactive fuzzy programming approach. Eur J Oper Res 141:167–185

SanieeMonfared MA, Safi M (2013) Network DEA: an application to analysis of academic performance. J Ind Eng Int 9:15

Seiford LM, Zhu J (1999) Profitability and marketability of the top 55 US commercial banks. Manage Sci 45(9):1270–1288

Sexton TR, Lewis HF (2003) Two-stage DEA: an application to major league baseball. J Prod Anal 19:227–249

Stackelberg HV (1952) The theory of the market economy. Oxford University Press, Oxford

Tavana M, Khalili-Damghani K (2014) A new two-stage Stackelberg fuzzy data envelopment analysis model. Measurement 53:277–296

Tavana M, Khalili-Damghani K, Sadi-Nezhad S (2013) A fuzzy group data envelopment analysis model for high-technology project selection: a case study at NASA. Comput Ind Eng 66(1):10–23

Tavana M, Khalili-Damghani K, Rahmatian R (2014) A hybrid fuzzy MCDM method for measuring the performance of publicly held pharmaceutical companies. Ann Oper Res 226(1):589–621

Tohidi GH, Khodadadi M (2013) Allocating models for DMUs with negative data. J Ind Eng Int 9:16

Tone K, Tsutsui M (2009) Network DEA: a slacks-based measure approach. Eur J Oper Res 197(1):243–252

Wang H (1997) Contingent valuation of environmental resources: a stochastic perspective, Ph.D. Dissertation. Department of Environmental Sciences and Engineering, School of Public Health, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

Wu D (2010) Bi-level programming data envelopment analysis with constrained resource. Eur J Oper Res 207:856–864