Available online at http://sanad.iau.ir/journal/ijim/
Int. J. Industrial Mathematics (ISSN 2008-5621)
Vol.16, No. 2, 2024 Article ID IJIM-1629, 20 pages
Research Article

Science and Research Branch (IAU)

Cost Malmquist Productivity Index for Bi-level Units: A Case Study from Bank

M. Habibpoor, M.R. Alirezaee *, J. Rashidinia

Department of Mathematics, Iran University of Science and Technology, Tehran, Iran.

Submission Date: 2022/09/30, Revised Date: 2023/01/17, Date of Acceptance: 2024/04/28

Abstract

Traditional cost models ignore the internal structure of decision-making units (DMUs), so, may produce ambiguous outcomes and provide a biased assessment. In this paper, we evaluate the performance of the units by considering their internal structures. We proposed a new cost Malmquist index for measuring the cost productivity change of the units with bi-level structures. The bi-level structure is a special case of hierarchical structures with two levels, where the leader unit is positioned at the upper level and followers are located at the lower level. The overall system of bi-level units tries to use inputs and produce outputs in a cost-efficient way. However, each subunit performs according to its goals and limited resources. This research tries to develop a bi-level cost model that is suitable for measuring the cost efficiency of bi-level units. Based on this model, a new cost Malmquist index (CMI) is suggested to evaluate the productivity changes of bi-level units. This index presents a new aspect of CMI and provides the productivity changes of units by considering the impact of the leader's and the subunits' performance. In addition, similar to the traditional CMI, it decomposes into various components, such as cost efficiency changes and cost technological changes. The developed CMI is applied to a real-world case study to evaluate eight management regions which all together manage 198 branches. The results show that the proposed CMI provides a more meaningful evaluation of DMUs compared to the conventional CMI.

Keywords: Data envelopment analysis, Bi-level structure, Cost efficiency, Productivity, Cost Malmquist index.

* Corresponding author: Email: mralirez@iust.ac.ir

1. Introduction

Data envelopment analysis (DEA) is a powerful methodology for evaluating the relative efficiency of decision-making units (DMUs) with multiple inputs and outputs. This technique is formulated as a linear programming model by Charnes et al. [1]. The first model of DEA is known as Charnes, Cooper, and Rhodes (CCR) model under Constant Returns to Scale (CRS) technology. Later, Banker et al. [2] developed the CCR model for Variable Returns to Scale (VRS) and introduced Banker, Charnes, and Cooper (BCC) model. Currently, there are many DEA models that are developed for assessing the performance of units in many different activities.

Malmquist index (MI) is the most important index for measuring the productivity changes of DMUs in multiple time periods. The MI was first used in productivity literature by Caves et al. [3]. Fare et al. [4] developed a DEA-based decomposition known as FGLR decomposition consisting of two components, Technological Change (TC) and Efficiency Change (EC). FGLR decomposition uses the CRS model to calculate all distance functions. Three-component decomposition of MI was developed by Fare et al. [5] regarding both CRS and VRS technologies involving Pure Efficiency Change (PEC), Scale Efficiency Change (SEC), and TC; This decomposition is called FGNZ. MI components in FGNZ decomposition are calculated using CCR and BCC basic models.

Maniadakis et al. [6] proposed cost MI when producers are assumed as cost minimizers and input price data are available. The index is defined in terms of input cost rather than input quantity distance functions. In fact, the concept of Cost Efficiency (CE) is used in the definition of distance functions. CE is used to show the ability of DMUs to produce current outputs at a minimal level of cost [7]. The concept of CE was first developed by Fare et al. [8] using a linear programming model. They considered the CE of a DMU as the ratio of minimum production cost to actual observed cost. Tone [9] improved the CE model introduced by Fare et al. [8] by evaluating DMUs in a cost-based production possibility set instead of a conventional production possibility set. Tohidi and Khodadadi [10] introduced a new model to evaluate CE with negative data. They also demonstrated that CE is the product of locative and range directional measure efficiencies. Bahri and Tarokh [11] focused on the ‘‘seller-buyer’’ supply chain model with exponential distribution lead time and showed that their method can minimize costs compared with systems that ignore the relationship between seller and buyer. Jahanshahloo et al. [12] started CE evaluation considering fuzzy DEA. They supposed that input prices are certain at each DMU and input-output data are as fuzzy numbers. Dotoli et al. [13] present a new cross-efficiency fuzzy DEA method for evaluating DMUs under uncertainty. Camanho and Dyson [14] applied the standard weight restriction techniques in the form of input cone assurance regions to explore the assessment of CE in complex scenarios of incomplete price information. Mostafaee and Saljooghi [15] developed a pair of two-level mathematical programming problems for the estimation of upper and lower bounds based on the dual multiplier formulation of the CE model.

All these approaches view each DMU as a “black box” by considering the initial inputs depleted and final outputs. Although these approaches may be suitable for evaluating the units with single-level structure, they are not suitable for assessing the units with a hierarchical structure. Most organizations have a hierarchical structure. Hierarchical organizations have different levels and each level consists of a number of sub-units. A unit at the higher level is divided into several subordinate units at the lower level. Some larger units in the second level sometimes may be further divided into several subordinate units with different functions at the third level, and so on (see also [16]). Black-box models ignore the independent performance of each subunit and consider the total output of the subunits of a unit. Also, these models provide no insight regarding the locations of inefficiency and cannot provide specific guidance to DMU managers to improve DMU's performance. Therefore, a developed cost DEA model is needed to consider the internal structure of DMUs.

In the literature, little research has been done on the evaluation of units by considering their internal structures. Wu [17] 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 creating bi-level programming DEA model. Shafiee [18] proposed a bi-level DEA model and then used a mixed integer linear programming to solve the proposed model. Hakim [19] developed a bi-level model based on DEA for centralized resource allocation. Ray [20] analyzed the cost efficiency of Indian bank branches using a network model. Kao [21] developed a relational model to measure the MI and its decompositions for parallel production systems and evaluated the productivity changes of thirty-nine branches of an Iranian commercial bank. Wanke et al. [22] studied efficiency in MENA banks, applying a dynamic network DEA model for accounting and financial indicators. Wanke et al. [23] perform a super-efficiency analysis of OECD banks during 2004-2013 within the ambit of parametric and nonparametric dynamic network models. Yang et al. [24] established a bi-level DEA model with multiple followers for evaluating the efficiency of unattended convenience stores. They used weak disability technology to deal with undesirable intermediate measures. Pachar et al. [25] proposed a bi-level programming DEA approach to evaluate multiple retail stores' cost efficiency considering a network structure operating in a Stackelberg relationship and defining benchmarks for inefficient stores in India. Hua et al. [26] developed a bi-level DEA cost model to evaluate the efficiency of two subsystems and complete water systems in 10 cities in the Minjiang River Basin. Barat et al. [27] suggested a network DEA-based methodology to address the problem of non-homogeneity in settings where subunits operate in a mixed-network structure. They evaluated layers and the overall system efficiencies. Hosseini Monfared et al. [28] proposed a network DEA model based on a multiplier model for two-stage structures. They developed an additive efficiency decomposition approach wherein the overall efficiency is expressed as a weighted sum of the efficiencies of the individual stages. Mollaeian et al. [29] analyzed the overall efficiency of three-stage DMU by proposing a network DEA model. They used nine Iranian tomatoes paste supply chains by considering the sustainable development factors to demonstrate the application of their proposed model.

Our purpose in this paper is to evaluate the productivity changes of units by considering their internal structure and based on CE. Conventionally, the Cost Malmquist Index (CMI) measures productivity change of the whole system in which only the inputs consumed and the outputs produced by the system are considered.

In this paper, we define a new CMI for measuring the cost productivity changes of a unit by considering its internal structure in two time periods. This index presents a new aspect of CMI and provides the cost productivity changes of units by considering the impact of leader's and sub-DMUs' performance. In addition, similar to the traditional CMI, it decomposes into various components, such as CE changes and cost technological changes.

Hence, the remainder of this paper is organized as follows. Section 2 presents the concepts of bi-level structure. Traditional and bi-level CEs are described in sections 3 and 4, respectively. In section 5, the traditional and bi-level CMI and its components are presented. In section 6, a real case study on the branches of a specialized bank of Iran is done to show the ability and advantage of the suggested CMI. We will analyze and compare traditional CMI and bi-level CMI in this section. Finally, concluding remarks appear in section 7.

2. Bi-level structures

Most organizations in the real world have a hierarchical structure, with units at different levels. Hierarchical systems have attracted relatively little attention. An organization usually has several units at the first level, where each unit has a number of sub-units at the second level. Large subunits may be divided into several subunits with different functions at the third level, and then this divisional leveling can be continued if necessary. For example, a bank in a country has n subdivisions in different provinces of the country, where each province includes a number of regions, and each region is divided into different branches. Bank managers distribute employees and budgets in different provinces, and the decision makers of provincial banks also distribute the employees and budgets received in such a way that their sub-units generate the most output. The management of each region is interested in improving the performance of its branches (Figure 1).

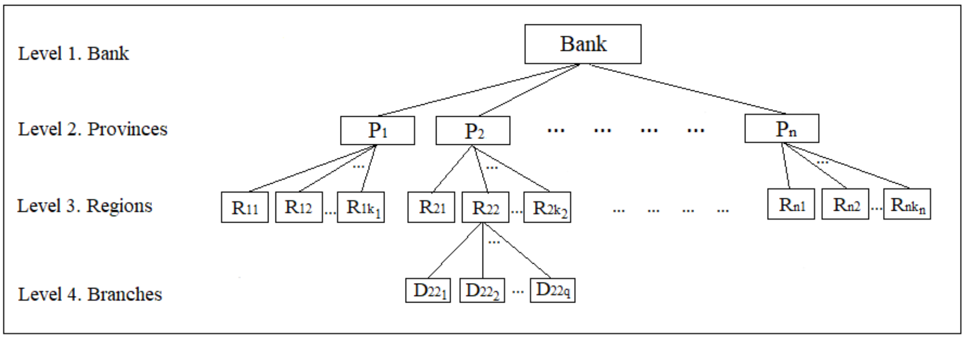

Figure 1. Hierarchical system of a bank.

Conventional DEA models are not suitable for evaluating the performance of such structures. In these models, the performance of subunits is ignored and the integration of inputs and outputs of sub-units is considered.

In this paper, we focus on a bi-level structure of hierarchical structures. The proposed model can be generalized to multi-level structures. Consider a system with a two-level structure, (see Figure 2). The system at the upper level has J units, numbered as $1, 2, ..., J$. Each $DMU_j, j = 1, 2, ..., J$ has R_j sub-DMUs at the lower level. Note that the units at the same level do not need to have the same number of subordinate units.

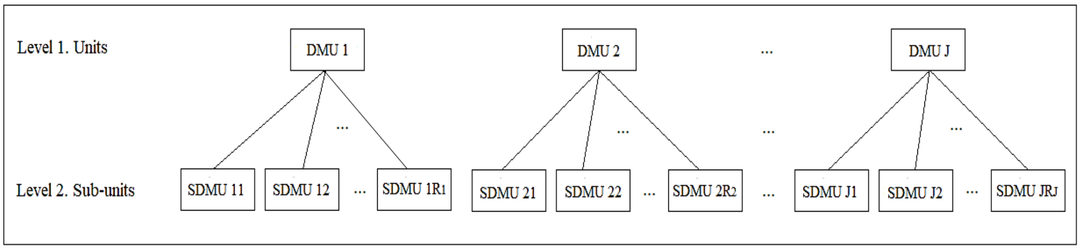

Figure 2. Bi-level system.

The way that a hierarchical unit operates is that the top management of the organization allocates the inputs to the divisions at the first level. The divisions at the first level consume some of the inputs that they received from the headquarters to generate their independent outputs and allocate the other amount to their subordinate divisions at the second level. Followers in the second level consume the inputs allocated to them by their leaders to produce the outputs. The leader may use some of the inputs to produce the intermediate outputs, which we do not consider here. Therefore, the outputs of a unit consist of its independent outputs and the total outputs of its subunits. Figure 3 shows the internal structure of a unit.

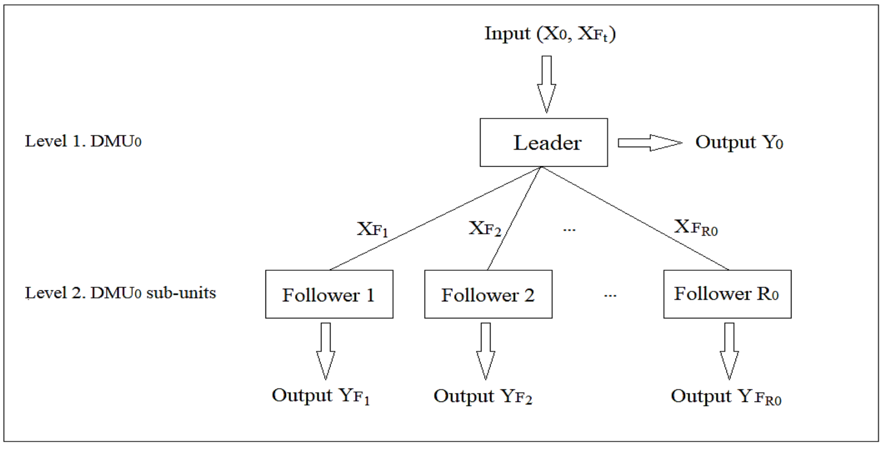

Figure 3. Internal structure of a unit.

3. Traditional cost efficiency

CE model is used to show the ability of DMUs to produce current outputs at a minimal level of cost. The traditional cost model treats the system as a "black box" with a number of initial inputs producing several final outputs, and the relationship and interactions between the leader and followers are not exposed. A conceptual "black box" system is depicted in Figure 4.

Consider J DMUs to be evaluated. Each DMU_j, $j=1,2,...,J$ uses inputs $X_j \in R^m$ to produce outputs $Y_j \in R^s$. To evaluate the cost efficiency of the DMU_0, $0 \in \{1,2,...,J\}$, we solve the linear programming problem below [7]:

$$
\begin{aligned}
C(Y_0,W_0) = \underset{X,\lambda}{Min}\; & W_0 X \\
s.t.\; & \sum_{j=1}^{J} \lambda_j Y_j \geq Y_0 \\
& \sum_{j=1}^{J} \lambda_j X_j \leq X \\
& \lambda \geq 0 \\
& X \geq 0
\end{aligned} \qquad (1)
$$

where (X,λ) are decision variables and W_0 is the cost vector of DMU_0 which may vary from one DMU to another. The objective function of model is to minimize the total cost of the DMU_0.

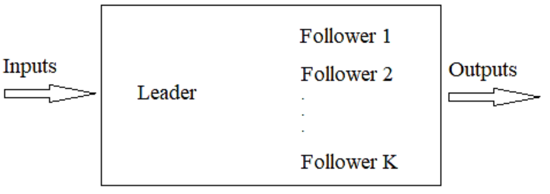

Figure 4. Structure of a “black box” system.

Based on an optimal solution (X^*,λ^*) of the above model, the traditional CE of DMU_0, is defined as:

$$
CE = \frac{C^*(Y_0,W_0)}{W_0 X_0} = \frac{W_0 X^*}{W_0 X_0} \qquad (1)
$$

where X_0 is the existing input vector of DMU_0.

4. Bi-level cost efficiency

Wu [17] proposed a cost DEA model for n bi-level decision systems in which each DMU includes two decentralized subsystems: a leader and a follower. Since most systems have more than one follower at the lower level, we extend Wu’s model to evaluate bi-level DMUs with one leader at the upper level and a number of followers at the lower level.

As mentioned, in bi-level structures, the performance of a unit depends on the optimal performance of its subunits and independent performance. Therefore, the CE of a unit with a bi-level structure should be measured based on the independent CE of the unit at the upper level and the CE of its subunits at the lower level. The proposed model for measuring the CE of DMU_0, can be expressed as:

$$
Min\; C_{BLP}(Y_0, Y_{F_k}^0, W_0, W_{F_k}^0) = W_0 \bar{X}_0 + \sum_{k=1}^{R_0} W_{F_k}^0 \bar{X}_{F_k}^0 \qquad (3)
$$

$$s.t.\begin{cases} \sum_{j=1}^{J}\lambda_j X_j \le \bar{X}_0 \\ \sum_{j=1}^{J}\lambda_j Y_j \ge Y_0 \\ \lambda_j \ge 0,\ \forall j=1,...,J \\ \bar{X}_0 \ge 0 \\ Min\ \ W_{F_1}^0\ \bar{X}_{F_1}^0 \\ s.t.\begin{cases} \sum_{j=1}^{J}\sum_{q=1}^{R_j}\mu_{F_{1q}}^j\ X_{F_q}^j \le \bar{X}_{F_1}^0 \\ \sum_{j=1}^{J}\sum_{q=1}^{R_j}\mu_{F_{1q}}^j\ Y_{F_q}^j \ge Y_{F_1}^0 \\ \mu_{F_{1q}}^j \ge 0,\ \ \forall q=1,...,R_j,\ \forall j=1,...,J \\ \bar{X}_{F_1}^0 \ge 0 \end{cases} \\ \vdots \\ Min\ \ W_{F_{R_0}}^0\ \bar{X}_{F_{R_0}}^0 \\ s.t.\begin{cases} \sum_{j=1}^{J}\sum_{q=1}^{R_j}\mu_{F_{R_0q}}^j X_{F_q}^j \le \bar{X}_{F_{R_0}}^0 \\ \sum_{j=1}^{J}\sum_{q=1}^{R_j}\mu_{F_{R_0q}}^j\ Y_{F_q}^j \ge Y_{F_{R_0}}^0 \\ \mu_{F_{R_0q}}^j \ge 0,\ \ \forall q=1,...,R_j,\ \forall j=1,...,J \\ \bar{X}_{F_{R_0}}^0 \ge 0 \end{cases} \end{cases}$$

where i, r are the indexes of the generic input and output, respectively. (X_j, Y_j) is the activity vector of the leader in DMU_j. $(X_{F_k}^j, Y_{F_k}^j)$ is the activity vector of follower k in DMU_j. W_0, $W_{F_k}^0$ are the unit cost associated with the leader and follower k.

At the upper level, the leader aims to minimize the total costs of the leader and the followers. Objective function of the upper level is specified by the sum of the products of each input in its unit cost as $W_0\bar{X}_0 + \sum_{k=1}^{R_0} W_{F_k}^0 \bar{X}_{F_k}^0$. The leader determines the inputs including $\bar{X}_0$ and an optimal multiplier λ. Then, at the lower level, each follower determines its own optimal input $\bar{X}_{F_k}^0$ in order to minimize the cost $W_{F_k}^0 \bar{X}_{F_k}^0$ subject to its constraint conditions. In the model, the CE of the leader is measured in comparison with other leaders, and the CE of each follower is calculated by comparing with all the followers in all of the units. In fact, in this model, the units at the same level are compared with each other and the effect of different levels on cost efficiency of whole unit is applied. This model can also be easily extended to multi-level structures.

After solving bi-level cost model (3), the optimal solution $\bar{X}_0^*, \bar{X}_{F_k}^{0\ *}, \lambda^*, \mu^*$ are obtained. Based on which, the CE of the follower k in DMU_0 is defined as:

$$CE_{F_k}^{0} = \frac{W_{F_k}^{0} \bar{X}_{F_k}^{0\,*}}{W_{F_k}^{0} X_{F_k}^{0}} \qquad (4)$$

DMU_0's follower k is termed cost-efficient if and only if $CE_{F_k}^{0} = 1$.

The cost efficiency of DMU_0 is defined as:

$$CE_0 = \frac{W_0 \bar{X}_0^{\,*} + \sum_{k=1}^{R_0} W_{F_k}^{0} \bar{X}_{F_k}^{0\,*}}{W_0 X_0 + \sum_{k=1}^{R_0} W_{F_k}^{0} X_{F_k}^{0}} \qquad (5)$$

DMU_0 is termed cost-efficient if and only if $CE_0 = 1$.

5.Cost Malmquist productivity index

5.1. Background

Consider panel data of $j = 1, 2, ..., J$ producers and $t = 1, 2$ time periods. In time period t, producers use inputs $X^t \in R^m$ to produce outputs $Y^t \in R^s$. Define now the production technology of period t in terms of the input requirement set, which is:

$$T_C^t = \left\{ (X^t, Y^t) : X^t \ can\ produce\ Y^t \right\} \qquad (6)$$

The subscript 'C' indicates that the production technology satisfies CRS. Assume that T_C^t is non-empty, closed, convex and bounded. Alternatively, define the technology of production in terms of the input distance functions as:

$$D^t(Y^t, X^t) = \sup\left\{ \varphi :\ (X^t/\varphi, Y^t) \in T_C^T,\ \varphi > 0 \right\} \qquad (7)$$

$D^t(Y^t, X^t)$ is the largest factor by which the input levels in X^t can be divided while (X^t, Y^t) remains in T_C^t. If $D^t(Y^t, X^t) \geq 1$, then it is sufficient for $(X^t, Y^t) \in T_C^t$ and $D^t(Y^t, X^t) = 1$, if and only if (X^t, Y^t) lie on the frontier. $D^t(Y^t, X^t)$ is reciprocal to input-oriented measure of technical efficiency which is:

$$TE^t(Y^t, X^t) = \min\left\{ \theta : (\theta X^t, Y^t) \in T_c^t, \theta > 0 \right\} \qquad (8)$$

Taking time period t as the reference period, the input-oriented Malmquist index (MI_t) can be stated as follow:

$$MI_t = \frac{D^t(X^{t+1}, Y^{t+1})}{D^t(X^t, Y^t)} \qquad (9)$$

MI_t compares (X^{t+1}, Y^{t+1}) to (X^t, Y^t) by measuring their respective distances from the production boundary of the reference period t. Similarly, with reference period $t+1$, the following index can be defined:

$$MI_{t+1} = \frac{D^{t+1}(X^{t+1}, Y^{t+1})}{D^{t+1}(X^t, Y^t)} \qquad (10)$$

MI_{t+1} measures the distance of (X^{t+1}, Y^{t+1}) and (X^t, Y^t) from the production boundary of the reference period $t+1$. To avoid an arbitrary choice of a reference period, the geometric mean of MI_t and MI_{t+1} is used as:

$$MI = \left[\frac{D^t(X^{t+1}, Y^{t+1})}{D^t(X^t, Y^t)} \times \frac{D^{t+1}(X^{t+1}, Y^{t+1})}{D^{t+1}(X^t, Y^t)} \right]^{1/2} \qquad (11)$$

Distance functions in Eq. (11) are defined in Eq. (7) regarding the CRS production boundaries. Clearly, given the definition of the distance function in (8), when $MI > 1$, we have a deterioration in productivity between t and $t+1$. Productivity remains unchanged if $MI = 1$ and improve if $MI < 1$.

5.2. Traditional cost Malmquist productivity index

Maniadakis and Thanassoulis [6] proposed a productivity index applicable when producers are cost minimizers and input prices are known. In particular, a cost Malmquist index, which is defined in terms of cost rather than input distance functions, is developed and computed using DEA models.

When input prices $W^t \in R^m$ are available, the technology can be defined in terms of a cost function:

$$C^t(Y^t, W^t) = \min\left\{ W^t X^t : (X^t, X^t) \in T_c^t, W^t > 0 \right\} \qquad (12)$$

$C^t(Y^t, W^t)$ is the minimum cost of producing a given output vector Y^t and the input prices W^t and the technology of the period t. The set of input vectors X^t which correspond to the scalar $C^t(Y^t, W^t)$ lie on an iso-cost line which defines a cost boundary.

$$Iso\,C^t(Y^t, W^t) = \left\{ X^t : W^t X^t = C^t(Y^t, W^t) \right\} \qquad (13)$$

This boundary contains the input vectors that can have the minimum cost with their price W^t.

The input-oriented measure of overall (or cost) efficiency for (X^t, Y^t) under input prices W^t, is defined as follows:

$$OE^t(Y^t, X^t, W^t) = \frac{C^t(Y^t, W^t)}{W^t X^t} \qquad (14)$$

By using cost efficiency, input's price productivity changes are determined. Specifically, in the spirit of the indexes in Eq. (9), Eq. (10), and Eq. (11) define a dual Cost Malmquist (CM) productivity index of periods t and $t+1$, and their geometric mean as follows:

$$CM_t = \frac{\dfrac{W^t X^{t+1}}{C^t(Y^{t+1}, W^t)}}{\dfrac{W^t X^t}{C^t(Y^t, W^t)}}, \qquad (15)$$

$$CM_{t+1} = \frac{\dfrac{W^{t+1} X^{t+1}}{C^{t+1}(Y^{t+1}, W^{t+1})}}{\dfrac{W^{t+1} X^t}{C^{t+1}(Y^t, W^{t+1})}}, \qquad (16)$$

$$CM = \left[\frac{\dfrac{W^t X^{t+1}}{C^t(Y^{t+1}, W^t)}}{\dfrac{W^t X^t}{C^t(Y^t, W^t)}} \times \frac{\dfrac{W^{t+1} X^{t+1}}{C^{t+1}(Y^{t+1}, W^{t+1})}}{\dfrac{W^{t+1} X^t}{C^{t+1}(Y^t, W^{t+1})}} \right]^{1/2} \qquad (17)$$

where $W^t X^t = W_0 X_0$ and $C^t(Y^t, W^t)$ as defined in Eq. (12).

The cost ratio $W^t X^t / C^t(Y^t, W^t)$ measures the extent to which the aggregate production cost in period t can be reduced while still securing the output vector Y^t under the input price vector W^t. This ratio measures the distance between the observed cost $W^t X^t$ and the cost boundary $C^t(Y^t, W^t)$. This distance will have a minimum value of 1, (when the two costs coincide), and the larger it is the larger the factor by which the observed cost of securing output Y^t can be reduced. This (cost) distance is the reciprocal of the input-oriented measure of overall efficiency defined in Eq. (14). The rest of the cost ratios in Eq. (15), Eq. (16), and Eq. (17) are defined analogously.

Thus, the CM index is defined in terms of distances by which input costs can be deflated to reach a cost boundary. This is in contrast to the MI which is defined in terms of distances by which input quantities can be deflated to reach a production boundary. Note at this point that again we do not need to make assumptions about the returns to scale that characterize the technology. Here, we use the CRS cost boundaries as benchmarks for productivity measurement. Just as with the MI index, the CM index value of less than 1 implies productivity progress, a value greater than 1 means regress, and a value of 1 indicates constant productivity.

The CM index can be decomposed into Cost (overall) Efficiency Changes (CEC) and Cost Technical Changes (CTC) as follows:

$$CM = CEC \times CTC \qquad (18)$$

where

$$CEC = \frac{\dfrac{W^{t+1} X^{t+1}}{C^{t+1}(Y^{t+1}, W^{t+1})}}{\dfrac{W^t X^t}{C^t(Y^t, W^t)}} \qquad (19)$$

and

$$CTC = \left[\frac{\dfrac{W^t X^{t+1}}{C^t(Y^{t+1}, w^t)}}{\dfrac{W^{t+1} X^{t+1}}{C^{t+1}(Y^{t+1}, w^{t+1})}} \times \frac{\dfrac{W^t X^t}{C^t(Y^t, w^t)}}{\dfrac{W^{t+1} X^t}{C^{t+1}(Y^t, w^{t+1})}} \right]^{1/2} \qquad (20)$$

CEC denotes the cost efficiency changes between periods t and $t+1$. The CEC indicates whether the DMU ‘catches up’ the cost boundary when moving from period t to $t+1$. CEC<1, CEC>1 and CEC=1 means that overall efficiency, increase, decrease or be constant between periods t and $t+1$, respectively. Similar concept is applied to CTC that measures the shift of the cost boundary evaluated at the input prices W^t and W^{t+1}.

5.3. Bi-level cost Malmquist productivity index

In this section, we propose a new cost Malmquist index that it is suitable for evaluating the cost productivity changes of DMUs with bi-level structures. Towards this end, based on definitions in previous section, the bi-level cost function $C_{BLP}^{t}(Y_{BLP}^{t}, W_{BLP}^{t})$ is defined as follows:

$$\begin{aligned} C_{BLP}^{t}(Y_{BLP}^{t}, W_{BLP}^{t}) &= C_{BLP}^{t}(Y_0^{\,t}, Y_{F_1}^{0t}, \ldots, Y_{F_{R_0}}^{0\ t}, W_0^{\,t}, W_{F_1}^{0t}, \ldots, W_{F_{R_0}}^{0\ t}) \\ &= \min \left\{ \begin{array}{l} W_0 \bar{X}_0 + \sum_{k=1}^{R_0} W_{F_k}^{0} \bar{X}_{F_k}^{0}; \\ (\bar{X}_0^{\ t}, \bar{X}_{F_1}^{0\,t}, \ldots, \bar{X}_{F_{R_0}}^{0\ t}, Y_0^{\,t}, Y_{F_1}^{0t}, \ldots, Y_{F_{R_0}}^{0\ t}) \in T_{BLP}^{t} \end{array} \right\} \end{aligned} \tag{21}$$

where T_{BLP}^{t} is the bi-level production possibility set and is defined in constraint set of model (3).

$C_{BLP}^{t}(Y_{BLP}^{t}, W_{BLP}^{t})$ is the minimum cost of producing a given output vector $Y_{BLP}^{t} = (Y_0^{\,t}, Y_{F_1}^{0t}, \ldots, Y_{F_{R_0}}^{0\ t})$ and the input prices $W_{BLP}^{t} = (W_0^{\,t}, W_{F_1}^{0t}, \ldots, W_{F_{R_0}}^{0\ t})$ and the technology of period t.

Based on the bi-level cost distance function and CE definition for the bi-level unit, the bi-level cost Malmquist (BCM) productivity index, also termed as the cost Malmquist productivity index of the DMU_0 with bi-level structure is defined as follows:

$$BCM = \left[\frac{\dfrac{W_{BLP}^{t} X_{BLP}^{t+1}}{C_{BLP}^{t}(Y_{BLP}^{t+1}, W_{BLP}^{t})}}{\dfrac{W_{BLP}^{t} X_{BLP}^{t}}{C_{BLP}^{t}(Y_{BLP}^{t}, W_{BLP}^{t})}} \times \frac{\dfrac{W_{BLP}^{t+1} X_{BLP}^{t+1}}{C_{BLP}^{t+1}(Y_{BLP}^{t+1}, W_{BLP}^{t+1})}}{\dfrac{W_{BLP}^{t+1} X_{BLP}^{t}}{C_{BLP}^{t+1}(Y_{BLP}^{t}, W_{BLP}^{t+1})}} \right]^{1/2} \tag{22}$$

where $W_{BLP}^{t} X_{BLP}^{\,t} = W_0 X_0 + \sum_{k=1}^{R_0} W_{F_k}^{0} X_{F_k}^{0}$.

Similar to the traditional CM index, the BCM index can be decomposed into bi-level cost efficiency changes (BCEC) and bi-level cost technological changes (BCTC) as follows:

$$BCM = BCEC \times BCTC \tag{23}$$

where

$$BCEC = \frac{\dfrac{W_{BLP}^{t+1} X_{BLP}^{t+1}}{C_{BLP}^{t+1}(Y_{BLP}^{t+1}, W_{BLP}^{t+1})}}{\dfrac{W_{BLP}^{t} X_{BLP}^{\,t}}{C_{BLP}^{t}(Y_{BLP}^{t}, W_{BLP}^{t})}} \tag{24}$$

and

$$BCTC = \left[\frac{\dfrac{W_{BLP}^{t} X_{BLP}^{t+1}}{C_{BLP}^{t}(Y_{BLP}^{t+1}, W_{BLP}^{t})}}{\dfrac{W_{BLP}^{t+1} X_{BLP}^{t+1}}{C_{BLP}^{t+1}(Y_{BLP}^{t+1}, W_{BLP}^{t+1})}} \times \frac{\dfrac{W_{BLP}^{t} X_{BLP}^{\,t}}{C_{BLP}^{t}(Y_{BLP}^{t}, W_{BLP}^{t})}}{\dfrac{W_{BLP}^{t+1} X_{BLP}^{\,t}}{C_{BLP}^{t+1}(Y_{BLP}^{t}, W_{BLP}^{t+1})}} \right]^{1/2} \tag{25}$$

BCEC denotes the cost efficiency changes of a DMU with bi-level structure between periods t and $t+1$. BCTC indicates the cost technology changes of a DMU with bi-level structure between periods t and $t+1$.

BCEC<1, BCEC>1 and BCEC=1 imply the cost efficiency, increase, decrease and constant between periods t and $t+1$, respectively. Similar concept is applied to BCTC that measures the shift of the cost boundary evaluated at the input prices W^{t}_{BLP} and W^{t+1}_{BLP}.

The cost function of DMU_0's follower $F_k, k=1,2,...,R_0$ is defined as follows:

$$C^{0\,t}_{F_k}(Y^{0\,t}_{F_k}, W^{0\,t}_{F_k}) = \min\left\{ W^{0\,t}_{F_k}\bar{X}^{0\,t}_{F_k} :\ (\bar{X}^{0\,t}_{F_k}, Y^{0\,t}_{F_k}) \in S^{0\,t}_{F_k},\ W^{0\,t}_{F_k} > 0 \right\} \tag{26}$$

where

$$S^{0\,t}_{F_k} = \left\{ \begin{array}{l} \sum_{j=1}^{J}\sum_{q=1}^{R_j} \mu^{j}_{F_{kq}} (X^{j}_{F_k})^t \le (\bar{X}^{0}_{F_k})^t, \\ \sum_{j=1}^{J}\sum_{q=1}^{R_j} \mu^{j}_{F_{kq}} (Y^{j}_{F_k})^t \ge (Y^{0}_{F_k})^t, \\ \mu^{j}_{F_{kq}} \ge 0,\ \forall q=1,...,R_j,\ \forall j=1,...,J, \\ (\bar{X}^{0}_{F_k})^t \ge 0 \end{array} \right\} \tag{27}$$

Similarly, the CMI of DMU_0's follower $F_k, k=1,2,...,R_0$ is defined as:

$$CM_{F_k} = \left[\frac{\dfrac{W^{0\,t}_{F_k}(X^{0}_{F_k})^{t+1}}{C^{0\,t}_{F_k}\left((Y^{0}_{F_k})^{t+1}, W^{0\,t}_{F_k}\right)}}{\dfrac{W^{0\,t}_{F_k}(X^{0}_{F_k})^{t}}{C^{0\,t}_{F_k}\left((Y^{0}_{F_k})^{t}, W^{0\,t}_{F_k}\right)}} \times \frac{\dfrac{W^{0\,t+1}_{F_k}(X^{0}_{F_k})^{t+1}}{C^{0\,t+1}_{F_k}\left((Y^{0}_{F_k})^{t+1}, W^{0\,t+1}_{F_k}\right)}}{\dfrac{W^{0\,t+1}_{F_k}(X^{0}_{F_k})^{t}}{C^{0\,t+1}_{F_k}\left((Y^{0}_{F_k})^{t}, W^{0\,t+1}_{F_k}\right)}} \right]^{1/2} \tag{28}$$

where $C^{0\,t}_{F_k}\left((Y^{0}_{F_k})^t, W^{0\,t}_{F_k}\right)$ is the minimum cost of producing a given output vector $(Y^{0}_{F_k})^t$ and the input vectors $(X^{0}_{F_k})^t$ and input prices $W^{0\,t}_{F_k}$ in the period t. Other cost distance functions have similar meaning.

6. A real case study

In this section, we calculate the BCM index on a real-world case study from Maskan Bank of Iran located in Tehran for two time periods 2017-2018, and then analyze the results. This study relates to eight regions located in the capital city of Tehran and involves 198 branches. The number of branches in each region is between twenty and thirty-two. In each region, there is a supervisory branch as the leader unit and a number of branches as the follower units. The number of followers is 198. It is noted that Maskan Bank is the largest Iranian governmental bank in the housing sector. The results of the case study can be useful for managers to understand the effect of the subunits̓ performance on the unit productivity changes and also to find out how they can manage any budget for improving the unit productivity growth.

Production analysis is one of the most significant dimensions of bank branch performance [30]. In this case study, we measure the performance of bank branches with respect to this perspective. From the production perspective, the regions and their branches are considered as producers of services for taking deposits, making loans and providing other diverse banking services using personnel expenses and location index as inputs.

The input of personnel expenses includes all the quantities and quality factors related to the staff of a branch. The input of location includes all the quantities and qualities entities related to the physical location of a branch. The planning and programming department of the bank has done a project for this index and they considered all the related factors in the developed location index. The most important factors considered in computing the location index are branch customers' specifications, the physical location of the branch, and branch staff characteristics. We used the data of the location index in our evaluation.

The output of deposit includes all kinds of methods of gathering money by a branch. The planning and programming department of the bank has done a project for this index and they considered a weighted sum of all kinds of accounts considering their values and number of transactions for the calculation of the deposit index and we used the data of the deposit index in our evaluation .

The output of loans includes all the money given as all kinds of loans and mortgages by a branch and similar to the deposit index some calculations have been done. Finally, the output of services is an index that includes all kinds of services presented by a branch to its customers.

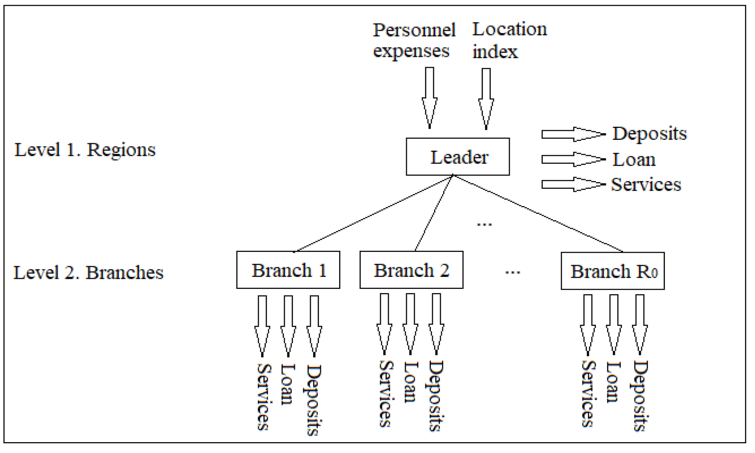

Figure 5. Bi-level production structure in our case study.

Descriptive statistics of inputs and outputs of leaders and followers for two time periods are given in Tables 1 and 2. Measurement unit of personnel expenses is 1000000 Rials. Other indices have no units because they are normalized. All the values and results are rounded in two digits.

Table 1. Data statistics of leaders.

	2017				2018			
Inputs	**Min**	**Max**	**Mean**	**STD**	**Min**	**Max**	**Mean**	**STD**
Personnel expenses	8727.07	14119.22	11029.86	1782.46	7071.13	16172.11	10178.99	3437.57
Location index	961.20	1192.00	1084.03	73.43	961.20	1192.00	1084.03	73.43
Outputs								
Deposits	2697.00	7540.00	4666.38	1579.2	3497.00	5491.00	4340.63	683.45
Loans	1018.00	3104.00	1610.88	752.19	932.10	2539.00	1495.81	663.46
Services	1111.00	7239.00	3075.25	2217.59	1433.00	5574.00	2740.63	1367.87

Table 2. Data statistics of followers.

	2017				2018			
Inputs	**Min**	**Max**	**Mean**	**STD**	**Min**	**Max**	**Mean**	**STD**
Personnel expenses	1115.09	18396.58	4663.19	2407.84	1384.62	16214.78	4507.42	2265.56
Location index	384.00	1212.00	945.79	153.59	384.00	1212.00	946.76	152.33
Outputs								
Deposits	154.50	3798.00	1218.25	674.36	86.80	3590.00	1216.70	669.42
Loans	35.07	18300.00	1221.25	1632.95	62.53	16127.00	1206.90	1581.19
Services	35.45	22045.00	1108.68	2042.54	110.60	8472.00	967.84	908.53

Descriptive statistics of input costs are given in Table 3. w_1 is the cost of personnel expenses which contain all expenses related to the staff of the branch such as pay, pension, etc. w_2 is the cost of the project for implementing and computing location index for each branch. Since the cost information is used in the form of proportion in the model, so their relative value is only important.

Table 3. Descriptive statistics of input costs.

	2017				2018			
leader	**Min**	**Max**	**Mean**	**STD**	**Min**	**Max**	**Mean**	**STD**
w_1	1.90	4.90	3.48	1.25	1.50	5.00	3.13	1.33
w_2	3.20	3.20	3.20	0.00	3.30	3.30	3.30	0.00
followers	**Min**	**Max**	**Mean**	**STD**	**Min**	**Max**	**Mean**	**STD**
w_1	0.70	5.00	3.29	1.20	1.00	5.00	3.32	1.19
w_2	3.20	3.30	3.26	0.05	3.20	3.30	3.24	0.05

Both CMI and BCMI with their components were calculated for all selected regions. The results of CMI and its decompositions are shown in Table 4 and the results BCMI and its decompositions are shown in Table 5.

Table 4. The results of traditional CMI and their components.

Regions	CEC	CTC	CM
1	0.98	0.98	0.96
2	1.00	1.16	1.16
3	1.00	1.03	1.04
4	0.81	1.03	0.83
5	1.00	1.01	1.01
6	1.23	0.92	1.13
7	0.96	1.04	0.99
8	1.26	1.00	1.26

The results of the CMI calculation show that regions 1 and 4 have been productive. In regions 5 and 7, there are almost no changes in their productivity growth. Other regions have been non-productive. Negative growth in CEC leads to a negative effect on the cost productivity growth in the non-productive regions 6 and 8. In non-productive regions 2 and 3, negative growth in CTC leads to a negative effect on the cost productivity growth. Traditional CMI does not indicate whether the negative changes in CTC or CEC in a region are due to negative performance at the upper level or lower level, and if it is at the lower level, which branches at the lower level had poor performance.

Table 5. The results of BCMI and their components.

Regions	BCEC	BCTC	BCM
1	1.02	0.95	0.97
2	0.97	1.04	1.01
3	0.99	0.91	0.90
4	0.90	0.87	0.79
5	1.02	0.92	0.94
6	1.42	0.86	1.23
7	0.96	0.89	0.85
8	1.41	0.88	1.24

The results of the BCMI calculation indicate that regions 1, 3, 4, 5 and 7 have been productive. In region 2, there are almost no changes in its productivity growth. Other regions have been non-productive. Region 7 has a positive growth in the cost efficiency change and frontier shifts which leads to a direct effect on the BCMI and has changed the rank of this region from the three using traditional CMI to the first.

Traditional CMI ignores the independent performance of subunits and the bi-level structure within units. However, by using BCMI, the cost productivity changes of the units are calculated more accurately, because the proposed bi-level cost model, which is used as the basic model in the calculation of BCMI, is embedded according to the bi-level structure of the units. We generalized a new model in which the impact of upper and lower levels' performance is included in the unit's performance evaluation. Therefore, BCMI increases the accuracy of data analysis.

Also, the traditional CMI calculation results may show that the regression of one of its components has caused the negative growth of cost productivity, but this is only true for single-level units. Traditional CMI does not show the sources of negative growth of components at different levels in bi-level units. Finding out which subunit or subunits in a region have performed poorly helps managers. BCMI helps managers to provide effective improvement solutions for hierarchically structured units.

Using BCMI, the sources of cost productivity regression of the bi-level units can be identified. For instance, we focus on the performance of region 6.

The results of the CMI and their components for the upper and lower levels units in region 6 are shown in Table 6. By examining the results, we can find out which levels of that region have negative growth in cost productivity and what factors affect the upper and lower levels' regression. CMI values for the lower-level units in Table 6 are calculated using relation (28).

Table 6. The results of the CM index and their components for the upper and lower levels units in region 6.

DMUs	CEC	CTC	CM
Upper level	1.78	0.92	1.64
Lower level			
1	0.73	0.86	0.63
2	1.71	0.79	1.35
3	2.13	0.89	1.90
4	1.55	0.84	1.30
5	2.08	0.87	1.81
6	1.71	0.88	1.51
7	1.06	0.83	0.88
8	2.27	0.89	2.01
9	1.54	0.91	1.40
10	1.42	1.21	1.72
11	1.54	0.90	1.39
12	1.57	0.81	1.27
13	0.75	0.91	0.68
14	1.74	0.88	1.53
15	1.75	0.90	1.57
16	0.96	0.85	0.82
17	1.76	0.86	1.51
18	1.73	0.78	1.35
19	1.14	0.85	0.98
20	1.50	1.05	1.58
21	0.99	0.85	0.84
22	0.90	0.87	0.78
23	2.01	1.01	2.03
24	0.96	0.84	0.81
25	1.08	0.86	0.92
26	1.45	0.81	1.18

Table 6 shows which factors at each level have influenced the results of BCMI of region 6. At the upper level, there is a positive change in the cost technology; but there is a significant negative change in the cost efficiency; this leads to a significant reduction in the cost productivity growth. At the lower level, some branches have negative growth in cost productivity. For each branch, the factors that caused its poor performance can be found. Although the upper level has negative growth in cost productivity, there are some branches at the lower level that have performed well. In addition to the impossibility of finding the exact sources of unproductive regions, conventional CMI ignores the positive performance of such branches.

7. Conclusion

Most organizations in the real world have a hierarchical structure, with units at different levels. The conventional cost model for measuring cost efficiency ignores the internal structure of a DMU. The knowledge of the internal structure of the DMUs might give further insights into their performance evaluation. In this paper, we focus on the bi-level structure because the results can be generalized to the multi-level structure. We developed a new cost model in order to provide more meaningful evaluations of DMUs with a bi-level structure. Based on the bi-level cost model, the BCM index was defined. Traditional CMI views each DMU as a "black box" by considering the initial inputs depleted and final outputs. It does not explain the underlying reasons why the DMUs with the hierarchical structure are non-productive or which subunit's performance requires further improvement. Compared to the traditional CMI, the BCMI can identify the subunits and factors that make a DMU non-productive. Also, by using BCMI, productivity changes are more accurate because the proposed two-level model is generalized according to the hierarchical structure of units. Therefore, BCMI provides potential managerial insights to improve the unit's performance. We utilized a real-world case study selected from bank branches to validate our BCMI and compare it with traditional CMI. Since the input costs in this research were known, the BCMI can be developed for the cases where the input costs are unknown or interval and it can be proposed as an approach for future research. Also, the proposed method can be extended to cases where intermediate outputs exist between the leader and the follower.

References

[1] Charnes, A., Cooper, W.W., Rhodes E., (1978), "Measuring the efficiency of the decision-making units", European Journal of Operational Research, 2, 429-444. https://doi.org/10.1016/0377-2217(78)90138-8

[2] Banker, R., Charnes, A., Cooper, W.W., (1984), "Some models for estimating technical and scale inefficiencies in data envelopment analysis", Management Science, 30 (1984) 1078-1092. https://doi.org/10.1287/mnsc.30.9.1078

[3] Caves, D.C., Christensen, L.R., Dievert, W.E., (1982), "The economic theory of index number and the measurement of input, output, and productivity", Econometrica, 50, 1393-1414. https://www.jstor.org/stable/1913388

[4] Fare, R., Grosskopf, S., Lindgren, B., Roose, P., (1992), "Productivity change in Swedish analysis pharmacies 1980-1989: A nonparametric Malmquist approach", Journal of Productivity, 3, 85-102. https://doi.org/10.1007/BF00158770

[5] Fare, R., Grosskopf, S., Lovell, C.K., (1985), "The Measurement of Efficiency of Production", Kluwer Academic Publishers, Dordrecht. https://doi.org/10.1007/978-94-015-7721-2

[6] Maniadakis, N., Thanassoulis, E., (2004), "A cost Malmquist productivity index", European Journal of Operational Research, 154, 396-409. https://doi.org/10.1016/S0377-2217(03)00177-2

[7] Cooper, W.W., Seiford, L.M., Tone, K., (2007), "Data envelopment analysis: a comprehensive text with models, applications, references and DEA-solver software", Springer, New York. https://www.scirp.org/reference/referencespapers?referenceid=1695927

[8] Fare, R., Grosskopf, S., Norris, M., Zhang, A., (1994), "Productivity growth, technical progress, and efficiency changes in industrial country", American Economic Review, 84, 66-83. https://www.jstor.org/stable/2117971

[9] Tone, K., (2002), "A strange case of the cost and allocative efficiencies in DEA", Journal of Operational Research Society", 53(11), 1225-1231. https://doi.org/10.1057/palgrave.jors.2601438

[10] Tohidi, G., M. Khodadadi, Allocation models for DMUs with negative data", Journal of Industrial Engineering International, 9(16) (2013) 1-6. https://link.springer.com/article/10.1186/2251-712X-9-16

[11] Bahri, M., Tarokh, M.J., (2012), "A seller-buyer supply chain model with exponential distribution lead time", Journal of Industrial Engineering International, 8. https://doi.org/10.1186/2251-712X-8-13

[12] Jahanshahloo, G.R., Hosseinzadeh Lotfi, F., Alimardani Jondabeh, M., Banihashemi, S., Lakzaie, L., (2008), "Cost efficiency measurement with certain price on fuzzy data and application in insurance organization", Applied Mathematical Sciences, 2(1), 1-18. https://www.m-hikari.com/ams/ams-password-2008/ams-password1-4-2008/lotfiAMS1-4-2008-1.pdf

[13] Dotoli, M., Epicoco, N., Falagario, M., Sciancalepore, F., (2015), "A cross-efficiency fuzzy data envelopment analysis technique for performance evaluation of decision making units under uncertainty", Computers and Industrial Engineering, 79, 103-114. https://doi.org/10.1016/j.cie.2014.10.026

[14] Camanho, A.S., Dyson, R.G., (2005), "Cost efficiency measurement with price uncertainty: a DEA application to bank branch assessments", European Journal of Operational Research, 161, 432-446. https://doi.org/10.1016/j.ejor.2003.07.018

[15] Mostafaee, A., Saljooghi, F.H., (2010), "Cost efficiency measures in data envelopment analysis with data uncertainty", European Journal of Operational Research, 202, 595-603. https://doi.org/10.1016/j.ejor.2009.06.007

[16] Kao, C., (2015), "Efficiency measurement for hierarchical network systems", Omega, 51, 121-127. https://doi.org/10.1016/j.omega.2014.09.008

[17] Wu, D., (2010), "Bi-level programming data envelopment analysis with constrained resource", European Journal of Operational Research, 207, 856-864. https://doi.org/10.1016/j.ejor.2010.05.008

[18] Shafiee, M., Hosseinzadeh Lotfi, F., Saleh, H., Ghaderi, M., (2016), "A mixed integer bi-level DEA model for bank branch performance evaluation by Stackelberg approach", Journal of Industrial Engineering International, 12, 81-91. https://doi.org//10.1007/s40092-015-0131-9

[19] Hakim, S., A. Seifi, A. Ghaemi, "A bi-level formulation for DEA-based centralized resource allocation under efficiency constraints", Computers and Industrial Engineering, 93(2016) 28-35. https://doi.org/10.1016/j.cie.2015.12.020

[20] Ray, S., (2016), "Cost efficiency in an Indian bank branch network: A centralized resource allocation model", Omega, 65 69-81. https://ideas.repec.org/a/eee/jomega/v65y2016icp69-81.html

[21] Kao, C., (2017), "Measurement and decomposition of the Malmquist productivity index for parallel production systems", Omega, 67, 54-59. https://doi.org/10.1016/j.omega.2016.04.001

[22] Wanke, P., Kalam Azad, M., Emrouznejad, A., Antunes, J., (2019), "A dynamic network DEA model for accounting and financial indicators: A case of efficiency in MENA banking", International Review of Economics Finance, 61, 52-68. https://doi.org/10.1016/j.iref.2019.01.004

[23] Wanke, P., Tsionas, M.G., Chen, Zh., Antunes, J. M., (2020), "Dynamic network DEA and SFA models for accounting and financial indicators with an analysis of super-efficiency in stochastic frontiers: An efficiency comparison in OECD banking", International Review of Economics Finance, 69, 456-468. https://doi.org/10.1016/j.iref.2020.06.002

[24] Yang, Z., Hui, L., Chen, X., (2021), "Efficiency evaluation of unattended convenience store based on bi-level DEA with multi-followers", E3S Web of Conferences, 253, 1-8. https://doi.org/10.1051/e3sconf/202125303005

[25] Pachar, N., Gupta, A., Jha, P.C., (2020), "Bilevel programming DEA approach for efficiency evaluation: a case study of Indian electronics retail stores", Yugoslav Journal of Operations Research, 30(4), 461-481. https://doi.org/10.2298/YJOR191124026P

[26] Hu, Z., Yan, S., Yao, L., Moudi, M., (2018), "Efficiency evaluation with feedback for regional water use and wastewater treatment", Journal of Hydrology, 562, 703-711. https://doi.org/10.1016/j.jhydrol.2018.05.032

[27] Barat, M., Tohidi, G., Sanei, M., (2019), "DEA for nonhomogeneous mixed networks", Asia Pacific Management Review, 24, 161-166. https://doi.org/10.1016/j.apmrv.2018.02.003

[28] Hosseini Monfared, S. N., Mozaffari, M.R., Rostamy Malkhalifeh, M., (2022), "Classifying flexible measures in two-stage network DEA", International Journal of Industrial Mathematics, 14(3), 306-318. http://dx.doi.org/10.30495/ijim.2022.19684

[29] Mollaeian E., Hosseinzade Lotfi, F., Toloie-Eshlaghy., A., Rostamy Malkhalifeh, M., Afshar kazemi M.A., (2022), "Congestion-based benchmarking and accessing sustainability in network DEA, Case study: Nine Iranian Tomato paste supply chains", International Journal of Industrial Mathematics, 14(1), 15-28. http://dx.doi.org/10.30495/ijim.2021.18999

[30] Paradi, J. C., Zhu, H., (2013), "A survey on bank branch efficiency and performance research with data envelopment analysis", Omega, 41, 61-79. https://doi.org/10.1016/j.omega.2011.08.010