Adv. Math. Fin. App., 2023, **8(**1), P. 157-172

Advances in Mathematical Finance & Applications

www.amfa.iau-arak.ac.ir
Print ISSN: 2538-5569
Online ISSN: 2645-4610
Doi: 10.22034/AMFA.2021.1923522.1566

Research Paper

Identifying the Effective Factors in Banking Sector using Data Envelopment Analysis Considering System Efficiency

Sharifeh Soofizadeh, Reza Fallahnejad*

Department of Mathematics, Khorramabad Branch, Islamic Azad University, Khorramabad, Iran.

ARTICLE INFO
Article history:
Received 2021-02-19
Accepted 2021-09-12

Keywords:
Banking
Financial assessment
Data Envelopment Analysis
Modified Russell measure
Variable Selection

ABSTRACT

Bank efficiency is essential in the establishment of healthy financial systems in countries. In this respect, bank managers are expected to respond correctly to questions raised about the financial performance of banks, which is practically impossible without examining the efficiency of the branches under their oversight. In most previous studies, Data Envelopment Analysis was used for evaluating the efficiency of financial branches. A large number of evaluation factors in the analysis leads to an increase in the number of efficient units and thus a decrease in the power of discrimination. Considering a systematic view, in this study, a step-by-step method was developed for selecting the effective factors in the efficiency of different branches of one of the Iranian Banks based on the effect of each factor or indicator on the whole system's efficiency including the branches under evaluation. To this end, a new method was proposed for the evaluation of the system's efficiency, and some of its properties were stated.

1 Introduction

Due to the obvious critical role of banks in most financial activities, their productivity (productivity and efficiency) is of paramount importance in both the private and public sectors. Owing to the service-based aspect of the operations of the banks and the broad variety of services rendered, the performance measurement has had particular challenges that need more precision and more effective approaches. Efficacious banking will assist society in achieving financial growth and progress. For this purpose, one of the important problems that bank management and financial officials pay significant attention to is the efficiency of the banking sector. Data Envelopment Analysis (DEA), which is used as a non-parametric tool to measure the efficiency of Decision-Making Units (DMU), is one of the most appropriate instruments in this respect. The use of DEA techniques is increasingly expanding today and is frequently used to analyze diverse organizations and sectors, such as the banking sector, post offices, hospitals, educational centers, powerhouses, refineries, etc. In addition to assessing relative efficiency, the use of DEA models defines the weaknesses of the enterprise in differing respects and, by presenting their

---

* Corresponding author. Tel.: +98-9163672137
E-mail address: R.Fallahnejad@khoiau.ac.ir

optimum level, establishes the strategy for the organization towards improving efficiency and productivity. It also provides benchmarks for evaluating inefficient units in order to increase their efficiency.

Substantial studies on the efficiency assessment of banks have been carried out. Some of them are addressed in this study. Using DEA, Fadzlan Sufian [1] obtained the productivity of Malaysia's banking sector in the Asian financial crisis with the use of variables such as bank size, profitability, and ownership. Tang and Manandhar [2] used DEA and established a framework for bank branch performance evaluation using three forms of efficiency, including operating efficiency, server efficiency, and profitability. As an indicator for DEA and bank efficiency assessment, Frei and Harker [3] used the Analytical Hierarchical Method (AHP) and explored the relationship between organizational and financial performance. Lotto [4] For the period between 2000 and 2017, Lotto analyzed factors impacting the operational performance of 36 commercial banks in Tanzania. To approximate the relationship between bank operational effectiveness and its determinants, he used a robust random effect regression model. Using a model focused on auxiliary variables, Wang et al. [5] assessed the two-stage network of Chinese commercial banks with adverse effects.

Ahadzade Namin et al. [6] used the weight restriction model in DEA to assess the performance of first-class branches of a commercial bank in Iran. To begin, the key indicators for evaluating the performance of bank branches were selected based on prior research. Then, based on the opinions of banking professionals and DEA, the efficiency of first-grade bank branches was evaluated, which included two input indicators and four output indicators. Henriques et al. [7] used DEA on a dataset of 37 Brazilian banks to evaluate bank efficiency between 2012 and 2016. By utilizing the intermediation technique to select variables, analyzing the major reasons for bank inefficiency, and determining how inefficient institutions in size may improve their efficiency, they examined three gaps in research performed with Brazilian banks. The efficiency of listed banks was studied statically and dynamically using DEA and the Malmquist index [8]. A model for locating the closest target in the presence of weight constraints has been described in [9]. The closest target for each DMU was introduced while taking into consideration trade-offs and weight constraints. Finally, the least changes to inefficient branches were represented while administering the approach for assessing one of Iran's banks. In [10], Ihaddaden used a radial DEA model to assess the efficiency of the Eurosystem's central banking system. Yu et al. [11] addressed the heterogeneity issue associated with determining meta-technology while assessing bank performance in a dynamic context. They examined and compared the dynamic performance of the financial holding company and non-financial holding company banks in Taiwan. Using a three stages methodology consisting of measuring the level of bank efficiency using DEA, evaluating the effect of financial performance on DEA efficiency using the Tobit regression model, and determining whether there is a difference in the efficiency of categories banks using the Mann-Whitney test, the level of efficiency from 2017 to 2018 of 18 Regional Development Banks and 35 Conventional Commercial Banks has been assessed. [12]. The super-efficiency DEA model was used to rank and compare the effectiveness of different central banks by Shair et al. [13]. The purpose was to look at the efficiency and total factor productivity (TFP) growth of the Pakistani banking industry, as well as the influence of risk and competition on efficiency and TFP growth. A sample of the 16 largest commercial banks in China was examined using a new technique for finding the nearest targets in the network structure [14]. Asmild et al. [15] applied the Multi-directional efficiency analysis approach to Bangladeshi banks in order to investigate the variations in inefficiency patterns amongst various subgroups. By conducting a dynamic three-step (production, intermediation, and profitability) network DEA, Azad et al. [16] analyzed the efficiency of conventional and Islamic banks in Malaysia. DEA, stochastic frontier, and

ANOVA studies on a sample of 90 individual banks from four distinct global regions (Europe, the United States, China, and India) were used to assess cost efficiency over a 15-year period (2002–2016) in [17]. In [18], frontier methods based on DEA and directional distance functions were used to evaluate the technical performance of a sample of 124 Brazilian banks and data for the six-year period 2014–19. For more studies about bank efficiency using DEA, see [19], [20], [21], [22], [23], [24], [25], [26],[27], [28], and [29].

In DEA, performance assessment is carried out using operational variables that are separated into indices of input and output. The selection of these factors and their proper classification is one of the important issues in this regard. It is necessary to select the correct number of such variables because the use of a large number of these variables in performance evaluation limits the distinguishing ability of the DEA. A variety of experiments and methods have been suggested for variables to be chosen. These rankings give the chance to pick the most powerful variables from various variables and, if necessary, reduce the number of them. Fanchon [30] proposes a technique to evaluate the optimum number of variables, determining their contribution to the construction of the efficiency frontier. Janet et al. [31] recommended the approach of step-by-step variable selection in the DEA. After eliminating or adding inputs or outputs, they used a formula to maximize (or minimize) the average of performance shifts. A new approach to choosing an appropriate set of variables using the genetic algorithm was explored in [32] and has been extended to the Indian banking sector. A step-by-step algorithm called Allocative DEA for selecting variables was proposed by Fernando et al. [33]. This algorithm was based on an average amount that evaluates the contribution of each variable in the calculation of the efficiency score of each DMU. Adler and Yazhemsky [34] used Monte Carlo simulation to generalize and compare the PCA-DEA and partial covariance-based variable reduction (VR). The evaluation of the methodologies found that PCA-DEA provides a more efficient instrument than VR.

The studies mentioned above mainly focused on the effect of each indicator on the individual efficiency of the units under evaluation and almost paid no attention to the impact of indicators on the efficiency of the whole system. In many cases, however, the system's overall performance, including the units under evaluation, is of significant importance. To this end, in this study, first, a new model for estimating the system efficiency is proposed according to a centralized scenario in DEA and the use of a modified Russell model and some of its properties are mentioned. Afterward, considering the step-by-step approach, the variables are chosen based on their impact on the overall system efficiency.

The structure of this paper is as follows. In section 2, some preliminaries are addressed. Our new approach is presented in section 3. In section 4, the applicability of the proposed model is presented in the case of different branches of Refah -Kargaran Bank in Lorestan province of Iran. Finally, in section 5, the study's conclusions are included.

2 Preliminaries

Assuming there are n $\mathrm{DMU_S}$ which consume m inputs to produce s outputs. The production possibility set under the variable return-to-scale is described as:

$$T_V = \left\{ (X,Y) \,\middle|\, X \geq \sum_{j=1}^{n} \lambda_j X_j \ , Y \leq \sum_{j=1}^{n} \lambda_j Y_j \ , \sum_{j=1}^{n} \lambda_j = 1 \ , \ \lambda_j \geq 0 \ ; j = 1,\dots,n \right\}$$

DEA models are divided into radial and non-radial. In the radial models, the efficiency is separately examined for input-based and output-based measures. On the contrary, in the non-radial models, the efficiency values for both input-based and output-based measures are simultaneously examined and do not have input and output orientation. The modified Russell model [35], which has very good properties, is one of the most important models in DEA studies. It is categorized within the non-radial DEA models class. The modified Russell measure for the evaluation of the under-evaluation DMUo in T_v is as follows:

$$
\begin{aligned}
\min\ Z = & \frac{\sum_{i=1}^{m}\theta_i}{m} \Bigg/ \frac{\sum_{k=1}^{s}\varphi_k}{s} \\
s.t.\ & \sum_{j=1}^{n}\lambda_j x_{ij} \le \theta_i\, x_{io}, \quad i=1,\dots,m \\
& \sum_{j=1}^{n}\lambda_j y_{kj} \ge \varphi_k\, y_{ko}, \quad k=1,\dots,s \\
& \sum_{j=1}^{n}\lambda_j = 1, \\
& \lambda_j \ge 0, \qquad j=1,\dots,n \\
& \theta_i \le 1, \qquad i=1,\dots,m \\
& \varphi_k \ge 1, \qquad k=1,\dots,s
\end{aligned}
\tag{1}
$$

If the optimal value $Z^*=1$, DMU_o is efficient, otherwise it is inefficient. Model (1) offers an efficiency score for each DMU, separately. The remainder of the study is dedicated to providing a methodology for the evaluation of the whole system.

3 The Proposed Approach for the Ranking and Selection of Variables Based on the System's Efficiency.

This section is composed of two parts. In the first part, we introduce a new methodology that incorporates the centralized scenario and the modified Russell model to evaluate the system with n DMUs, and some of the properties of the developed model will be mentioned. This is accompanied by a step-by-step approach to estimating the importance of indicators based on their effect on system efficiency. The efficiency of each DMU is evaluated separately in the traditional DEA; however, it is presumed in this analysis that all DMUs are regulated by a central decision-making unit which is intended to analyze the efficiency of the system as a whole. The following non-radial model, which is a combination of the centralized scenario of Lozano et al. [36] and the modified Russell process, is proposed as the model (2) for the assessment of the under-evaluation system.

In this model, the efficiency of the whole system is obtained. Notice that $\left(\sum_{j=1}^{n}\lambda_{jr}x_{ij},\sum_{j=1}^{n}\lambda_{jr}y_{kj}\right)\in T_v$, and therefore $\left(\sum_{r=1}^{n}\sum_{j=1}^{n}\lambda_{jr}x_{ij},\sum_{r=1}^{n}\sum_{j=1}^{n}\lambda_{jr}y_{kj}\right)$ is the summation of n DMU belonging to the T_v, some of them may be the same. We call the set containing the elements like $\left(\sum_{r=1}^{n}\sum_{j=1}^{n}\lambda_{jr}x_{ij},\sum_{r=1}^{n}\sum_{j=1}^{n}\lambda_{jr}y_{kj}\right)$ as T_v^g and it's frontier as ∂T_v^g. Therefore, it can be inferred that model (2) evaluates the group with compare it to the elements on the ∂T_v^g. Model (2) is clearly feasible and its optimum value is between 0 and 1.

$$
\begin{aligned}
\min\ & R = \frac{\sum_{i=1}^{m}\theta_i}{m} \Bigg/ \frac{\sum_{k=1}^{s}\varphi_k}{s} \\
s.t.\ & \sum_{r=1}^{n}\sum_{j=1}^{n}\lambda_{jr}x_{ij} \le \theta_i \sum\nolimits_{j=1}^{n} x_{ij}, \quad i = 1,\dots,m \\
& \sum_{r=1}^{n}\sum_{j=1}^{n}\lambda_{jr}y_{kj} \ge \varphi_k \sum\nolimits_{j=1}^{n} y_{kj}, \quad k = 1,\dots,s \\
& \sum_{j=1}^{n}\lambda_{jr} = 1, \qquad r = 1,\dots,n \\
& \lambda_{jr} \ge 0, \qquad j, r = 1,\dots,n \\
& \theta_i \le 1, \qquad i = 1,\dots,m \\
& \varphi_k \ge 1, \qquad k = 1,\dots,s
\end{aligned}
\tag{2}
$$

Definition. A group is technically efficient if the optimal value of model (2) is 1. Otherwise, it is inefficient. We now provide some properties of our proposed model.

Theorem 1. R is units invariant.

Proof. Notice that θ_i and φ_k are unit-invariant so that R is also unit-invariant. To see that this is the case, we note that these inequalities may be formulated as equations without affecting the optimal value of R^*. Then, writing $\theta_i = \sum_{r=1}^{n}\sum_{j=1}^{n}\lambda_{jr}x_{ij} \Big/ \sum_{r=1}^{n} x_{ir}\, , i = 1,\dots,m$ and $\varphi_k = \sum_{r=1}^{n}\sum_{j=1}^{n}\lambda_{jr}y_{kj} \Big/ \sum_{r=1}^{n} y_{kr}\, , k = 1,\dots,s$ we see that we can multiply numerators and denominators by $p_i, q_k > 0, i = 1,\dots,m,\ k = 1,\dots s$ respectively, without affecting the values of θ_i and φ_k .

Theorem 2. Each unit $\left(\hat{x}_{ir}, \hat{y}_{kr}\right) = \left(\sum_{j=1}^{n}\lambda^*_{jr}x_{ij}\ \ \forall i\, , \sum_{j=1}^{n}\lambda^*_{jr}y_{kj}\ \ \forall k\right)$ is Pareto efficient on T_v .

*Proof***:** By contradiction, suppose that $\left(\hat{x}_{ir'}, \hat{y}_{kr'}\right), i = 1,\dots,m, k = 1,\dots,s,$ is not Pareto efficient, and others are efficient. Then vectors $(\lambda_{1r},\dots,\lambda_{1r})$ can be found satisfying $\sum_{j=1}^{n}\lambda_{jr} = 1$ and defining a virtual operating point

$$
\begin{cases}
\hat{\hat{x}}_{ir} = \sum_{j=1}^{n}\lambda_{jr}x_{ij} \le \hat{x}_{ir} & \forall i \\
\hat{\hat{y}}_{ir} = \sum_{j=1}^{n}\lambda_{jr}y_{kj} \ge \hat{y}_{ir} & \forall k
\end{cases}
$$

such that the above inequality is strict for at least one input or one output index of the *r*th unit. Without loss of generality, assume that strict inequality occurs in the *i'*th component of the units (the proof procedure is similar for the case where inequality occurs in the output index or both input and output indices). Then

$$\begin{cases} \hat{\hat{x}}_{i'r'} = \sum_{j=1}^{n} \lambda_{jr'} x_{i'j} < \hat{x}_{i'r'} \\ \hat{\hat{x}}_{ir'} = \sum_{j=1}^{n} \lambda_{jr'} x_{ij} \le \hat{x}_{ir'} \quad \forall i \neq i' \\ \hat{\hat{y}}_{ir'} = \sum_{j=1}^{n} \lambda_{jr'} y_{kj} \ge \hat{y}_{ir'} \quad \forall k \end{cases}$$

By summation on index *r*, we will have:

$$\begin{cases} for\ index\ i':\ \sum_{r=1}^{n} \hat{\hat{x}}_{i'r} < \sum_{r=1}^{n} \hat{x}_{i'r} & (a) \\ for\ others:\ \sum_{r=1}^{n} \hat{\hat{x}}_{ir} \le \sum_{r=1}^{n} \hat{x}_{ir} & (b) \\ \qquad\qquad \sum_{r=1}^{n} \hat{\hat{y}}_{kr} \ge \sum_{r=1}^{n} \hat{y}_{kr} & (c) \end{cases} \qquad (3)$$

Dividing the two sides of (7-a) by $\sum_{r=1}^{n} x_{i'r}$, those of (7-b) by $\sum_{r=1}^{n} x_{ir}$, and those of (7-c) by $\sum_{r=1}^{n} y_{kr}$ we get:

$$\bar{\theta}_{i'} = \sum_{r=1}^{n} \hat{\hat{x}}_{i'r} \Big/ \sum_{r=1}^{n} x_{i'r} < \sum_{r=1}^{n} \hat{x}_{i'r} \Big/ \sum_{r=1}^{n} x_{i'r} = \theta_{i'}^{*}$$

$$\bar{\theta}_{i} = \sum_{r=1}^{n} \hat{\hat{x}}_{ir} \Big/ \sum_{r=1}^{n} x_{ir} \le \sum_{r=1}^{n} \hat{x}_{ir} \Big/ \sum_{r=1}^{n} x_{ir} = \theta_{i}^{*}$$

$$\bar{\varphi}_{k} = \sum_{r=1}^{n} \hat{\hat{y}}_{kr} \Big/ \sum_{r=1}^{n} y_{kr} \ge \sum_{r=1}^{n} \hat{y}_{kr} \Big/ \sum_{r=1}^{n} y_{kr} = \varphi_{k}^{*}$$

Therefore, $\left(\bar{\lambda}_{jr}, \bar{\theta}_i, \bar{\varphi}_k\right)$ is a feasible solution for (2) and its objective function value is smaller than R^* , which contradicts the optimality of $\left(\lambda_{jr}^{*}, \theta_i^{*}, \varphi_k^{*}\right)$. The contradiction assumption is thus invalid, and hence $\left(\hat{x}_{ir'}, \hat{y}_{kr'}\right), i = 1,\dots,m, k = 1,\dots,s$ is Pareto efficient for every index *r*.

Theorem 3. $\left(\sum_{r=1}^{n}\sum_{j=1}^{n} \lambda_{jr}^{*} x_{ij} \quad \forall i, \sum_{r=1}^{n}\sum_{j=1}^{n} \lambda_{jr}^{*} y_{kj} \quad \forall k \right) \in T_v^g$ is Pareto efficient.

*Proof***.** Suppose that $\left(\lambda_{jr}^{*}, \theta_i^{*}, \varphi_k^{*}\right)$ is the optimal solution, R^*, of the model (2), and $\left(\sum_{r=1}^{n} \hat{x}_{ir}, \sum_{r=1}^{n} \hat{y}_{kr} \right) = \left(\sum_{r=1}^{n}\sum_{j=1}^{n} \lambda_{jr}^{*} x_{ij}, \sum_{r=1}^{n}\sum_{j=1}^{n} \lambda_{jr}^{*} y_{kj} \right)$. Since the constraints are expressed as equality at optimality, by Theorem (2), we have:

$$\theta_i^{*} = \sum_{r=1}^{n}\sum_{j=1}^{n} \lambda_{jr}^{*} x_{ij} \Big/ \sum_{r=1}^{n} x_{ir} = \sum_{r=1}^{n} \hat{x}_{ir} \Big/ \sum_{r=1}^{n} x_{ir}$$

$$\varphi_k^{*} = \sum_{r=1}^{n}\sum_{j=1}^{n} \lambda_{jr}^{*} y_{kj} \Big/ \sum_{r=1}^{n} y_{kr} = \sum_{r=1}^{n} \hat{y}_{kr} \Big/ \sum_{r=1}^{n} y_{kr}$$

Now, by contradiction suppose that $\left(\sum_{r=1}^{n}\hat{x}_{ir},\sum_{r=1}^{n}\hat{y}_{kr}\right)$ is not Pareto efficient. Then, there exists $\left(\sum_{r=1}^{n}\tilde{x}_{ir},\sum_{r=1}^{n}\tilde{y}_{kr}\right)\in T_v^g$ such that $\left(\tilde{x}_{ir},\tilde{y}_{kr}\right)\in T_v$ and

$$\left(\sum_{r=1}^{n}\tilde{x}_{ir},-\sum_{r=1}^{n}\tilde{y}_{kr}\right)\underset{\neq}{\leq}\left(\sum_{r=1}^{n}\hat{x}_{ir},-\sum_{r=1}^{n}\hat{y}_{kr}\right) \qquad (4)$$

Without loss of generality, assume that strict inequality occurs in the *i'*th component of the inputs (the proof procedure is similar for the case where inequality occurs in the output index or both input and output indices). Since $\left(\tilde{x}_{ir},\tilde{y}_{kr}\right)\in T_v$, there exists a $\lambda_{jr}\geq 0$ for every *r* such that $\sum_{j=1}^{n}\lambda_{jr}=1,\ \forall r$ and

$$\begin{cases}\tilde{\tilde{x}}_{ir}=\sum_{j=1}^{n}\lambda_{jr}x_{ij}\leq\tilde{x}_{ir}\\ \tilde{\tilde{y}}_{ir}=\sum_{j=1}^{n}\lambda_{jr}y_{kj}\geq\tilde{y}_{ir}\end{cases}$$

By summation on index *r*, we will have:

$$\begin{cases}\tilde{\tilde{x}}_{ir}=\sum_{r=1}^{n}\sum_{j=1}^{n}\lambda_{jr}x_{ij}\leq\sum_{r=1}^{n}\tilde{x}_{ir}\\ \tilde{\tilde{y}}_{ir}=\sum_{r=1}^{n}\sum_{j=1}^{n}\lambda_{jr}y_{kj}\geq\sum_{r=1}^{n}\tilde{y}_{kr}\end{cases}$$

Now, by (4) we will have:

$$\begin{cases}for\ index\ i':\ \sum_{r=1}^{n}\sum_{j=1}^{n}\lambda_{jr}x_{i'j}\leq\sum_{r=1}^{n}\tilde{x}_{i'r}<\sum_{r=1}^{n}\hat{x}_{i'r} & (a)\\ for\ others:\ \sum_{r=1}^{n}\sum_{j=1}^{n}\lambda_{jr}x_{ij}\leq\sum_{r=1}^{n}\tilde{x}_{ir}\leq\sum_{r=1}^{n}\hat{x}_{ir} & (b)\\ \qquad\qquad\quad \sum_{r=1}^{n}\sum_{j=1}^{n}\lambda_{jr}y_{kj}\geq\sum_{r=1}^{n}\tilde{y}_{kr}\geq\sum_{r=1}^{n}\hat{y}_{kr} & (c)\end{cases} \qquad (5)$$

By dividing the two sides of (9-a) by $\sum_{r=1}^{n}x_{i'r}$, those of (9-b) by $\sum_{r=1}^{n}x_{ir}$, and those of (9-c) by $\sum_{r=1}^{n}y_{kr}$, we get:

$$\bar{\theta}_{i'}=\sum_{r=1}^{n}\sum_{j=1}^{n}\lambda_{jr}x_{i'j}\Big/\sum_{r=1}^{n}x_{i'r}<\sum_{r=1}^{n}\hat{x}_{i'r}\Big/\sum_{r=1}^{n}x_{i'r}=\theta_{i'}^{*}$$

$$\bar{\theta}_{i}=\sum_{r=1}^{n}\sum_{j=1}^{n}\lambda_{jr}x_{ij}\Big/\sum_{r=1}^{n}x_{ir}\leq\sum_{r=1}^{n}\hat{x}_{ir}\Big/\sum_{r=1}^{n}x_{ir}=\theta_{i}^{*}$$

$$\bar{\varphi}_{k}=\sum_{r=1}^{n}\sum_{j=1}^{n}\lambda_{jr}y_{kj}\Big/\sum_{r=1}^{n}y_{kr}\geq\sum_{r=1}^{n}\hat{y}_{kr}\Big/\sum_{r=1}^{n}y_{kr}=\varphi_{k}^{*}$$

Therefore, $\left(\bar{\lambda}_{jr}, \bar{\theta}_i, \bar{\varphi}_k\right)$ is a feasible solution for (2) and its objective function value is smaller than R^*, which contradicts the optimality of $\left(\lambda^*_{jr}, \theta^*_i, \varphi^*_k\right)$. The contradiction assumption is thus invalid and $\left(\sum_{r=1}^{n} \hat{x}_{ir}, \sum_{r=1}^{n} \hat{y}_{kr}\right)$ is therefore Pareto efficient.

Theorem 4. Suppose that $\left(\sum_{r=1}^{n} x^A_r, \sum_{r=1}^{n} y^A_r\right) \in T^g_v$ and $\left(\sum_{r=1}^{n} x^B_r, \sum_{r=1}^{n} y^B_r\right) \in T^g_v$, and that R^A and R^B are the optimal values of the model (2) in evaluating them, respectively. So, if $\left(\sum_{r=1}^{n} x^A_r, -\sum_{r=1}^{n} y^A_r\right) \underset{\neq}{\leq} \left(\sum_{r=1}^{n} x^B_r, -\sum_{r=1}^{n} y^B_r\right)$, then $R^A < R^B$.

Proof. The proof is clear.

Suppose that $\left(\lambda'_{jr}, \theta'_{ir}, \varphi'_{kr}\right)$ is the optimal solution obtained by evaluating unit r by the enhanced Russell model and $\left(\sum_{j=1}^{n} \lambda'_{jr} x_{ij}, \sum_{j=1}^{n} \lambda'_{jr} y_{kj}\right)$ is the projection of unit r onto the frontier T_v.

Theorem 5. Suppose that R' is the optimal value of model (2) in evaluating $\left(\sum_{r=1}^{n}\sum_{j=1}^{n} \lambda'_{jr} x_{ij}, \sum_{r=1}^{n}\sum_{j=1}^{n} \lambda'_{jr} y_{kj}\right)$. Then $R' \leq R^*$.

Proof. The proof is clear.

3.1 Converting the Model to Linear Programming

Model (2) is a fractional programming model which can be transformed into the following programming model using Charnes and Cooper [37] transformations. For this purpose, new variables should be introduced.

$$\beta = \left(\frac{\sum \varphi_k}{s}\right)^{-1}, \qquad 0 < \beta \leq 1$$

$$u_i = \beta \theta_i, \quad i = 1, \dots, m$$

$$v_k = \beta \varphi_k, \quad k = 1, \dots, s$$

$$t_{jr} = \beta \lambda_{jr}, \quad j, r = 1, \dots, n$$

Using these transformations, model 2 becomes as follows:

$$
\begin{aligned}
\min\ & \frac{\sum_{i=1}^{m} u_i}{m} \\
s.t.\ & \sum_{r=1}^{n}\sum_{j=1}^{n} t_{jr} x_{ij} \le u_i \sum_{r=1}^{n} x_{ir} \qquad i=1,\dots,m \\
& \sum_{r=1}^{n}\sum_{j=1}^{n} t_{jr} y_{kj} \ge v_k \sum_{r=1}^{n} y_{kr} \qquad k=1,\dots,s \\
& \sum_{j=1}^{n} t_{jr} = \beta, \qquad r=1,\dots,n \\
& t_{jr} \ge 0, \qquad j,r=1,\dots,n \\
& u_i \le \beta, \qquad i=1,\dots,m \\
& v_k \ge \beta, \qquad k=1,\dots,s \\
& \sum_{k=1}^{s} v_k = s \\
& 0 \le \beta \le 1
\end{aligned} \tag{6}
$$

Using the following transformation, the optimal solution for Model (2) is obtained:

$$\theta_i = u_i / \beta, \quad i=1,\dots,m$$
$$\varphi_k = v_k / \beta, \quad k=1,\dots,s$$
$$\lambda_{jr} = t_{jr} / \beta, \quad j,r=1,\dots,n$$

Because of the linearity of the model (6), it can be easily solved.

3.2 Determining the Importance of Variables and Their Ranking

Assuming that there are n DMU with m input variable (i=1,…, m) and s output variable (r= 1,…, s), the following approach is proposed for determining the importance and priority of input and output variables:

1st step: Implementing model (2) to obtain the overall system's efficiency: Before removing any factor, we mark the system efficiency with *E'*.

2nd step: The elimination of inputs and/or outputs, one by one, and solving model (2). We call the system efficiency after elimination *E**.

3rd step: Calculating the difference between *E'* and *E** (D=|E- E*|), and determining the most important input or output based on a larger value for D.

4th step: Implementing the model until at least one input or output is left or stopping at a predetermined stage by the decision maker. The input or output variable is of high significance if the value of D is large, and we maintain the input or output. If it is small, it means that the input and output are of low importance and can be ignored. In this way, the inputs and outputs that are less significant are removed and the number of indicators decreases. Figure 1 shows the flowchart of the proposed method.

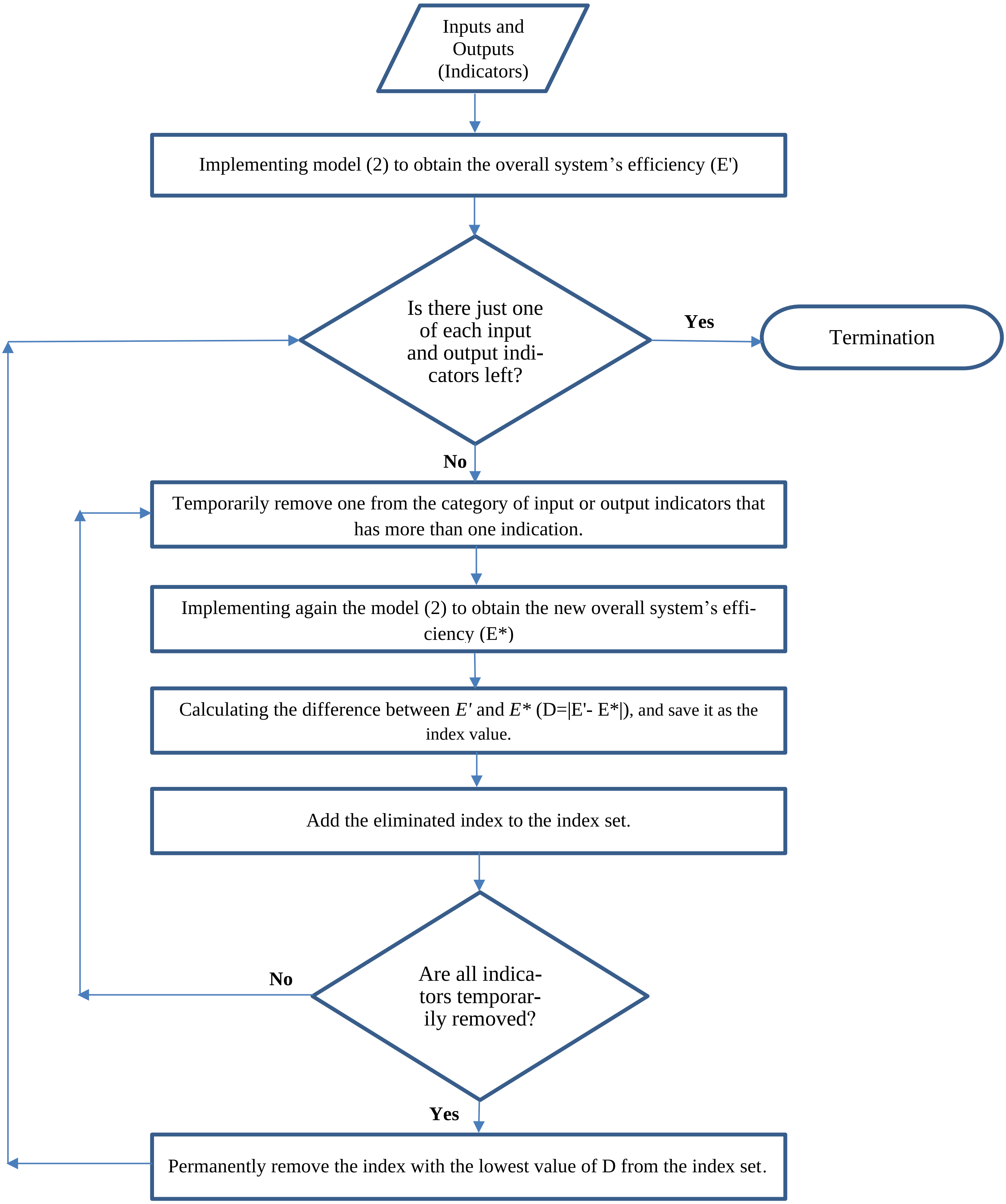


Fig.1: The Flowchart of the Proposed Method

4 Application in Bank Branches

In this section, the performance of 14 branches of Refah-Kargaran bank in Lorestan province is evaluated using the proposed algorithm in the previous section. Eight important indicators were considered for the assessment of the performance of the branches of Refah Bank. The mean of resources, the mean of expenditures, the volume of activity, and transactions have been considered output indicators. The mean of demands, the number of personnel, the mean of costs, and the physical space of the branches were considered as inputs. Data from 14 branches and all variables are seen in Table 1. In this table, the inputs and outputs are I and O, respectively.

I_1: Number of personnel	O_1: The mean of resources
I_2: The mean of costs	O_2: The mean of expenditures
I_3: Branch space	O_3: the volume of activity
I_4: The mean of demands	O_4: Transaction

Table1: Inputs and Outputs Data for the Practical Example

Branches	I_1	I_2	I_3	I_4	O_1	O_2	O_3	O_4
1	8	1274	135	3115	41705	30320	161409	1131
2	11	1541	535	20848	47089	105106	450970	10107
3	13	1877	542	6102	62469	61308	380183	3673
4	8	1242	295	4244	38308	39698	249834	2406
5	7	911	392	2469	25042	27418	176162	2231
6	13	1613	627	12341	102939	115268	356593	11169
7	5	863	124	13766	23742	41855	165422	3406
8	6	830	240	4715	37849	60049	192622	3779
9	7	1442	320	14284	44825	48750	251760	1357
10	4	560	165	3753	17856	21926	100930	1359
11	5	722	466	4802	25911	30376	166796	2859
12	6	901	552	4015	30402	26215	174644	3024
13	7	1154	274	731	42361	19558	250478	0
14	6	889	411	6112	26304	36388	216235	2828

Source: [38]

It is experimentally proposed that the number of branches should be more than three times the total number of inputs and outputs in order to maximize the discrimination power in the DEA {3(m+s)=< n}. This means that the number of indicators must be at most one-third of the number of units, which for the present example is 5. However, there are eight indicators available in this study. When all of the indicators are used to measure the efficiency of all branches using the modified Russel model (3), the values in Table 2 are obtained.

Table 2: Efficiency of All Branches using Russell Model and Considering all Indicators

DMU	DMU1	DMU2	DMU3	DMU4	DMU5	DMU6	DMU7
E	1	1	0.76	0.78	1	1	1
DMU	DMU8	DMU9	DMU10	DMU11	DMU12	DMU13	DMU14
E	1	1	0.58	0.68	0.61	0.91	0.74

As can be seen in Table 2, the number of efficient units is high and it is difficult to differentiate between units. Therefore, by reducing the number of inputs and outputs, a better distinction must be obtained. The proposed approach of system efficiency is used to reduce the number of indicators. While considering all the indicators, the efficiency of the system consisting of all bank branches using the model (3) proposed in this paper is equal to 0.7.

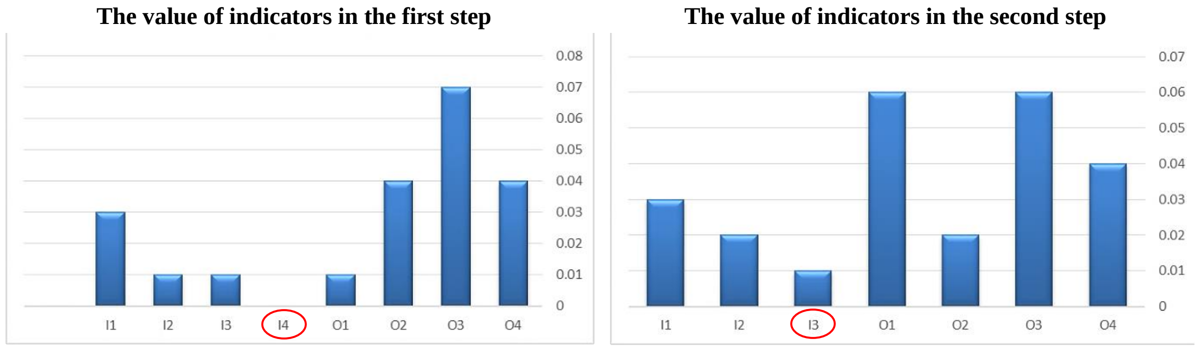
Table 3: The Results Obtained from Utilizing the Suggested Model

Efficiency	I_1	I_2	I_3	I_4	O_1	O_2	O_3	O_4
E^*_1	0.76	0.71	0.71	0.70	0.69	0.74	0.63	0.74
$D_1 = \|E^*_1 - E_1\|$	0.03	0.01	0.01	0.00	0.01	0.04	0.07	0.04
E^*_2	0.67	0.68	0.71		0.64	0.72	0.64	0.74
$D_2 = \|E^*_2 - E_2\|$	0.03	0.02	0.01		0.06	0.02	0.06	0.04
E^*_3	0.64	0.74			0.66	0.75	0.65	0.75
$D_3 = \|E^*_3 - E_3\|$	0.07	0.03			0.05	0.06	0.06	0.04
E^*_4	0.07				0.67	0.78	0.69	0.79
$D_4 = \|E^*_4 - E_4\|$	-				0.07	0.04	0.05	0.05
E^*_5	0.07				0.64		0.69	0.85
$D_5 = \|E^*_5 - E_5\|$	-				0.11		0.69	0.07
E^*_6	0.07				0.63		0.77	
$D_6 = \|E^*_6 - E_6\|$	-				0.22		0.13	

Now, we assess the effect of each factor on the efficiency of the system by eliminating the indicators one by one. If the first input is removed from the indicators' set and model 2 is solved again, the system's efficiency will be 0.76. This indicates that the amount of change in the system's efficiency is $|0.76 - 0.70| = 0.06$. If the second input is removed from the indicators' set, the system's efficiency will be 0.71 and the amount of change in the system's efficiency is $|0.71 - 0.70| = 0.01$. In this way, the system's efficiency is achieved by deleting individual inputs and outputs. These values are listed in the first row of Table 3. In the second row of the table, the difference in system efficiency before and after the removal of the indicators is shown. As can be seen from the second row of the table, the value of the difference in system performance before and after the removal of the fourth input is almost zero (numbers are rounded to two decimal places). This means that deleting this indicator does not cause much change in system performance. Since this index has the least impact, in the first step we exclude this index from the evaluation. We continue the analysis with 7 indicators including 3 inputs and 4 outputs and eliminate the other indicators one by one again. The system performance in this case, is shown with E^*_2, and the difference with the system performance before removing the second stage indicators can be seen in the line below. This time the third input is selected for deletion. We will continue this process. In the last row of Table 3, the third output (the volume of activity) has a lower impact on system efficiency than the first output (the mean of resources). But note that the choice is such that we always have at least one input and one output. This means that for example, not all 4 inputs or all 4 outputs can be deleted. As is seen in the last row of the table, the number of employees from the input indicators, and the mean of demands from the output indicators are the most important ones. See also Figure 2 which graphically describes the steps of the proposed model in ranking the criteria in the evaluation of bank Branches.

5 Conclusion

Evaluating the efficiency of organizations, especially banks, is one of the most important issues in the field of management. The DEA technique is a very powerful tool for evaluating organizations, taking into account various input and output indicators. Although DEA is considered non-parametric, the sample size may be a significant issue in establishing the efficiency ratings for the assessed units empirically, because the usage of too many inputs and outputs may result in a substantial number of DMUs being scored as efficient. Empirical guidelines have been created in the DEA literature to prevent too many DMUs from being evaluated as efficient.

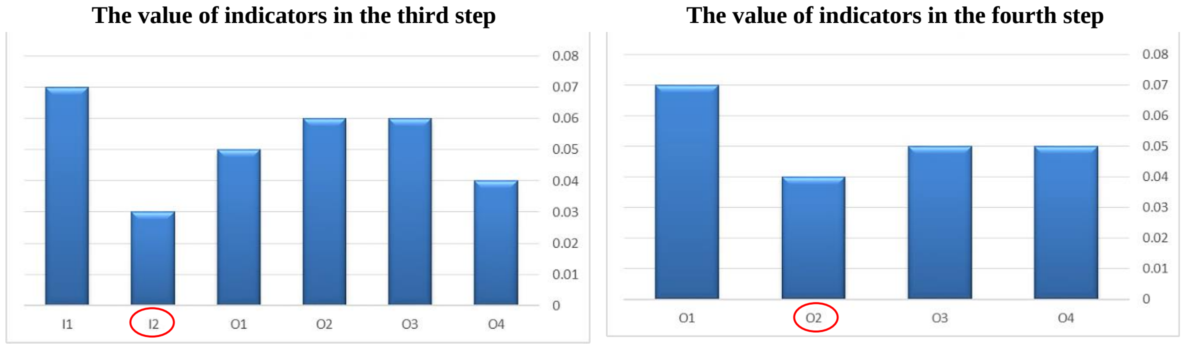


The value of indicators in the first step

The value of indicators in the second step

Input 4 has the lowest value and is removed forever..

Input 3 has the lowest value and is removed forever..

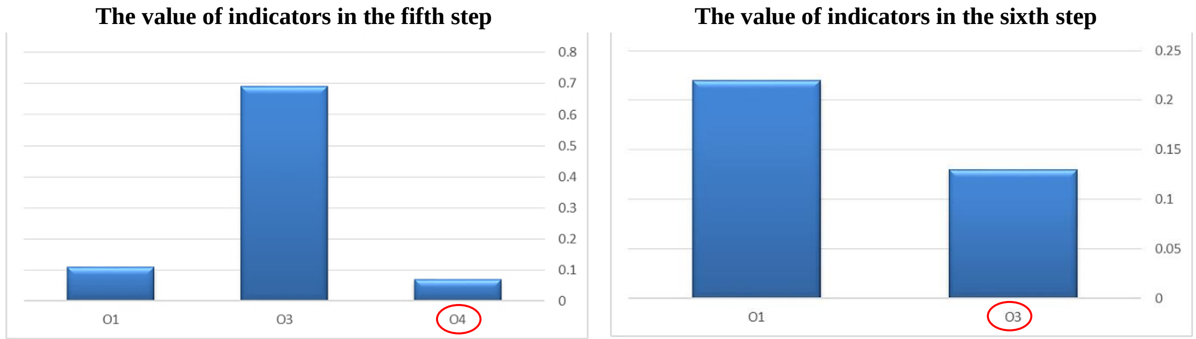


The value of indicators in the third step

The value of indicators in the fourth step

.Input 3 has the lowest value and is removed forever.

Since in the previous step only the first input remained, we do not examine it anymore, and in step 4 the output 2 has the lowest value and is permanently eliminated.



The value of indicators in the fifth step

The value of indicators in the sixth step

Output 4 has the lowest value and is permanently removed.

Output 3 has the lowest value and is eliminated. There are just two indications left, the first input and the first output.

Fig. 2: Steps of the Proposed Model for Applied Example

These empirical cutoffs link the number of variables to the number of observations. So in evaluating the bank branches, if the number of indicators used to evaluate bank branches is large, a large number of them are known as efficient in common DEA models. In this study, eight indicators were provided. When all of the measures were used to assess the efficiency of all 14 branches of Refah-Kargaran Bank in the province of Lorestan, 8 of 14, or 50%, were deemed efficient. As a result, they cannot be properly distinguished from each other in terms of performance. In order to minimize the number of input and output variables, in the present study, we first introduced a method based on a modified Russell model to evaluate system performance.

Using the model developed, the bank's efficiency with 14 branches was estimated to be 0.7 after considering all variables. Then we used a step-by-step approach to rank the indicators based on their effect on the efficiency of the whole system. In this way, an indicator whose removal causes a slight change in system performance is considered insignificant and can be excluded from the analysis. These factors can be ranked based on this criterion. These indicators can be ranked as follows in order of low value: the mean of demands, branch space, the mean of costs, the mean of expenditures, transactions, and the volume of activity. The results indicated that, among the input indicators, the average of demands was the least important and eliminated in the first phase, and the number of personnel was the most important of all. Also, among the output indicators, the mean of expenditures and the mean of resources were the least and the most important factors, respectively. In this study, the efficiency of bank branches was investigated in a situation in which all data were definite and related to a specific time interval. In further research, indefinite data can be considered. Also, the approach can be applied in a way that the inputs' prices are at hand. The reduction, selection, and ranking of indicators based on their impact on the productivity of each unit or the efficiency of the system can also be studied. Finally, the modern structures of DEA such as network structure can be investigated.

References

[1] Sufian, F., *Determinants of bank efficiency during unstable macroeconomic environment: Empirical evidence from Malaysia,* Res. Int. Bus. Financ, 2009, **23**, P.54–77, Doi:10.1016/j.ribaf.2008.07.002.

[2] Manandhar, R., Tang, J.C.S.,*The evaluation of bank branch performance using data envelopment analysis*, A framework. J. High Technol. Manag. Res, 2002, **13**, P. 1–17, Doi:10.1016/S1047-8310(01)00045-1.

[3] Frei, F.X., Harker, P.T., *Measuring aggregate process performance using AHP.* Eur. J. Oper. Res, 1999, **116**, P.436–442, Doi:10.1016/S0377-2217(98)00134-9.

[4] Lotto, J., Papavassiliou, V., *Evaluation of factors influencing bank operating efficiency in Tanzanian banking sector,* Cogent Econ. Financ, 2019, **7**, 1664192, Doi:10.1080/23322039.2019.1664192.

[5] Wang, K., Huang, W., Wu, J., Liu, Y.N., *Efficiency measures of the Chinese commercial banking system using an additive two-stage DEA,* Omega (United Kingdom) 2014, **44**, P.5–20, Doi:10.1016/j.omega.2013.09.005.

[6] Ahadzadeh Namin, M., Khamseh, E., Mohamadi, F., *Evaluate the performance of bank branches using the control approach in analyzing the data cover weight,* 2019, **10**, P.1–28.

[7] Henriques, I.C., Sobreiro, V.A., Kimura, H., Mariano, E.B., *Efficiency in the Brazilian banking system using data envelopment analysis,* Futur. Bus. J. 2018, **4**, P.157–178, Doi: 10.1016/j.fbj.2018.05.001.

[8] Jiang, H., He, Y., *Applying data envelopment analysis in measuring the efficiency of Chinese listed banks in the context of macroprudential framework. Mathematics 2018,* **6**, P.184, Doi:10.3390/math6100184.

[9] Razipour-GhalehJough, S., Hosseinzadeh Lotfi, F., Jahanshahloo, G., Rostamy-malkhalifeh, M., Sharafi, H. *Finding closest target for bank branches in the presence of weight restrictions using data envelopment analysis,* Ann. Oper. Res. 2020, **288**, P.755–787, Doi:10.1007/s10479-019-03166-6.

[10] Ihaddaden, M.E.F., *Investigating Eurosystem Central Banking Efficiency: A Data Envelopment Analysis Approach,* 2019, Doi:10.2139/SSRN.3326536.

[11] Yu, M.M., Lin, C.I., Chen, K.C., Chen, L.H. *Measuring Taiwanese bank performance: A two-system dynamic network data envelopment analysis approach.* Omega (United Kingdom) 2021, **98**, 102145, Doi: 10.1016/j.omega.2019.102145.

[12] abidin, z., prabantarikso, r.m., wardhani, r.a., endri, e. *Analysis of Bank Efficiency Between Conventional Banks and Regional Development Banks in Indonesia.* J. Asian Financ. Econ. Bus. 2021, **8**, P.741–750. Doi: 10.13106/jafeb.2021.vol8.no1.741.

[13] Shair, F., Shaorong, S., Kamran, H.W., Hussain, M.S., Nawaz, M.A., Nguyen, V.C. *Assessing the efficiency and total factor productivity growth of the banking industry: do environmental concerns matters?* Environ. Sci. Pollut. Res. 2021, **28**, P.20822–20838, Doi:10.1007/s11356-020-11938-y.

[14] An, Q., Wu, Q., Zhou, X., Chen, X., *Closest target setting for two-stage network system: An application to the commercial banks in China,* Expert Syst. Appl. 2021, **175**, P.114799, Doi: 10.1016/j.eswa.2021.114799.

[15] Asmild, M., Kronborg, D., Mahbub, T., Matthews, K., *Inefficiency patterns in family-owned banks in Bangladesh.* J. Econ. Stud. 2021, Doi:10.1108/JES-06-2020-0286.

[16] Azad, M.A.K., Talib, M.B.A., Kwek, K.T., Saona, P., *Conventional versus Islamic bank efficiency: A dynamic network Data-Envelopment-Analysis approach.* J. Intell. Fuzzy Syst. 2021, **40**, P.1921–1933, Doi:10.3233/JIFS-189196.

[17] Paulet, E., Mavoori, H. *Cross-regional analysis of banking efficiency drivers.* Appl. Econ. 2021, **53**, P.2042–2065, Doi:10.1080/00036846.2020.1855312.

[18] Sáez-Fernández, F.J., Picazo-Tadeo, A.J., Jiménez-Hernández, I. *Performance and risk in the Brazilian banking industry.* Heliyon 2021, **7**, e06524, Doi: 10.1016/j.heliyon.2021.e06524.

[19] Lartey, T., James, G.A., Danso, A., *Interbank funding, bank risk exposure and performance in the UK: A three-stage network DEA approach. Int.* Rev. Financ. Anal. 2021, **75**, 101753, Doi: 10.1016/j.irfa.2021.101753.

[20] Kazemi, S., Tavana, M., Toloo, M., Zenkevich, N.A. *A common weights model for investigating efficiency-based leadership in the russian banking industry.* RAIRO-Oper. Res. 2021, **55**, P.213–229, Doi:10.1051/ro/2020143.

[21] Bou-Hamad, I., Anouze, A.L., Osman, I.H., *A cognitive analytics management framework to select input and output variables for data envelopment analysis modeling of performance efficiency of banks using random forest and entropy of information.* Ann. Oper. Res. 2021, P.1–30, Doi:10.1007/s10479-021-04024-0.

[22] Lotfi, F.H., Lotfi, F.H., Navabakhs, M., Tehranian, A., Rostamy-malkhalifeh, M., Shahverdi, R. *Ranking Bank Branches with Interval Data the Application of DEA.* Int. Math. Forum 2007, **9**, P.429–440.

[23] Shafiee, M., *Bank branches efficiency assessment using dynamic data envelopment analysis approach to SBM.* J. New Res. Math. 2017, **3**, P.5–18.

[24] Charles, V., Aparicio, J., Zhu, J., *The curse of dimensionality of decision-making units: A simple approach to increase the discriminatory power of data envelopment analysis.* Eur. J. Oper. Res. 2019, **279**, P.929–940,

Doi: 10.1016/j.ejor.2019.06.025.

[25] Dar, Q.F., Ahn, Y.-H., Dar, G.F. *The Impact of International Trade on Central Bank Efficiency: An Application of DEA and Tobit Regression Analysis. Stat. Optim. Inf. Comput. 2021,* **9**, P.223–240, Doi:10.19139/soic-2310-5070-1077.

[26] Eskelinen, J., *Comparison of variable selection techniques for data envelopment analysis in a retail bank.* Eur. J. Oper. Res. 2017, **259**, P.778–788, Doi: 10.1016/j.ejor.2016.11.009.

[27] Niknafs, J., Keramati, M.A., Haghighatmonfared, J., *Estimating Efficiency of Bank Branches by Dynamic Network Data Envelopment Analysis and Artificial Neural Network. Adv.* Math. Financ. Appl. 2020, **5**, P. 377–390, Doi:10.22034/AMFA.2019.1585957.1192.

[28] Akbari, S., Heydari, J., Keramati, M., Keramati, A., *Designing A Mixed System of Network DEA for Evaluating the Efficiency of Branches of Commercial Banks in Iran,* Adv. Math. Financ. Appl. 2019, **4,** P.1–13, Doi:10.22034/AMFA.2019.582260.1165.

[29] Esfandiar, M., Saremi, M., Jahangiri Nia, H., *Assessment of the efficiency of banks accepted in Tehran Stock Exchange using the data envelopment analysis technique,* Adv. Math. Financ. Appl. 2018, **3**, P.1–11, Doi:10.22034/AMFA.2018.540815.

[30] Fanchon, P., *Variable selection for dynamic measures of efficiency in the computer industry,* Int. Adv. Econ. Res. 2003, **9**, P.175–188, Doi:10.1007/BF02295441.

[31] Wagner, J.M., Shimshak, D.G., *Stepwise selection of variables in data envelopment analysis: Procedures and managerial perspectives,* Eur. J. Oper. Res. 2007, **180**, P.57–67, Doi: 10.1016/j.ejor.2006.02.048.

[32] Madhanagopal, R., Chandrasekaran, R., *Selecting Appropriate Variables for DEA Using Genetic Algorithm (GA) Search Procedure.* Int. J. Data Envel. Anal. Operations Res. 2014, **1**, P.28–33, Doi:10.12691/IJDEAOR-1-2-3.

[33] Fernandez-Palacin, F., Lopez-Sanchez, M.A., Muñoz-Márquez, M., *Stepwise selection of variables in dea using contribution loads,* Pesqui. Operacional 2018, **38**, P.31–52, Doi:10.1590/0101-7438.2018.038.01.0031.

[34] Adler, N., Yazhemsky, E., *Improving discrimination in data envelopment analysis: PCA-DEA or variable reduction.* Eur. J. Oper. Res. 2010, **202**, P.273–284, Doi: 10.1016/j.ejor.2009.03.050.

[35] Pastor, J.T., Ruiz, J.L., Sirvent, I., *Enhanced DEA Russell graph efficiency measure,* Eur. J. Oper. Res. 1999, **115**, P.596–607, Doi:10.1016/S0377-2217(98)00098-8.

[36] Lozano, S., Villa, G., *Centralized resource allocation using data envelopment analysis,* J. Product. Anal. 2004, **22**, P.143–161, Doi:10.1023/b:prod.0000034748.22820.33.

[37] Charnes, A., Cooper, W.W. *Programming with linear fractional functionals,* Nav. Res. Logist. Q. 1963, **10**, P.273–274, Doi:10.1002/nav.3800100123.

[38] Moazzami Gudarzi, M., Jaberansari, M., Moallem, A., Shakiba, M., *Applying Data Envelopment Analysis (DEA) for Measuring Relative Efficiency and Ranking Branches of Refah Kargaran Bank in Lorestan Province.* QJER 2014, **14**(1), P.115-126.