Comprehensive Performance Evaluation of Banking Branches: A Three-Stage Slacks-Based Measure (SBM) Data Envelopment Analysis

Mohammad Zarei Mahmoudabadi, Ali Emrouznejad

1 Management Department, Meybod University, Meybod, Iran
2 Aston Business School, Aston University, Birmingham, UK

Abstract

The productivity of the banks in any country is a key factor in the growth and development of that country's economy. Recently, the evaluation and improvement of the productivity of the banking industry has been taken into much consideration in Iran. Data Envelopment Analysis (DEA) is a comprehensive and accepted approach for assessing the performance of banking industry. Although extensive studies have been done on banking industry using standard DEA models, they are, in fact, they ignore the internal structure of bank performance. Since the overall operational process of the banking system is made up of several partial processes, network DEA models are used to take into account all the internal components of the process and the coherence of the whole process. This is also done as the evaluation of the efficiency of partial processes helps to identify the sources of inefficiency of the overall banking system. In the present study, a network Slacks-Based Measure (SBM) DEA model is used in which the efficiency of the overall system is equal to the weighted average of the efficiency of the individual stages. The main advantage of this model is its ability to provide better efficiency criteria, calculate the weight of each stages separately, and simultaneously evaluate the mediator variables as both input and output. Finally, the comprehensive performance evaluation of banking industry is designed in three divisions, namely, production, intermediation, and social welfare approach. The model is applied to simultaneously evaluate operational efficiency, service effectiveness, and social effectiveness for 37 branches of one of the largest commercial banks in Iran.

Keywords: Network DEA; Slacks-Based Measure (SBM); Production approach; Intermediary approach; Banking efficiency.

* Corresponding author: Ali Emrouznejad, Professor and Chair in Business Analytics, Aston Business School, Aston University, Birmingham, UK, Web: www.Emrouznejad.com, Email: a.emrouznejad@aston.ac.uk
1. Introduction

Banking industry, as the main source of financial intermediation and payment system channel, plays a crucial role in the development of financial system (Paradi and Zhu, 2013). The efficiency and quality of services offered by banks not only have a significant impact on a country's economic growth, but also affect different aspects of people's daily life. With the increase in domestic and foreign competition and the provision of various services and products by banks, there is a serious need to improve the performance of the bank branches in order for them to stay in competition. The productivity of the banks of any country is a key factor in the growth and development of its economy and creation of job opportunities; Iran is no exception in this regard. Assessing and improving the efficiency of the banking system of Iran has been very seriously addressed recently. Performance evaluation can be seen as a benchmark tool by which a company's director can estimate his operational activity, examine the extent of achieving his goals, identify the position of his company in the same industry, and present some guidelines for the improvement of the units that perform inefficiently (Lin and Chiu, 2013). From the strategic management viewpoint, it is also really important for banks to continuously evaluate their efficiency through self-assessment, because doing so is very effective in achieving their performance improvement in the long run. In today's competitive business environment, there is a growing need to identify inefficient resources that may not be found using common approaches. In the recent decades, managerial and academic literature has increasingly focused on the measurement of organization's performance, as it leads to staff motivation, decision support, improvement in organizational learning, continuous improvement, and increased communication and coordination (Chiesa et al., 2009). Due to the internationalization and liberalization of global banking, banks have spread rapidly over the past two decades. To counter this competitive environment, many bank officials and academic researchers have worked to find ways to improve the performance of banks (Huang et al., 2012). In addition to high economic importance, the growing competitive market has highlighted the importance of assessing banks' performance in order to continuously improve their operations and monitor their financial conditions. There are lots of implementations for assessing and analyzing banks' performance with regard to the identification of sources of operational inefficiencies, gaps in the allocation of effective sources, effects of the changes in current laws and regulations on bank operations, and their capability of enabling bank managers to redirect their businesses towards the current business trends, profitable business processes, etc. (Paradi and Zhu, 2013).
The main objective of any financial organization is to improve its performance; thus, performance evaluation is considered as one of the best ways for operational improvement of organizations. By applying a variety of performance measurement methods, organizations can evaluate the effectiveness and efficiency of processes that are relevant to the organization's strategic goals. In addition, performance measurement tools can help organizations to allocate and distribute resources in a better way. Also, the advances in technology and globalization of financial services as well as the removal of government supervision from the banking industry have made banks vulnerable to the increasing competitive pressure (Roghanian et al., 2012). Since performance awareness is vital in all aspects of organizations, various methods have been proposed for measuring performance. The rapid change of banking services and other environmental drivers set continuously new demands for bank’s performance measurement systems. As a result, results gained from traditional measures are not adequate anymore (Harker and Zenios, 2000). Intensive competition in the banking industry has been one of the main drivers of change towards the more holistic performance measurement and management. It has made it necessary for financial institutions and banks to use new and comprehensive approach in performance evaluation, and thus make it possible for the bank to achieve its strategic and operative goals by providing adequate information. Due to the nature of the business and focus on creating money in banking industry, financial institutions’ performance was before usually measured only with a wide range of financial indicators. Banks have, however, realized that relying only on financial indicators is inadequate in measuring bank performance pervasively. Although several traditional indicators and performance measures can be used to evaluate the performance of a bank, the research findings and experience of industry experts has shown that these indicators and performance measures can be very misleading.

There are numerous techniques used to measure financial institutions and bank branches operational efficiency, such as ratios (Schweser and Temte, 2002), indices (Coelli et al., 1998), (among these traditional indicators are return on equity (ROE), assets, number of bank branches, profit and loss, return on asset (ROA), and return on investment (ROI)) and regression analyses (Murphy and Orgler, 1982). While effective in many circumstances, traditional techniques and performance measures have a number of inherent limitations making them inappropriate for fully reflecting the increasingly complex nature of banking industry. For example, traditional financial ratio analysis does not allow for impartially combining independent evaluations into a single performance score and it is difficult to use
for comparative purposes. A bank branch might have strong performances for some ratios but show poorly in others making it difficult to judge whether the bank branch is, on average or on some other basis, efficient or not. Simply aggregating these results and performances together can give a misleading indicator of performance or worse, hide under-performing business components within the overall scores. Although, some more complex ratios can take the form of indicator numbers, determining the weights to be used and discovering under-performing activities due to aggregated numbers are just two of the difficulties using Indicators. Another way to measure efficiency is regression analysis, a parametric method that requires a general production model to be specified. Moreover, regression analysis is a central tendency method and is only suitable to model single input-multiple outputs or multiple inputs-single output systems (Paradi et al., 2011).

One of the most effective and widely used performance evaluation methods being Data Envelopment Analysis (DEA) models. DEA is a comprehensive and accepted approach used to evaluate performance in the banking industry; this method is widely accepted and have been applied in many applications (Emrouznejad and Yang, 2018), mainly because of multiple inputs and outputs used in this model and its appropriateness for examining nonlinear relationships in analyses (Chang et al., 2011, Hu et al., 2019).

However, DEA models used to evaluate the performance of banks function as a black box. The performance in some cases such as technical, cost, profitability, and productivity are calculated by assuming that the inputs are consumed to produce outputs (Aly et al., 1990; Berger and Humphrey, 1991; Chan and Liu, 2006; Chou et al., 2002; Ferrier and Lovell, 1990; Grabowski et al., 1993; McAllister and McManus, 1993). Yet, as noted above, the major drawback of these models is that the decision making units (DMUs) remain as black box and they are not checked and looked at in-depth; that is to say that the inputs and outputs of the model are taken into consideration. Consequently, the unit manager cannot identify the source of inefficiency within their unit and only sees whether it is efficient or not (Lewis and Sexton, 2004). Therefore, if the performance of the sub-processes is not taken into account in performance measurement, there is actually a mock estimate of the performance quality. In order to analyze the performance of the sub-processes and the source of inefficiency, some studies have systematically decomposed the overall efficiency into partial efficiencies (Huang et al., 2012). For instance, Färe and Grosskopf (2000) established the relationships between various production processes in the Network DEA (NDEA) model, which allowed for more structures to be added to the model to fit more specific applications.
With a survey of DEA research in the banking industry we found that in all studies conducted in the field of performance evaluation of bank branches, a superficial evaluation has been made. So that the newest researches are limited to the use of network models (to open the black box of performance) that lack the advantages mentioned for the model used in this research. In these models, the weights of each division are pre-determined, and the mediator variables are only evaluated as inputs or outputs. It is not possible to determine reference units and calculate virtual units for inefficient units and the model only determines that the unit under study (bank branch) is efficient or inefficient. Divisional efficiency and overall efficiency are not measurable in an integrated approach. The ability to run the model with both the constant returns to scale (CRS) and the variable returns to scale (VRS) does not exist. Some models only apply for a two-stage process; there was also no possibility of exogenous inputs and outputs for intermediate processes.

Also, the long-term impacts of the banking system on society have not been overlooked, and both efficiency approaches of the banking industry (production and intermediation approach) have not been studied at the same time. Operational efficiency, service effectiveness, social efficiency is not evaluated simultaneously. Thus, using NDEA approach in this research, we intend to eliminate the weaknesses of traditional DEA models in performance evaluation and at the same time, evaluate the efficiency, effectiveness, and, in general, productivity of 37 branches of one of the largest commercial banks of Iran in three divisions (namely, production, intermediation, and social welfare approach). Doing so, we also have the intention of designing a comprehensive and integrated model for measuring productivity in the banking industry.

A lot of studies have been conducted in the field of bank performance evaluation and banking industry using various DEA models. Due to lack of space, in this paper we only outline some of the most important ones, along with the names of their authors, the country or region under study, the number of investigated DMUs, the list of their inputs and outputs, as shown in Table 1.

Table 1. A survey of DEA researches in banking industry

<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>No. of Bank branches</th>
<th>Input variables</th>
<th>Outputs variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Akther et al., 2013)</td>
<td>Bangladesh</td>
<td>21</td>
<td>Stage 1: employees, physical capital, equity, bad loans (the previous period); stage 2: deposits</td>
<td>Stage 1: deposits; stage 2: portfolios of loans, secure investments, bad loans</td>
</tr>
<tr>
<td>(Al-Tamimi and Lootah, 2007)</td>
<td>UAE</td>
<td>15</td>
<td>Model 1: operational cost, interest costs; model 2: interest costs, employees expenses, other operational costs</td>
<td>Model 1: interest income, non-interest income, loans, deposits; model 2: interest income, non-interest income</td>
</tr>
<tr>
<td>(Asmild and Tam, 2007)</td>
<td>7 countries</td>
<td>115</td>
<td>Rewards and wages, interest costs, other costs</td>
<td>Interest income, non-interest income</td>
</tr>
<tr>
<td>(Athanassopoulos and Curram, 1996)</td>
<td>UK</td>
<td>250</td>
<td>ATMs, employees, counter transactions, potential market</td>
<td>Loans sales, liability sales, investments and insurance policies sold</td>
</tr>
<tr>
<td>(Cook and Zhu, 2010)</td>
<td>Canada</td>
<td>100</td>
<td>Sales staff, service staff, other staff</td>
<td>Service, sale</td>
</tr>
<tr>
<td>(Cook et al., 2004)</td>
<td>Canada</td>
<td>1200</td>
<td>FTE, operating expense</td>
<td>Service transactions, sale transactions</td>
</tr>
<tr>
<td>(Fujii et al., 2014)</td>
<td>India</td>
<td>24</td>
<td>Employees, deposits, fixed assets</td>
<td>Acquired assets, customer loans, bad loans</td>
</tr>
<tr>
<td>(Fukuyama and Matousek, 2017)</td>
<td>Japan</td>
<td>-</td>
<td>Stage 1: Labor, capital; Stage 2: deposits</td>
<td>Stage 1: deposits; Stage 2: loans, various nonperforming loans, securities investments</td>
</tr>
<tr>
<td>(Gaganis et al., 2009)</td>
<td>Greece</td>
<td>458</td>
<td>interest costs, non-interest costs, missed loans</td>
<td>Non-interest income, interest income</td>
</tr>
<tr>
<td>(Huang et al., 2018)</td>
<td>U.S.</td>
<td>6182</td>
<td>Labor, physical capital</td>
<td>Total loans, investments, non-interest income; intermediate output: purchased funds</td>
</tr>
<tr>
<td>(Jablonsky et al., 2004)</td>
<td>Czech Republic</td>
<td>81</td>
<td>Employees, operating expense, space</td>
<td>Number of accounts, number of transactions, savings</td>
</tr>
<tr>
<td>(Lin and Chiu, 2013)</td>
<td>Taiwan</td>
<td>30</td>
<td>Production inputs: fixed assets, operating expense, capital; Mediator inputs: deposits, loans</td>
<td>Mediator outputs: deposits, loans; Profitability outputs: non-interest income, interest income, profit</td>
</tr>
<tr>
<td>(Ohsato and Takahashi, 2015)</td>
<td>Japan</td>
<td>105</td>
<td>Division 1: interest on deposits, branches; Division 2: employees, deposits minus interest on deposits; Division 3: &lt;deposits minus interest on deposits&gt; minus &lt;loans and bills discounted minus interest on loans and bills discounted&gt;</td>
<td>Division 1: fees and commissions, deposits minus interest on deposits; Division 2: interest on deposits and bills discounted, &lt;deposits minus interest on deposits&gt; minus &lt;loans and bills discounted minus interest on loans and bills discounted&gt;; Division 3: interest and dividends on securities, gains on sales of bonds</td>
</tr>
<tr>
<td>(Sakar, 2006)</td>
<td>Turkey</td>
<td>11</td>
<td>Branch numbers, employees per branch, assets, loans, deposits</td>
<td>ROA, ROE, interest income/assets, interest income/operating income, non-interest income/assets</td>
</tr>
<tr>
<td>(Wang et al., 2014)</td>
<td>China</td>
<td>16</td>
<td>Employees, fixed assets; Mediator input: deposits</td>
<td>Non-interest income, interest income, non-performing loans; Mediator outputs: deposits</td>
</tr>
<tr>
<td>(Wanke and Barros, 2014)</td>
<td>Brazil</td>
<td>40</td>
<td>Number of branches, employees; mediator inputs: administrative expenses, personnel expenses</td>
<td>Equity, permanent assets; mediator outputs: administrative expenses, personnel expenses</td>
</tr>
<tr>
<td>(Wanke et al., 2016)</td>
<td>Mozambique</td>
<td>117</td>
<td>Total costs, employee costs</td>
<td>Total deposits, income before tax, total credit</td>
</tr>
<tr>
<td>(Wu et al., 2006)</td>
<td>Canada</td>
<td>142</td>
<td>Personnel, other public expenses</td>
<td>Deposits, incomes, loans</td>
</tr>
</tbody>
</table>

2. Literature review
2.1. Data envelopment analysis (DEA)

The original DEA idea was to provide a methodology to identify the cases with the best performance among a set of comparable decision making units (DMUs), and thereby shape the efficient frontier. In addition, this methodology is able to measure the level of performance of non-frontier units. In order to identify the reference set against other inefficient units, some comparisons are made (Azar et al., 2016). DEA is a nonparametric linear programming technique widely used to evaluate the relative efficiency of a set of homogeneous units with similar inputs and outputs (Zerafat Angiz et al., 2012). Mathematically, DEA is a linear programming methodology that calculates the performance of each DMU based on the production possibility set (PPS) that is determined by all DMUs. Its significant advantage is that it doesn’t require the parametric specification, such as production function, to obtain efficiency scores (Siriopoulos and Tziogkidis, 2010). In the DEA model, performance is relative, and the efficiency frontier is created by a convex combination of efficient units. Therefore, each DMU that is located on the efficient frontier is efficient, and otherwise, it will be inefficient. In order to make an inefficient unit efficient, some changes must be made to its inputs and outputs.

2.2. Network Data Envelopment Analysis (NDEA)

In performance evaluation on the basis of production view, the traditional DEA models function as a black box, so that the inputs into these boxes (DMUs) turn into outputs, while the actual conversion process is generally not clearly modeled (Kao, 2014). This suffices for the time when it is only aimed at identifying inefficient decision making units and determining the extent of their inefficiencies. However, the management is sometimes trying to identify the inefficiencies of the internal processes of DMUs. Here we need an introspection of the DMUs and its internal processes so that we can convert each DMU into a network of more partial decision making units. Each of these partial DMUs can use a set of sources produced by other partial DMUs, or generate a set of resources for other partial DMUs. Network DEA (NDEA) models help managers of each DMU to focus more on the strategy of enhancing the efficiency of the unique stages of the production process (Lewis and Sexton, 2004). NDEA models measure the overall efficiency of the organization and the efficiency of each of the organization’s sub-processes. In addition, these models allow the overall efficiency to be broken down using mathematical relationships between organizational efficiency and sub-processes efficiency. In the NDEA models, the network structure is used instead of the hierarchical structure of activities (Hsieh and Lin, 2010). Network DEA is
related to the use of DEA technique in order to measure the relative efficiency of a system with respect to its internal structure (Wanke et al., 2019). In the NDEA, more meaningful and informative results are obtained compared to the common black box approach (Kao, 2014). Another issue that needs to be addressed is that ignoring the operation of the process components may have misleading results; there are numerous examples in the literature that show the entire system can be efficient, even though all process components are not (Kao and Hwang, 2008). There are also cases where all the process components of a DMU have a worse performance than those of another DMU, while the overall performance of the system is better (Kao and Hwang, 2010). These results indicate that the network DEA model is required to generate more accurate and reliable results, especially when the problem has a network structure. Unlike the conventional DEA models, the network DEA models do not have a single standard format; rather, their format depends on the network under study (Kao, 2009).

3. Methodology
Given that there are a large number of DEA models in the literature and that researchers have used different DEA models for the performance evaluation of bank branches, we inevitably deal with deciding about choosing a specific DEA version and the dilemma of choosing proper inputs and outputs. Although there is no comprehensive agreement on how to make these decisions, a systematic approach is proposed for decision making in each case, which can be very useful.

3.1. Selecting variables (input and output selection)
Studies have shown that the results of performance assessment largely depend on the selection of inputs and outputs (Berg et al., 1991; Tortosa-Ausina, 2002). In recent years, a lot of research has been done to measure the efficiency in the banking industry. One of the most important components and stages of these studies is making a decision on the definition of proper inputs and outputs in the banking industry. This is not a simple issue and has remained a widespread and unresolved disagreement in the literature. The selection of inputs and outputs is different for assessing the performance of the banking industry since it depends on the research objectives. However, except for deposits, there is generally an agreement on the main categories of inputs and outputs of banks' performance assessment. For instance, conventional inputs include physical capital and employee, and conventional outputs include total loans and other incomes or investments. Nevertheless, regarding the role of deposits,
there is a controversy in the research literature that deposits should be considered as inputs or outputs in the banking system (Deposits have some input characteristics since they are the raw materials for investment and are used to pay part of the interest. Deposits also have some output specifications since they are related to the amount of liquidity, security, and services provided to the depositors); this by itself depends on whether the banking system is considered as a production process or as an intermediation process. In the studies regarding the efficiency of the banking industry, there are two main approaches (production and intermediation approach) to explain and define inputs and outputs that depict different views of banking activities. In the production approach, banks are regarded as firms that use human and other physical resources as inputs to generate different types of deposit accounts and other outputs (such as number of transactions). In this approach, banks provide customers with financial transactions, maintenance of customer deposits, customers' financial documents processing, customer credit reports, cash checks, and other financial assets management. The production approach emphasizes operational activities, and thus, banks are primarily considered as service providers to customers (account holders). The set of inputs of this approach only includes physical inputs (such as employees, materials, space, or equipment) or their related costs, since only physical inputs are required for conducting transactions, processing financial documents, or providing customers with consulting services. Given the fact that only the operational process is important in the production approach, there is no interest expense in this approach. The outputs of this approach represent the services provided to customers; thus, it is better to evaluate the processed documents or specific services provided over a specific time period using the number and types of transactions. In addition, considering the fact that the details on some of the transactions are not usually available, the deposit is regarded as the indicator of the level of the service provided by bank branch, one of the most important outputs of this stage (Camanho and Dyson, 2005). In intermediation approach, banks are considered as the financial service intermediaries (i.e. the resource mediators between savers and investors). In this approach, the bank takes deposits from customers and converts them into loans and gives them to the applicants. The mediating role of the bank branches is mainly used to examine how efficient a branch is in systematically collecting deposits and other funds from customers (inputs) and lending the money in various forms of loans or profit-earning processes. In this approach, financial resources (deposits) and expenses related to the interests on deposits are the inputs, since deposits are the most important raw materials that are converted in a financial intermediary process and loans and other revenue generating activities are outputs (Mester, 1997; Berger and Humphrey, 1997).
In the intermediation approach, banks have two important functions, namely moving and distributing resources efficiently in order to finance investment activities in the economy.

Bank deposits are among the important sources of a bank to provide various facilities. As stated, researchers have proposed different opinions about the definition of deposit. Deposits were previously considered as output, and recently they are considered as input (Paradi and Zhu, 2013). For example, Berg et al. (1991) considers deposit as output since it shows a resources consumer activity. However, Berger (1993) has a different point of view and believes that deposits are input in the models that pay attention to the interests paid on the purchased fund resources. Such problems make it impossible to compare the results of various studies.

Berger and Humphrey (1997) argued that none of these two approaches alone is complete and cannot reflect the secondary roles of banks as (1) providers of processing services for accounts and transactions and (2) financial intermediaries that transfer the funds (deposits) from savers to investors. In addition, deposits are, on the one hand, the raw materials for banks to invest. On the other hand, these raw materials provide liquidity, security, and payment services to depositors. Nevertheless, each of these approaches has its advantages. Using these approaches jointly reflects a comprehensive image of banking industry including business activity (e.g. customer interaction) and intermediation activity (associated with financial performance). Therefore, in order to conduct a comprehensive assessment, we are going to evaluate the performance of bank branches under both the production and intermediation approaches; taking advantage of a new concept, we are also going to consider deposits in evaluating the efficiency of bank branches in an NDEA structure. In the network DEA model, deposits are considered as intermediate product, meaning that they are initially considered as outputs of previous sub-process and then as inputs of the next sub-process. Considering the extensive review of the literature as well as involving the stakeholders of the results of banking activities, asking some experts of the field and taking into account the production and intermediation approaches, the research variables were identified as follows. Then, the operational process of Iran's banking system is divided into three sub-processes, respectively consisting of production process (the process of the production of deposit, account, and transaction using employees, fixed assets, and non-operating costs), intermediation process (the process of earning through lendable deposits) in which the intermediating input/output of the deposits connects these two processes, and social welfare
process (job creation through granting facilities to the industry and mining, service, and agriculture division).

Finally, the identified variables to be used for evaluating the performance of the banking industry in this study are as follows: (the inputs of the initial sub-process), employees ($X_1$), fixed assets ($X_2$), and non-operating costs ($X_3$) which included all bank costs other than deposit interest expense. In addition, the outcomes of the banking industry (the output of the second sub-process) are as follows: bank facilities ($W_1$), interest income ($W_2$), and non-interest income ($W_3$). The deposit interest expenses ($X_4$) is considered as the input of the intermediation process (second sub-process), the bank deposits ($Z_1$) is considered as the output of the banking system (the outputs of the first sub-process), and the input of the intermediation process is regarded as linking the first and second sub-process. Also, the number of transactions ($Z_2$) and the number of accounts ($Z_3$) are considered as other outputs. Ultimately, the job creation rate (employment) ($Y_1$) is identified as the output of the third sub-process (social welfare) and the impact of the banking industry in Iran (long-term development that is the logical result of achieving a set of outputs and outcomes) and bank facilities (the output of the intermediation process) is recognized as the input of the social welfare process (the mediating variable between sub-processes 2 and 3).

Social welfare is an expression referring to the economic, social, and political situation that aims at preserving human dignity and the accountability of people in society towards each other and improving their abilities. Indicators that are considered for social welfare in Iran include employment, education, health, treatment, nutrition, housing, and per capita income. Employment is selected in the social welfare division as an effect of banking activities due to the fact that the associated information is available, credible and observable, and is more comprehensive compared to other social welfare indicators since high level of employment increases in the individuals' income and, as a result, enhances the national income and, ultimately, the per capita income (per capita income is also an important indicator of social welfare). In addition, the overall result that can be inferred from these studies is that in order to enhance the overall satisfaction of society, the best way is to fight poverty, and as noted, employment contributes to increasing people's incomes and, as a result, reducing poverty. Thus, it is more comprehensive than the other social welfare indicators.

The employment index is obtained on the basis of the banks' granted facilities from the employment ratio to the value of the products of various divisions of economy as an indicator
for job creation. The desired divisions in this study are the industry and mining, service and agriculture Divisions; job creation indices for each Division are calculated as follows:

\[ l_{it} = \frac{TL_{it}}{Q_{it}} \quad ; \quad t = 1, 2, 3, K , T \]

Where \( l_{it} \) represents the employment index for the \( i^{th} \) division in the year \( t \). Also, \( TL_{it} \) is the total employment created in \( i^{th} \) division and \( Q_{it} \) is the value of the products (value added) of \( i^{th} \) division at time \( t \).

By multiplying the calculated index in the banks’ granted facilities to each division, the total employment generated from facilities granted by banks to the \( i^{th} \) division is obtained. This index is calculated as follows:

\[ L_{it} = CR_{it} \times l_{it} \]

Where \( L_{it} \) represents the employment generated by banks’ granted facilities to the \( i^{th} \) division at time \( t \), \( CR_{it} \) is banks’ granted facilities to the \( i^{th} \) division at time \( t \), and \( l_{it} \) is the employment index for division \( i \) at time \( t \). \( L_{t} \) is the total employment in all divisions, that is obtained by the total employment rate generated by the banking facilities in all the three divisions and is calculated as follows:

\[ L_{t} = \sum_{i=1}^{3} L_{it} = \sum_{i=1}^{3} (CR_{it} \times l_{it}) \]

Finally, the network model of performance evaluation of banking industry of Iran is developed in accordance with Fig. 1.

**Fig. 1.** Banking three-stage SBM model
The model is applied to simultaneously estimate operational efficiency, service effectiveness, social effectiveness and productivity for 37 branches of one of the largest commercial banks in Iran.

Our proposed evaluation framework for Iranian banking operational performance can be separated into three main dimensions: production, intermediation, and social welfare, as depicted in Fig. 1. which shows a comprehensive evaluation of banking performance, allowing managers to identify major sources of and opportunities to improve it.

Finally, for the purpose of including the most appropriate and significant items of the banking industry and considering the most commonly used variables for efficiency evaluation in the literature, this study regards the inputs of the banking industry (the inputs of the first sub-process: production) as (i) employees (X1), which refers to the number of full-time employees hired, (ii) fixed assets (X2), which refer to the asset value of physical capital, and (iii) non-operating costs (X3), which refer to the all bank costs other than deposit interest expense. The outputs of the banking system (the outputs of the first sub-process) are as follows: (i) number of transactions (Z2); (ii) number of accounts (Z3); and (iii) bank deposits (Z1), which include current deposits and time deposits, are treated as intermediate inputs/outputs (the outputs of the first sub-process and the input of the second sub-process) in this study. The inputs of the second sub-process (Intermediation) are as follows: (i) interest expenses (X4), which refers to fees paid for the profits of deposits, and (ii) bank deposits (Z1). The outputs of the second sub-process are as follows: (i) interest incomes (W2), which refers to incomes that are primarily derived from loans; (ii) non-interest incomes (W3), which includes fees, commissions, investment and other business income; and (iii) bank facilities (W1), which refers to the credit line(s) such as overdrafts, loans, import and export lines, etc. offered by the bank to the customer, are treated as intermediate inputs/outputs (the outputs of the second sub-process and the input of the third sub-process) in this study. The input of the third sub-process (social welfare) is (i) bank facilities (W1), and the output is (i) employment (Y1), which refers to the number of jobs created by the banks' granted facilities to various economic sectors. The three-stage SBM structure for banking efficiency evaluation and the inputs, outputs (includes outputs, outcomes and impact), and intermediate measure selection are shown in Fig. 1.

3.2. The orientation of DEA model

Selecting the type of model's orientation depends on the extent of control a decision maker can exercise over the inputs or outputs. In this research, based on the model used to evaluate
the banking industry in Iran, a non-oriented DEA model is used to evaluate the performance of this industry since the mediating variables of the model are both input and output.

3.3. The type of return to the scale of the DEA model

The constant returns to scale (CRS) assumption is valid only when all DMUs are operating at an optimal scale, and the application of variable returns to scale (VRS) setting will be more appropriate in banking efficiency evaluation (Wang et al., 2014). As full competition in Iranian banks is unlikely, the assumption that all the investigated banks operate at the optimal scale does not seem logical. However, in this research, CRS and VRS assumptions are used. The analysis of the efficiency of bank branches in the CRS can be considered as a long-term situation of the banks and VRS can be regarded as a short-term condition of them.

3.4. The applied NDEA model

Network DEA models that stand on the radial measure of efficiency, e.g. the CCR (Charnes et al., 1978) or the BCC (Banker et al., 1984) models as the basic DEA methodology and production possibility set, should observe the assumption of the appropriate changes to the reduction of inputs and increase of outputs. The slacks-based measure (SBM) is a non-radial method and is suitable for measuring efficiencies when inputs and outputs may change in a non-proportional way (Tone and Tsutsui, 2009). The SBM model deals with directly with input surpluses and output shortfalls. Having some features such as invariability with regard to the data unit and the uniform increase in each of the input and output slack variables, the SBM model compensates the shortcomings of other models. The main advantage of the SBM model over the other models is that it able to appropriately measure the efficiency of weakly efficient DMUs (Arabi et al., 2015).

Let \( n \) DMUs \((1, K, n)\) consisting of \( K \) process \((1, K, K)\), \( m_k \) represent the number of inputs of process \( k \), and \( r_k \) be the number of outputs of process \( k \). Also, we denote the link leading from process \( k \) to process \( h \) by \((k, h)\). \( X_{jk}^{(k)} \) is the input resources to \( DMU_j \) in process \( k \), \( Y_{jk}^{(k)} \) is the output products of \( DMU_j \) at process \( k \), and \( Z_{jk}^{(h,k)} \) is the linking intermediate products from process \( k \) to process \( h \) (the output of process \( k \) and the input of process \( h \)). Also, \( s_i^{(k)} \) and \( s_o^{(k)} \) are the slack variables corresponding to the inputs and outputs of process \( k \), respectively. \( s_i^{(k,h)} \) and \( s_o^{(k,h)} \) are the slack variables corresponding to communication links as inputs and outputs. If we optimize the number of communication links from process \( k \) to process \( h \) as the inputs of process \( h \), we will have \( t_{hk}^{(k,h)} \) and if we optimize the number of
communication links from process \( k \) to process \( h \) as the output of process \( k \), we will have \( t_{k,h}^{(k,h)} \). Then, the network SBM model used in the research is presented as follows (See also Zarei Mahmoudabadi et al., 2018 and Li et al., 2018):

\[
\begin{align*}
\theta_0^{\text{Global}} &= \text{Min} \left[ kt - \sum_{k=1}^{K} \left( \frac{1}{m_k} + \sum_{h=2}^{K} t_{k,h}^{(k,h)} \right) \left( \sum_{l=1}^{n_k} S_{l}^{(k,l)} + \sum_{l=1}^{K} S_{l}^{(k,h)} \right) \right],
\end{align*}
\]

s.t. \( kt + \sum_{k=1}^{K} \left( \frac{1}{r_k} + \sum_{k=1}^{K-1} t_{k,h}^{(k,h)} \right) \left( \sum_{j=1}^{n_k} S_{j}^{(k,l)} + \sum_{j=1}^{K} S_{j}^{(k,h)} \right) = 1, \)

\[
\begin{align*}
\sum_{j=1}^{n_k} \lambda_j^{(k)} X_{j}^{(k)} + S_{i}^{(-k)} &= tX_{i0}^{(k)}, \quad (k = 1, K, k), \\
\sum_{j=1}^{n_k} \lambda_j^{(k)} Y_{j}^{(k)} - S_{i}^{(k)} &= tY_{i0}^{(k)}, \quad (k = 1, K, k), \\
\sum_{j=1}^{n_k} \lambda_j^{(h)} Z_{j}^{(h)} &= \sum_{j=1}^{n_k} \lambda_j^{(k)} Z_{j}^{(k,h)}, \quad (\forall (k, h)), \\
\sum_{j=1}^{n_k} \lambda_j^{(h)} Z_{j}^{(h)} + S_{h}^{(-k)} &= tZ_{h0}^{(k,h)} + S_{h}^{(k,h)}, \quad (\forall (k, h)), \\
\sum_{j=1}^{n_k} \lambda_j^{(k)} Z_{j}^{(k,h)} - S_{k}^{(-k)} &= tZ_{k0}^{(k,h)} - S_{k}^{(k,h)}, \quad (\forall (k, h)), \\
S_{i}^{(k,h)} &\leq Md^{(k,h)}, \quad (\forall (k, h)), \\
S_{h}^{(-k)} &\leq M (1 - d^{(k,h)}), \quad (\forall (k, h)), \\
\lambda_j^{(k)}, S_{i}^{(-k)}, S_{i}^{(k)}, S_{h}^{(-k)}, S_{h}^{(k,h)} &\geq 0, \quad d^{(k,h)} \in \{0, 1\}, (\forall j, i, r, k).
\end{align*}
\]

Then weight of each process is defined as (this indicates the contribution of each sub-process to the overall inefficiency of the system):

\[
\begin{align*}
w_k = \frac{1 + \frac{1}{r_k} + \sum_{k=1}^{K-1} t_{k,h}^{(k,h)} \left( \sum_{j=1}^{n_k} S_{j}^{(k,l)} + \sum_{j=1}^{K} S_{j}^{(k,h)} \right)}{\left( \sum_{k=1}^{K} \left( \frac{1}{r_k} + \sum_{k=1}^{K-1} t_{k,h}^{(k,h)} \left( \sum_{j=1}^{n_k} S_{j}^{(k,l)} + \sum_{j=1}^{K} S_{j}^{(k,h)} \right) \right) \right)}, \quad \text{s.t. } 1 + \frac{1}{r_k} + \sum_{k=1}^{K-1} t_{k,h}^{(k,h)} \left( \sum_{j=1}^{n_k} S_{j}^{(k,l)} + \sum_{j=1}^{K} S_{j}^{(k,h)} \right). 
\end{align*}
\]

Among the advantages of this network model is that the weight of each of the components or sub-processes is obtained by the particular model and conditions of the unit under consideration and without any external interference that was found in the research literature as a negotiable and unresolved issue. In addition, the linking activity or intermediate variables are simultaneously examined as both inputs and outputs. Again, without any external interference and through the model and the best existing conditions for the unit under consideration, it is determined that the mediating variable operates as input and thus, its
reduction is desirable, or it is specified that the mediating variable functions as output and its increase is desirable for the unit under study.

If $S_{k}^{+} = 0$ in the above model, the constraints of the linking activity as outputs are redundant, meaning that the linking activity is considered as input; and if $S_{h}^{-} = 0$, the constraints of the linking activity as input are redundant, meaning that the linking activity is considered as output. In addition, if VRS is the case, based on the variable change, convex constraint $\sum_{j=1}^{n} \lambda_{j}^{(k)} = 1, (\forall k)$ is added to the set of the above constraints.

After calculating the overall efficiency of the system, the efficiency of each of the sub-processes is calculated based on the following procedure. Regarding the first sub-process, based on the previous prioritization, i.e. maintaining the overall efficiency of the system, the current model determines its efficiency $\theta_{0}$ as follows:

$$
\theta_{0} = \text{Min} \left[ t - \frac{1}{m_{1}} \sum_{j=1}^{m_{1}} S_{j}^{(1,0)} \right]
$$

s.t. $t + \frac{1}{r_{1} + \sum_{k=1}^{K} t_{r_{1}+1}^{(k, h)}} \left( \sum_{r=1}^{s} Y_{r}^{(k, h)} + \sum_{k=1}^{K} Z_{0}^{(k, h)} \right) = 1,$

$$
\sum_{j} \lambda_{j}^{(k)} X_{ij}^{(k)} + S_{j}^{(k)} = tX_{i0}^{(k)}, \quad (k = 1, K, k),
$$

$$
\sum_{j} \lambda_{j}^{(k)} Y_{ij}^{(k)} - S_{j}^{(k)} = tY_{i0}^{(k)}, \quad (k = 1, K, k),
$$

$$
\sum_{j} \lambda_{j}^{(k)} Z_{j}^{(k, h)} = \sum_{j} \lambda_{j}^{(k)} Z_{j}^{(k, h)}, \quad (\forall (k, h)),
$$

$$
\sum_{j} \lambda_{j}^{(k)} Z_{j}^{(k, h)} + S_{h}^{(k, h)} = tZ_{0}^{(k, h)} + S_{k}^{(k, h)}, \quad (\forall (k, h)),
$$

$$
\sum_{j} \lambda_{j}^{(k)} Z_{j}^{(k, h)} - S_{h}^{(k, h)} = tZ_{0}^{(k, h)} - S_{h}^{- (k, h)}, \quad (\forall (k, h)),
$$

$$
S_{k}^{(k, h)} \leq Md^{(k, h)}, \quad (\forall (k, h)),
$$

$$
S_{h}^{(k, h)} \leq M (1 - d^{(k, h)}), \quad (\forall (k, h)),
$$

$$
\left[ \begin{array}{c}
kt - \frac{1}{m_{1} + \sum_{k=1}^{K} t_{r_{1}+1}^{(k, h)}} \left( \sum_{r=1}^{s} Y_{r}^{(k, h)} + \sum_{k=1}^{K} Z_{0}^{(k, h)} \right) \\
\end{array} \right] -
$$

$$
\theta_{0}^{\text{global}} \left( (k - 1)r + \sum_{k=1}^{K} \frac{1}{r_{1} + \sum_{k=2}^{K} t_{r_{1}+1}^{(k, h)}} \left( \sum_{r=1}^{s} Y_{r}^{(k, h)} + \sum_{k=1}^{K} Z_{0}^{(k, h)} \right) \right) = \theta_{0}^{\text{global}},
$$

$$
\lambda_{j}^{(k)} , S_{j}^{(k, k)} , S_{j}^{(- (k, h))} , S_{h}^{(k, h)} , S_{k}^{(- (k, h))} \geq 0, \quad d^{(k, h)} \in \{0,1\}, \quad (\forall j, i, r, k).
$$

Using the same method, the efficiency of the rest of the sub-processes can also be calculated. If $\left( \theta^{*}, \tau^{*}, \lambda_{j}^{(k)} , S_{j}^{(- (k, h))} , S_{j}^{(k, h)} , S_{h}^{(k, h)} , S_{k}^{(- (k, h))} \right)$ is the optimal solution obtained by solving (1), the optimal solution for the proposed Network SBM (NSBM) model is defined as follows:
Based on the proposed NSBM model, a DMU is efficient when it operates efficiently in all its sub-processes, that is $\theta_0^* = 1$, and consequently, $E_0^* = 1$. This condition is equivalent to the fact that all the surplus and shortfall variables are zero in the optimal solution; that is to say that there is no surplus input and shortfall output in the optimal solution. In the end, for each inefficient DMU, the efficient target (projected on the frontier) can be calculated from the following equations:

\[
E^* = \theta^*; \quad A_j^{(k)^*} = \frac{Z_j^{(k)^*}}{t}; \quad S_j^{(k)^*} = \frac{S_j^{(k)^*}}{t}; \quad S_i^{(k)^*} = \frac{S_i^{(k)^*}}{t}; \quad S_h^{(k,h)^*} = \frac{S_h^{(k,h)^*}}{t}; \quad S_k^{(k,h)^*} = \frac{S_k^{(k,h)^*}}{t}.
\]

(4)

4. Results and discussion

One of the important issues in the world of competition is the extraction and analysis of the market share of industry activists. Market share plays an important role in analyzing competitive position and examining the success factors. In the banking system of Iran, many examples can be considered to examine the market share of banks. Among them, the market share of banks in attracting resources is one of the most important and most effective indicators. Accordingly, the market share of the bank under review in attracting deposits (resources) from 2007 to 2014 is 10.11, 10.26, 10.89, 11.91, 12.02, 11.06, 9.16 and 9.57, respectively, which indicates the relative importance of this bank compared to other banks in Iran.

Also, some of the awards and honors of the bank under review that indicate the importance of this bank is as follows:

- Ranking top position in Iran's commercial banks;
- Earn green banking innovation award;
- Earn Paris international quality golden star award;
- The best bank of Iran in different years based on the ranking of international institutions.

Table 2a and 2b summarizes the data for inputs, outputs, and the intermediate (link) variables of 37 branches of one of Iran's largest commercial banks.

<p>| Table 2a. Selection of input and output variables in the current study |
|---------------------------------|-------------------|-------------------|</p>
<table>
<thead>
<tr>
<th>Country</th>
<th>No. of Bank branches</th>
<th>Input variables</th>
<th>Outputs variables</th>
</tr>
</thead>
</table>
4.1. The results of applying network SBM approach

Rapidly changing banking services and other environmental drivers set continuously new demands for bank’s performance measurement systems. As a result data gained from traditional measures are not adequate anymore (Harker and Zenios, 2000). Intensive competition in the banking industry has been one of the main drivers of change towards the more holistic performance measurement and management. It has made it necessary for financial institutions and banks to use new and comprehensive approach in performance evaluation, and thus make it possible for the bank to achieve its strategic and operative goals by providing adequate information. Due to the nature of the business and focus on creating money in banking industry, financial institutions’ performance was before usually measured only with a wide range of financial indicators. Banks have however realized that relying only on financial indicators is inadequate in measuring bank performance pervasively. Although
several traditional indicators and performance measures can be used to evaluate the performance of a bank, the research findings and experience of industry experts has shown that these indicators and performance measures can be very misleading. Among these traditional indicators are return on equity (ROE), debts, assets, number of bank branches, profit and loss, return on asset (ROA), and return on investment (ROI).

There are numerous techniques used to measure financial institutions and bank branches operational efficiency, such as ratios (Schweser and Temte, 2002), indices (Coelli et al., 1998), and regression analyses (Murphy and Orgler, 1982). While effective in many circumstances, traditional techniques and performance measures have a number of inherent limitations making them inappropriate for fully reflecting the increasingly complex nature of banking industry. For example, traditional financial ratio analysis does not allow for impartially combining independent evaluations into a single performance score and it is difficult to use for comparative purposes. A bank branch might have strong performances for some ratios but show poorly in others making it difficult to judge whether the bank branch is, on average or on some other basis, efficient or not. Simply aggregating these results and performances together can give a misleading indicator of performance or worse, hide under-performing business components within the overall numbers. Although, some more complex ratios can take the form of indicator numbers, determining the weights to be used and discovering under-performing activities due to aggregated numbers are just two of the difficulties using Indicators. Another way to measure efficiency is regression analysis, a parametric method that requires a general production model to be specified. Moreover, regression analysis is a central tendency method and is only suitable to model single input-multiple outputs or multiple inputs-single output systems (Paradi et al., 2011). However, the traditional measures ignore the dynamic links between variables in the bank. In this division, we return to the NDEA model and consider the internal communication of the black box model. To calculate the overall efficiency of bank branches in the network system (Fig. 1), the proposed model is used under the CRS and VRS assumptions. Also, divisional efficiency is obtained through model 3. Unlike the separation approach in which the weight of each division (sub-process) is externally imposed on the model and is the same for all DMUs, here in the network model used in this study, the weight of each division is determined using (2) by the optimal solution of the model and based on the special conditions of the unit under study. Table 3 shows the overall efficiency of the network system and the efficiency of each division along with the weight of each division under the CRS assumption (the numbers in
parentheses indicate the rank of the branches in that division). As already stated, the analysis of the efficiency of bank branches in the CRS assumption can be considered as the long-term condition of branches (it should be mentioned that given the importance of CRS and to avoid prolonged explanation, the whole analyses in this section are based on the CRS assumption, and the VRS assumption analyses are therefore avoided, though most of these analyses are also valid for the VRS assumption). Also, Table 3 shows the reference set with their shadow price values (numbers within parentheses) under CRS assumption for each division; for instance, A1 represents DMU_{AF} in division 1. This means that in the optimal solution we have $\lambda_{AF}^{(1)} f 0$. According to the results reported in Table5, the virtual unit for evaluating the efficiency of branch A in division 1 is the combination of 1.031 of branch AF, 0.379 of branch AI, and 2.566 of branch AJ. Division 2 is made up of the combination of 1.629 of branch C and 1.357 of branch Q, and division 3 is constructed of 7. 232 of branch AB (The way virtual unit is constructed is explained in section 3.4.).

Table 3. Results of NSBM (CRS assumption)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.3127 (13)</td>
<td>0.4059 (36)</td>
<td>0.5586 (5)</td>
<td>0.1970 (14)</td>
<td>0.0609</td>
<td>AF1(1.031), AI1(0.379), AJ1(2.566)</td>
<td>C2(1.629), Q2(1.357)</td>
<td>AB3(7.232)</td>
</tr>
<tr>
<td>B</td>
<td>0.2376 (20)</td>
<td>0.5762 (26)</td>
<td>0.2423 (12)</td>
<td>0.1661 (18)</td>
<td>0.5299</td>
<td>V1(0.099), AF1(0.759), AI1(0.189), AJ1(0.335)</td>
<td>C2(0.279), Q2(0.712)</td>
<td>AB3(3.287)</td>
</tr>
<tr>
<td>C</td>
<td>0.8241 (1)</td>
<td>1.0000 (1)</td>
<td>1.0000 (1)</td>
<td>0.6096 (1)</td>
<td>0.4506</td>
<td>C1(1.000)</td>
<td>C2(1.000)</td>
<td>AB3(0.877)</td>
</tr>
<tr>
<td>D</td>
<td>0.3676 (8)</td>
<td>0.6243 (23)</td>
<td>0.3866 (7)</td>
<td>0.2632 (6)</td>
<td>0.4907</td>
<td>X1(0.338), AI1(0.306)</td>
<td>C2(0.049), Q2(0.612)</td>
<td>AB3(2.659)</td>
</tr>
<tr>
<td>E</td>
<td>0.0743 (37)</td>
<td>0.4287 (35)</td>
<td>0.0569 (37)</td>
<td>0.0651 (35)</td>
<td>0.4944</td>
<td>AF1(0.259), AI1(0.484), AJ1(0.149)</td>
<td>Q2(0.985)</td>
<td>AB3(4.210)</td>
</tr>
<tr>
<td>F</td>
<td>0.1775 (27)</td>
<td>0.5941 (25)</td>
<td>0.1693 (28)</td>
<td>0.1292 (25)</td>
<td>0.5516</td>
<td>C1(1.200), V1(0.857), AI1(0.014)</td>
<td>C2(0.235), Q2(0.381), AH2(0.163)</td>
<td>AB3(2.824)</td>
</tr>
<tr>
<td>G</td>
<td>0.3265 (10)</td>
<td>1.0000 (1)</td>
<td>0.2330 (23)</td>
<td>0.2540 (7)</td>
<td>0.4331</td>
<td>G1(1.000)</td>
<td>C2(0.538), AH2(0.108)</td>
<td>AB3(1.125)</td>
</tr>
<tr>
<td>H</td>
<td>0.1674 (26)</td>
<td>0.9260 (19)</td>
<td>0.1735 (31)</td>
<td>0.1045 (28)</td>
<td>0.5339</td>
<td>V1(0.316), AC1(0.185), AF1(0.085), AI1(0.207), AJ1(0.440)</td>
<td>Q2(0.618), AH2(0.313)</td>
<td>AB3(4.543)</td>
</tr>
<tr>
<td>I</td>
<td>0.4791 (4)</td>
<td>0.3903 (37)</td>
<td>0.6614 (3)</td>
<td>0.4334 (2)</td>
<td>0.4602</td>
<td>K1(0.574), AJ1(1.451)</td>
<td>C2(0.247), Q2(0.770)</td>
<td>AB3(3.508)</td>
</tr>
<tr>
<td>J</td>
<td>0.1827 (26)</td>
<td>0.5182 (29)</td>
<td>0.1805 (27)</td>
<td>0.1315 (24)</td>
<td>0.5302</td>
<td>AF1(0.525), AI1(0.320), AJ1(0.220)</td>
<td>C2(0.227), Q2(0.775)</td>
<td>AB3(4.425)</td>
</tr>
<tr>
<td>K</td>
<td>0.4322 (5)</td>
<td>1.0000 (1)</td>
<td>0.5206 (6)</td>
<td>0.2487 (8)</td>
<td>0.5922</td>
<td>K1(1.000)</td>
<td>C2(0.213), Q2(0.394), AH2(0.150)</td>
<td>AB3(2.780)</td>
</tr>
<tr>
<td>L</td>
<td>0.2652 (17)</td>
<td>1.0000 (1)</td>
<td>0.2723 (16)</td>
<td>0.1506 (21)</td>
<td>0.5870</td>
<td>L1(1.000)</td>
<td>C2(0.180), Q2(0.043), AH2(0.132)</td>
<td>AB3(1.145)</td>
</tr>
<tr>
<td>M</td>
<td>0.4894 (2)</td>
<td>1.0000 (1)</td>
<td>0.5681 (4)</td>
<td>0.2967 (5)</td>
<td>0.5497</td>
<td>M1(1.000)</td>
<td>C2(0.694), Q2(0.043), AH2(0.020)</td>
<td>AB3(0.926)</td>
</tr>
<tr>
<td>N</td>
<td>0.1016 (34)</td>
<td>0.0339</td>
<td>0.0644 (16)</td>
<td>0.0775 (33)</td>
<td>0.4373</td>
<td>N1(1.000)</td>
<td>C2(0.270), Q2(0.400), AH2(0.264)</td>
<td>AB3(3.550)</td>
</tr>
<tr>
<td>O</td>
<td>0.2022 (23)</td>
<td>0.8091 (20)</td>
<td>0.3802 (8)</td>
<td>0.9496 (30)</td>
<td>0.7390</td>
<td>K1(0.489), AF1(0.170), AI1(0.083), AJ1(0.135)</td>
<td>C2(0.801), Q2(0.191)</td>
<td>AB3(1.520)</td>
</tr>
<tr>
<td>P</td>
<td>0.2972 (15)</td>
<td>0.5656 (27)</td>
<td>0.3159 (12)</td>
<td>0.2212 (12)</td>
<td>0.5169</td>
<td>K1(0.375), AI1(0.118), AJ1(0.296)</td>
<td>Q2(0.523)</td>
<td>AB3(2.236)</td>
</tr>
<tr>
<td>Q</td>
<td>0.4812 (3)</td>
<td>1.0000 (1)</td>
<td>1.0000 (1)</td>
<td>1.2362 (10)</td>
<td>0.6792</td>
<td>Q1(1.000)</td>
<td>Q2(1.000)</td>
<td>AB3(4.276)</td>
</tr>
<tr>
<td>R</td>
<td>0.0974 (35)</td>
<td>0.5263 (28)</td>
<td>0.1432 (30)</td>
<td>0.0541 (36)</td>
<td>0.6951</td>
<td>C1(0.276), AI1(0.543), AJ1(0.431)</td>
<td>C2(0.476), Q2(0.882), AH2(0.121)</td>
<td>AB3(5.473)</td>
</tr>
<tr>
<td>S</td>
<td>0.1210 (33)</td>
<td>0.4464 (32)</td>
<td>0.1256 (32)</td>
<td>0.0842 (32)</td>
<td>0.5636</td>
<td>V1(0.590), AF1(0.227), AI1(0.056), AJ1(0.711)</td>
<td>C2(0.530), Q2(0.441)</td>
<td>AB3(2.350)</td>
</tr>
<tr>
<td>T</td>
<td>0.0839 (33)</td>
<td>0.6502 (22)</td>
<td>0.0651 (35)</td>
<td>0.0331 (37)</td>
<td>0.4791</td>
<td>C1(0.099), K1(0.326), V1(0.129), Q2(0.647), AH2(0.032)</td>
<td>AB3(2.960)</td>
<td></td>
</tr>
</tbody>
</table>
Based on the results reported in Table 3, the weighted average of the efficiency of the division is equal to the overall efficiency. Furthermore, one of the notable results obtained here is that the weighted average of the ranks of divisions is a more appropriate representative for the overall ranking, compared to the simple arithmetic average of the ranking of divisions. It is in such a way that the rank correlation between the weighted average of the ranks of divisions and the rank of the overall efficiency is 99%, but the rank correlation between the simple arithmetic average of the rank of the divisions and the rank of the overall efficiency is 90%.

Fig. 2 demonstrates the comparison of the efficiency scores of the separation and network approaches under CRS assumption. As is clear, the efficiency scores in the network approach are, in all cases, smaller than the efficiency scores in the separation approach, which indicates the high discriminating power of the network DEA model.
Fig. 3 shows the comparison of the scores of the overall efficiency of network system and the efficiency of each division under CRS assumption. Based on this graph, the efficiency scores of sub-process 1, or in other words the operational efficiency of all the branches (except for Branch A and I) are higher than that of sub-process 2 and 3, indicating that almost all branches function well in the operational efficiency (the average operational efficiency of bank branches is 77%). Also, the efficiency of sub-process 2, or, in other words, the service effectiveness of almost all branches (except for branches A, C, I, and Q) has been lower than their operational efficiency. Therefore, the branches of the bank under study are worse in terms of the service effectiveness than operational efficiency and should focus more on this field (the average service effectiveness of bank branches is 30%). A worthwhile consideration is the inappropriate situation of almost all bank branches in the field of social effectiveness (efficiency of sub-process 3), so that the social effectiveness of most bank branches is at the lowest position (the average social effectiveness of the bank branches is 18%); this indicates that the bank branches have not operated very well in the area of social welfare, and that the policy-makers in this area should give more serious attention to this issue. The weights assigned by the model to each sub-process also confirm the above-mentioned explanations, so that the total weights assigned to the bank branches in sub-process 1, 2, and 3 are 4.0309, 13.1186, and 19.8505, respectively. This indicates the contribution of each sub-process to the overall inefficiency of the system (the total weight in all sub-processes and bank branches is equal to 37). Therefore, the contribution of sub-process 1, 2, and 3 to the overall inefficiency of the system is 10.89%, 35.46%, and 53.65%, respectively.
In terms of the overall efficiency of the system, or, in other words, productivity, it should be noted that since the productivity is obtained from the weighted average of divisional efficiencies (operational efficiency, service effectiveness, and social effectiveness); productivity is always within the range of the lowest scores and the highest scores of the divisional efficiencies. According to Fig. 3, the social effectiveness (sub-process 3) of all bank branches except branch Z is lower than the productivity (global); indicating the inappropriate condition of the studied bank branches regarding social effectiveness has led to a dramatic drop in productivity. Another evidence for this claim is the rank correlation between the rank of divisions (sub-processes) and the productivity rank, which confirms the above statement. Thus, the rank correlation between overall efficiency (productivity) rank and the rank of operational efficiency, service effectiveness, and social effectiveness is 33.4%, 84.0%, and 95.3%, respectively. One of the interesting results of this analysis is that the best branch in terms of productivity (branch C) is a branch that operates at an acceptable level in all fields; this indicates that one-dimensional activity or ignoring some dimensions will lead to a sharp drop in productivity. For instance, the productivity of branch Q and branch I is almost equal (48%), with the difference that branch Q, is the efficient branch in terms of operational efficiency and service effectiveness and obtaining the 100% score. However, branch I has not obtained a score of 100% in any areas, though having an acceptable activity in all the three areas (operational efficiency: 39%, service effectiveness: 66%, and social effectiveness 43%). The reason for the sharp drop in productivity in branch Q, despite the operational efficiency and service effectiveness score of 100%, is the lack of attention to social effectiveness (social effectiveness score: 23%). As another witness to this issue, it can
be seen that branch M, which is characterized by two inefficient areas (service effectiveness: 57% and social effectiveness: 30%) and only one efficient area (operational efficiency: 100%), has obtained a higher productivity score compared to branch Q. Therefore, with regard to the above-mentioned evidence, the importance of paying attention to all aspects and not neglecting certain dimensions is so clear, and managers of bank branches should avoid one-dimensional attitude and move to multidimensional attitude and pay special attention to all aspects.

Table 4 shows the overall efficiency of the network system and the efficiency of each division, along with the weight of each division, under the VRS assumption, so that the average weighted efficiency of the divisions is equal to overall efficiency. Also, Table 4 shows the reference units with their shadow price values for each division. As previously stated and shown in Table 4, the total values of shadow prices in each division is equal to one under the VRS assumption.

### Table 4. Results of NSBM (VRS assumption)

<table>
<thead>
<tr>
<th>DMU</th>
<th>Overall Score (Rank)</th>
<th>Divisional Score</th>
<th>Reference Set</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Overall (Score Rank)</td>
<td>W1</td>
<td>W2</td>
</tr>
<tr>
<td>A</td>
<td>1.0000 (1)</td>
<td>1.0000 (1)</td>
<td>1.0000 (1)</td>
</tr>
<tr>
<td>B</td>
<td>0.5174 (15)</td>
<td>0.2756 (5)</td>
<td>0.2756 (5)</td>
</tr>
<tr>
<td>C</td>
<td>0.8267 (7)</td>
<td>0.7669 (8)</td>
<td>0.6140 (8)</td>
</tr>
<tr>
<td>D</td>
<td>0.7774 (7)</td>
<td>0.2591 (15)</td>
<td>0.3406 (15)</td>
</tr>
<tr>
<td>E</td>
<td>0.1072 (37)</td>
<td>0.4219 (36)</td>
<td>0.0576 (36)</td>
</tr>
<tr>
<td>F</td>
<td>0.2665 (36)</td>
<td>0.1614 (33)</td>
<td>0.2461 (28)</td>
</tr>
<tr>
<td>G</td>
<td>0.4881 (18)</td>
<td>0.3702 (22)</td>
<td>0.4090 (18)</td>
</tr>
<tr>
<td>H</td>
<td>0.2946 (28)</td>
<td>0.0982 (6)</td>
<td>0.609 (21)</td>
</tr>
<tr>
<td>I</td>
<td>0.6984 (37)</td>
<td>0.4494 (35)</td>
<td>0.6780 (10)</td>
</tr>
<tr>
<td>J</td>
<td>0.2703 (30)</td>
<td>0.5702 (31)</td>
<td>0.1833 (31)</td>
</tr>
<tr>
<td>K</td>
<td>0.6041 (12)</td>
<td>0.5626 (12)</td>
<td>0.4568 (14)</td>
</tr>
<tr>
<td>L</td>
<td>0.6195 (10)</td>
<td>0.6800 (9)</td>
<td>0.4215 (16)</td>
</tr>
<tr>
<td>M</td>
<td>0.7916 (6)</td>
<td>0.2639 (11)</td>
<td>0.5587 (11)</td>
</tr>
<tr>
<td>N</td>
<td>0.1386 (36)</td>
<td>0.0653 (36)</td>
<td>0.2464 (32)</td>
</tr>
<tr>
<td>O</td>
<td>0.2467 (32)</td>
<td>0.3858 (21)</td>
<td>0.1229 (36)</td>
</tr>
<tr>
<td>P</td>
<td>0.6134 (11)</td>
<td>0.2789 (11)</td>
<td>0.4643 (12)</td>
</tr>
<tr>
<td>Q</td>
<td>0.8821 (4)</td>
<td>0.2940 (5)</td>
<td>0.7183 (5)</td>
</tr>
<tr>
<td>R</td>
<td>0.2385 (33)</td>
<td>0.4837 (34)</td>
<td>0.2318 (29)</td>
</tr>
<tr>
<td>S</td>
<td>0.3057 (27)</td>
<td>0.2916 (26)</td>
<td>0.1857 (33)</td>
</tr>
</tbody>
</table>
4.2. Projection onto efficient frontiers

Depending on the efficient frontier, DMUs are divided into two groups of efficient and inefficient units. Efficient units are units that are located on the efficient frontier, and inefficient units are units that are not located on the efficient frontier. As previously stated, we defined the projection of DMU 0 by (5). Here, for example, the projection of inefficient unit A is calculated with the CRS assumption. The optimal solution of the NSBM model for unit A is as follows:

\[ \theta_{0\text{glob}} = 0.3127405, \]
\[ t^* = 0.1183748, \quad S_{1, 2}\text{eff} = 0.000000, \quad S_{3, 4}\text{eff} = 1.482219, \quad S_{5, 6}\text{eff} = 3.098995, \quad S_{7, 8}\text{eff} = 0.000000, \]
\[ S_{1, 2}\text{eff} = 22329.88, \quad S_{3, 4}\text{eff} = 0.000000, \quad S_{5, 6}\text{eff} = 0.617841, \quad S_{7, 8}\text{eff} = 0.000000, \quad S_{9, 10}\text{eff} = 191.4504, \]
\[ S_{1, 2}\text{eff} = 0.000000, \quad S_{3, 4}\text{eff} = 0.000000, \quad S_{5, 6}\text{eff} = 0.3130049, \quad S_{7, 8}\text{eff} = 34.74193, \quad \lambda_{0}^{(1)} = 0.1220906, \]
\[ \lambda_{0}^{(1)} = 0.04482023, \quad \lambda_{0}^{(2)} = 0.3038014, \quad \lambda_{0}^{(3)} = 0.1928685, \quad \lambda_{0}^{(4)} = 0.1606845, \quad \lambda_{0}^{(5)} = 0.8561047 \]

Using the existing equations in the NSBM model (by the formula (4)), we have:
Table 5 reports the projection of inputs, outputs and links of branch A under the CRS assumption onto the efficient frontiers by the formula (5).

**Table 5. Projection of DMU A onto efficient frontiers**

<table>
<thead>
<tr>
<th>Variables</th>
<th>DMU A</th>
<th>Projection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employees</td>
<td>19</td>
<td>19=0.19</td>
</tr>
<tr>
<td>Fixed assets</td>
<td>17.654</td>
<td>17.654=12.52141=5.13259</td>
</tr>
<tr>
<td>Non-operating costs</td>
<td>71.749</td>
<td>71.749=26.17952=5.13259</td>
</tr>
<tr>
<td>Number of transactions</td>
<td>108416</td>
<td>108416=188637.10857=297053.10857</td>
</tr>
<tr>
<td>Number of accounts</td>
<td>11378</td>
<td>11378=0=11378</td>
</tr>
<tr>
<td>Bank deposits</td>
<td>593.964</td>
<td>593.964=2.64419=596.60819</td>
</tr>
<tr>
<td>Interest expenses</td>
<td>5.648</td>
<td>5.648=0.5648</td>
</tr>
<tr>
<td>Interest income</td>
<td>3.157</td>
<td>3.157=5.21936=8.37636</td>
</tr>
<tr>
<td>Non-interest income</td>
<td>1.753</td>
<td>1.753=0.1753</td>
</tr>
<tr>
<td>Bank facilities</td>
<td>408.961</td>
<td>408.961=293.49093=702.45193</td>
</tr>
<tr>
<td>Employment</td>
<td>396.794</td>
<td>396.794=1617.32396=2014.11796</td>
</tr>
</tbody>
</table>
According to the results of the optimal solution, since the slack variables of $S_{2}^{3.2^*}, S_{3}^{3.3^*} = 0$, the linking activities (bank deposits and bank facilities) are considered as output and as output variables are optimized by the model. The results shown in Table 5 also confirm this.

5. Policy Implication
The main motive for doing this research is that standard DEA models cannot provide the management with sufficient and adequate details to identify the specific inefficient resources in the interactions between segments of a bank, and also the existing models have some deficiencies in different areas. In the era of global economics, characterized by fierce competition without borders, bank branches are increasingly forced to seek new ways of identifying inefficiencies in order to maintain their competitive advantage in the market. This research shows how the NSBM model can meet the above-mentioned needs by providing insight into specific inefficient resources of organizational processes in Iranian banks. By opening the black box of the production process in the banking industry, the NSBM model provides access to basic and underlying diagnostic information in different sections that will otherwise remain unknown and unidentified. NDEA models are able to explain why some DMUs are in fact inefficient with the full efficiency score obtained using conventional black box models and considering the operation of the network process components in efficiency calculations. This research has the intention of designing a comprehensive and integrated model for measuring performance in the banking industry. For this purpose, the most important inputs, outputs, outcomes and impact of the banking system were identified.

There has been much debate on whether to include deposits as an input or as an output in performance evaluation of banking branches, because deposits have both input and output characteristics (based on whether the production approach is intended or an intermediation approach). On the other hand, none of these two approaches alone is complete and cannot reflect the comprehensive role of banks, and the use of each of these approaches has its own benefits. Therefore, in order to conduct a comprehensive assessment, this research is intended to evaluate the performance of bank branches under both the production and intermediation approaches. But an unresolved issue in this regard is the role of the decision maker in determining the deposit as an input or as an output, which reduces the value and confidence of the research results. Therefore, in this study, a specific model of data envelopment analysis was used in performance evaluation that simultaneously evaluates deposits as inputs and outputs (without external interference).
Unlike other models, which the weight of each division (sub-process) is externally imposed on the model and is the same for all branches, here in the network model used in this study, the weight of each division is determined using NSBM model by the optimal solution of the model and based on the special conditions of the unit under study.

The components’ efficiency and the weights associated with them help to identify the most influential factors in the system performance, and proper control of these factors will effectively improve the performance of the system.

6. Concluding remarks and direction for future research

Performance evaluation of banking industry is very useful for the management that monitors the performance and for the lawmakers who monitor the financial stability, if they are to identify disturbances. Additionally, in their investment portfolios, investors and market analysts are also interested in including the ranking of financial institutions and banks. Other benefits are also derived from banking performance evaluation studies, which include ranking of bank branches based on estimating the efficiency, monitoring the performance of bank branches, transferring resources from one bank to another, measuring the effect of internal and external factors on the efficiency of bank branches, finding solutions to improve inefficient branches by establishing virtual units for inefficient branches, and understanding regional and geographical differences in efficiency frontiers, and so on. The development of reliable performance measures also lead to better decisions by the executives who usually seek the synergy of the integration of the branches that provides complementary services. Therefore, in this study, in order to establish a comprehensive assessment in the banking industry, for the first time, bank branches are evaluated under both the production and intermediation approaches in the form of a new model of DEA. Doing so, the controversy about the role of the deposit as the output of the production approach and the input of the intermediation approach is resolved using an integrated and coherent approach. In the present study, an SBM model was used for network systems so that the system efficiency is equal to the weighted mean of its components efficiencies as well as the weight of each process is its contribution to the inefficiency of the overall system. Weights are not predetermined by the decision maker; instead, they will be the most desirable weights in calculating the efficiency of the overall system, and this is more consistent with the nature of DEA since it is obtained based on the optimal solution of the model, the specific conditions of the branch, and without any external intervention. When compared with radial models, it can be said that the main
advantage of slacks-based measure model is its ability in providing appropriately measure the efficiency of weakly efficient DMUs. Another positive feature of the model used in this study is the simultaneous evaluation of the mediator variables as both input and output and its determination through the model. In the present study, the necessity of considering the internal processes of a DMU in the performance evaluation of that DMU was addressed. Subsequently, the proposed NSBM model was used to evaluate the performance of 37 branches of one of the largest commercial banks in Iran. The results showed that proposed model has high discriminating power than traditional black box models. In addition, the system’s efficiency is in fact the weighted average of its components' efficiency. The components' efficiency and the weights associated with them help to identify the most influential factors in the system performance, and proper control of these factors will effectively improve the performance of the system. Another finding of this study indicates that in structures that have network and communication effects between divisions, the use of black box and separation approaches does not provide a true and accurate evaluation of performance; instead, appropriate network models should be used in these cases since in the network approaches the results reflect system performance in a better and more realistic way.

Acknowledgment
The authors would like to thank, Professor Carl Chen, the editor of International Review of Economics and Finance, and two anonymous reviewers for their insightful and constructive comments and suggestions, as results the paper has been improved substantially.

References


