Accounting and Auditing with Application

www.aaa.reapress.com

Acc. Aud. Appl. Vol. 2, No. 3 (2025) 183–189.

Paper Type: Original Article

Deep Analysis of Financial Indicators Affecting Bank Efficiency Using the Monte Carlo Simulation Technique

Zahra Joorbonyan*

Department of Management, Faculty of Humanities, University of Guilan, Guilan, Iran; zahrajbn20@gmail.com.

Citation:

Received: 24 February 2025 *Revised: 19 April 2025* *Accepted: 17 June 2025*	Joorbonyan, Z. (2025). Deep analysis of financial indicators affecting bank efficiency using the monte carlo simulation technique. *Accounting and Auditing with Application, 2*(3), 183-189.

Abstract

In today's competitive environment, evaluating bank branch performance plays a crucial role in managerial decision-making. Inefficient branches continuously strive to improve their efficiency, while efficient ones seek to maintain their superior positions. Discriminant Analysis is a common classification method in banking, used to predict the status of new branches based on data from existing ones. However, predictions from this method often involve uncertainty. This study introduces a confidence level metric to determine the status of new branches more accurately. Utilizing sensitivity analysis based on Monte Carlo simulation, the impact of various financial indicators on this confidence level is assessed, identifying key indicators that influence the classification of branches as efficient or inefficient. The results reveal that long-term deposits hold significant importance, whereas variables such as number of personnel, overdue receivables, and Qarz al-Hasna deposits have negligible effects on efficiency classification. These findings provide valuable insights for bank managers in establishing and managing new branches, and enable targeted planning to reform and guide inefficient units towards enhanced efficiency.

Keywords: Monte Carlo simulation, Discriminant analysis, Confidence level, Sensitivity analysis, Bank branch efficiency.

1 | Introduction

The banking system plays an important role in economic stability as a factor in implementing monetary policies. For this reason, the profitability and income of banks have always been a concern for experts and the general public alike, as the optimal functioning of banks significantly impacts the economic growth and development of the country. In recent years, threats and pressures resulting from globalisation and the growth

Corresponding Author: zahrajbn20@gmail.com

10.22105/aaa.vi.73

of non-bank financial and credit institutions have forced banks to improve their performance in order to survive and compete in the market. Globalisation has led to the establishment of research centres and the conduct of research activities comparing their situation with that of other banks. Performance measurement is one of the most effective ways for banks to obtain information for decision-making purposes. Many studies have been conducted to measure bank efficiency and evaluate their ranking, and numerous models from mathematics, statistics, and operations research have been designed to evaluate banks accurately. However, many of these models are classical and cannot evaluate thoroughly and optimally, paving the way for more complex models to enter this field. One basic approach to setting up productivity and efficiency improvement programmes at bank level is for each bank to examine its future situation, identify the factors affecting the inefficiency of its branches, and plan accordingly to guide inefficient units towards improvement [1].

Discriminant analysis can be applied to many cases of decision-making. Since Fisher introduced his model, most subsequent models have been deterministic, whereas in practice, most variables are uncertain [2]. One important feature of banking is uncertainty in banking parameters and indicators, which can affect bank efficiency [3]. One relatively new approach to optimisation under conditions of data uncertainty is the use of simulation techniques. Banking indicators are usually subject to high levels of uncertainty due to measurement errors, a lack of information, and an incomplete understanding of the factors affecting them. Incomplete knowledge reduces confidence in the model output. This study aims to investigate which of the considered banking indicators, such as number of personnel, types of deposits, types of facilities, overdue claims, and banking fees, have the greatest impact on the ranking of the bank branch in question. The results can help banking managers and experts to identify effective performance evaluation indicators and facilitate decision-making [4].

2 | Confidence Level

Discriminant analysis is a decision-making tool for predicting the classification of new observations and assigning them to previously defined categories. In this method, a group of observations whose membership in different groups is known is used to estimate the weights of a discriminant function. Mangsarian introduced a linear discriminant function using linear programming for the case where there are two sets of linearly separable observations [5]. Later studies developed linear programming methods using criteria such as minimizing the sum of deviations or maximizing the smallest deviation from the discriminant function produced for the case where the two sets are not linearly separable [6], [7].

In a number of studies, they combined discriminant analysis and goal programming to introduce various models, considering criteria such as minimising maximum deviation, maximising minimum deviation, minimising the sum of internal deviations, minimising the sum of deviations, minimising misclassified observations, and maximising the ratio of internal to external deviations [8–10]. Assuming the membership of observations in groups is known in these models, a hyperplane is defined that separates the two groups using a set of weights and a threshold value. This hyperplane can then be used to predict the group to which new observations belong.

Siyoshi presented a new model of discriminant analysis by combining the collective model of data envelopment analysis and discriminant analysis, and using objective programming [11], [12]. This model is known as the DEA/DA method. One problem with existing discriminant analysis methods is that they only classify new observations, without providing information about the degree of confidence of their membership of the specified group. To address this issue, this study defines a degree of confidence for the new observation, providing more details that can inform better decision-making. In the DEA/DA model, the classification of the new observation is determined by solving two models. In order to determine the degree of confidence, first determine the state in which the new observation $r_o=(r_{1o},\dots,r_{mo})^T$ belongs to the first group. The hyperplane produced in the first stage $\sum_{i=1}^{m}\lambda_i^* r_{io} = a_o^*$ is, so the distance r_o is obtained from this hyperplane. Suppose $\lambda^* = (\lambda_1^*,\dots,\lambda_m^*)$ Therefore, the distance $r_o=(r_{1o},\dots,r_{mo})^T$. And this hyperplane is obtained from the following relationship:

$$\bar{\bar{d}}_1 = \frac{\|\lambda^*.r_o - a_o^*\|_p}{\|\lambda^*\|_p}. \tag{1}$$

For simplicity, *Eq. (1)* in the second norm is written as follows:

$$\bar{\bar{d}}_1 = \frac{|\lambda^*.r_o - a_o^*|}{\sqrt[2]{{\lambda_1^*}^2 + \cdots + {\lambda_m^*}^2}}. \tag{2}$$

The distance of all observations of the first group from the hyperplane $\sum_{i=1}^{m} \lambda_i^* r_{io} = a_o^*$ is calculated as follows:

$$\tilde{d}_{1j} = \frac{|\lambda^*.r_j - a_o^*|}{\sqrt[2]{{\lambda_1^*}^2 + \cdots + {\lambda_m^*}^2}}, \quad j \in G_1. \tag{3}$$

And the distance of the farthest point from this hyperplane is equal to $\hat{d} = \max\{\tilde{d}_{1j} | j \in G_1\}$. Therefore, to normalize the value of $\bar{\bar{d}}_1$, it is divided by the maximum distance.

$$d_1 = \frac{\bar{\bar{d}}_1}{\hat{d}}.$$

Similarly, the observation distance from the hyperplane generated in the second step $\sum_{i=1}^{m} \mu_i^* r_{io} = c^*$, called d2, is calculated.

In this study, the distance from both hyperplanes is used to determine the degree of confidence. For this reason, the degree of confidence is defined as follows:

$$d = d_1 \times d_2. \tag{4}$$

A confidence level close to one indicates that the observation belongs to the group with high confidence, while a confidence level close to zero indicates that the new observation belongs to the group with low confidence. If the new observation is the most distant from the hyperplane, the values $d_1=1$ or $d_2=1$ are considered, meaning the new observation is also included in the set of observations used to calculate the confidence level. Similarly, if the new observation belongs to the second group, the confidence level is obtained in a similar way.

3 | The Role of Financial Indicators Using Monte Carlo Simulation

In this section, the role of financial indicators affecting the degree of confidence is examined. For this purpose, suppose that the new observation is represented as $r_o = (r_{1o}, \ldots, r_{mo})^T$, whose parameters are independent and have a probability density function *f*, so it is represented as follows [13]:

$$d = f(r_o), \qquad r_o = (r_{1o}, r_{2o}, \ldots, r_{mo}) \in K^m \equiv [0,1]^m. \tag{5}$$

The sensitivity index is obtained by decomposing the function f into additive sums in the following form [14]:

$$f(r_o) = f(r_{1o}, r_{2o}, \ldots, r_{mo}) = f_c + \sum_{i=1}^{m} \sum_{q_1<\cdots<q_i} f_{q_1,\ldots,q_i}(r_{q_1}, r_{q_2}, \ldots, r_{q_i}), \tag{6}$$

where f_c is the median value of the function, and the integral of each sum over its variables is zero. Therefore:

$$\int_{U^n} f_{i_1,\ldots,i_t} f_{j_1,\ldots,j_q} \, dr_o = 0, \quad (i_1, \cdots, i_t) \neq (j_1, \cdots, j_q). \tag{7}$$

Which are calculated using the following multidimensional integrals:

$$f_c = \int_{K^m} f(r_o)\, dr_o, f_i(r_{io}) = -f_c + \int_0^1 \cdots \int_0^1 f(r_o) dr_{o\sim i}, \; f_{ij}(r_{io}, r_{jo}) = -f_c - f_i(r_{io}) - f_j(r_{jo}) + \int_0^1 \cdots \int_0^1 f(r_o) dr_{o\sim(ij)}, \tag{8}$$

denote the integral over all variables except r_{io}, r_{io}, and r_{jo}, respectively. To calculate the sensitivity where $dr_{o\sim(ij)}$ and $dr_{o\sim i}$ index, SI_i, the variance of the confidence level must be obtained as follows:

$$V_d = \int_{K^m} f^2(r_o)\, dr_o - f_c^2. \tag{9}$$

By squaring *Eq. (8)* and integrating over K^m and using the property given in *Eq. (7)*, we obtain:

$$V_d = \sum_i V_i + \sum_{i<h} V_{ih} + \sum_{i<h<k} V_k + \cdots + V_{1,2,\cdots,m}. \tag{10}$$

$$V_i = V[E(d|r_{io} = r_{io}^*)]. \tag{11}$$

This expression yields the sensitivity of d to the r_{io} factor. Accordingly, the sensitivity index is defined as follows:

$$SI_i = \frac{V_i}{V_d}. \tag{12}$$

Initially, this section classifies the 78 Mellat Bank branches with more than 20 employees in terms of efficiency using a collective model. Next, the efficiency or inefficiency of a new branch is predicted. According to the method presented in this study, the degree of confidence in this new branch is calculated. Finally, the sensitivity of this new branch's efficiency to the bank's indicators is evaluated. When classifying these branches, the number of employees, interest paid, and overdue receivables are considered inputs, while facilities, long-term deposits, current deposits, Qarz-ol-Hasana deposits, short-term deposits, interest received, and fees received are considered output variables. Using this collective model, 45 of the 78 branches studied are considered efficient, while 33 are considered inefficient.

The values of the weighted estimates of the discriminant functions obtained in both stages are given in *Table 1*. To control the imbalance between the data and the difference in the importance of efficient and inefficient data, the weight w is assigned to the groups in the second stage of DEA-DA. In this section, the correctly classified data are more important, so w is set to 1. Consider now the new observation A, whose information is given in *Table 2*. Using the results of the first stage of DEA-DA, it is found to be within the overlap, and the second stage concludes that it is inefficient. Using *Table 1*, the obtained confidence level is 0.0283, corresponding to the overlap of this observation. Given that the confidence level of this observation is low, it should be classified as inefficient with caution. Also, greater care should be taken when collecting or estimating its data.

Sensitivity analysis using the Monte Carlo method relies on repeated random sampling to calculate the results [12]. The Monte Carlo method uses random numbers to simulate the parameters. In other words, by taking into account the Coefficient of Variation (COV) for each parameter, a set of random numbers is generated whose mean equals the exact value of the variable. The COV indicates the amount of dispersion around the mean value and is defined as follows:

$$COV = \frac{\sigma}{\mu}. \tag{13}$$

In this context, σ represents the standard deviation and μ represents the mean value. The following pattern is used to generate random variables relating to each parameter:

$$r_{ij} = \mu_{r_{ij}} \left(1 + COV_{r_{ij}} \alpha_l\right), \tag{14}$$

where $\mu_{r_{ij}}$ and $COV_{r_{ij}}$ express the mean values and coefficients of variation of random parameters, respectively, and α_l is a random parameter with a mean of zero that is used in a Monte Carlo simulation. A large amount of uncertainty usually accompanies the collection of banking data, so the COV of these variables plays an important role in efficiency changes. However, as obtaining the COV requires extensive research in banking, which is beyond the scope of this study, we consider the COV of all variables to be 0.05 in this section. The sensitivity of the indicators to changes in the degree of confidence is examined using the Monte Carlo method. The results of the sensitivity analysis for different parameters are shown in *Fig. 1*. These results show the extent to which changes or uncertainty in each parameter affect the degree of confidence in the desired observation. This diagram shows the relative contribution of each parameter to changes in the degree of confidence.

The total sensitivity index for different parameters is approximately 99%, so the higher order indices have very small values, meaning their effect on the degree of confidence is very insignificant; therefore, their calculation has been omitted. The results given in *Fig. 1* indicate that the impact of long-term deposits on changes in the degree of confidence is 65%, and changes in the three indicators of the number of personnel, Qarz-ul-Hasana deposits, and overdue claims have the least effect on changes in the degree of confidence. Therefore, the most important source of uncertainty in the degree of confidence is due to long-term deposits and interest received, and their values should be collected and estimated more accurately.

Table 1. Hyperplane weights obtained from DEA-DA.

Index	First stage	Second Stage
personnel	-0.86752	-0.00100
Facilities	-0.00100	0.001833
Long-term deposits	0.001747	0.038314
Current deposits	0.001031	0.002730
Qarz al-Hasana deposits	0.002191	0.008920
Short-term deposits	0.001000	0.010257
Interest paid	-0.02534	-0.47351
Interest received	0.060138	0.001000
Commission received	0.039035	-0.043627
Deferred claims	-0.00100	-0.02617

Table 2. New observation information.

Index	C
personnel	29
Facilities	10^8×899
Long-term deposits	10^8×259
Current deposits	10^8×390
Qarz al-Hasana deposits	10^8×90
Short-term deposits	10^8×201
Interest paid	10^8×22
Interest received	10^8×13
Commission received	10^6×2016
Deferred claims	10^8×103

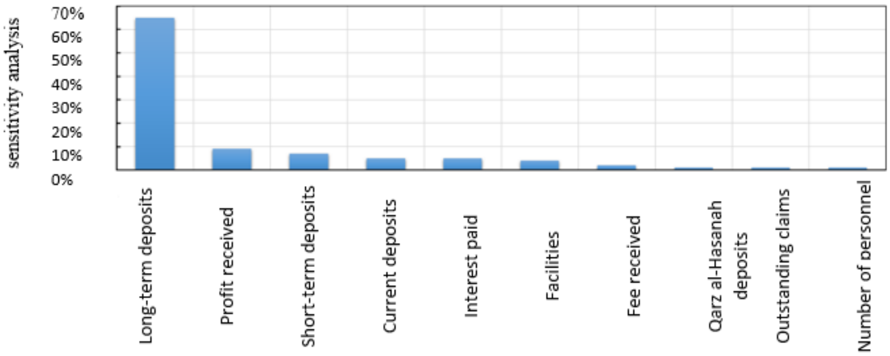

Fig. 1. Results of sensitivity analysis on new observation.

4 | Conclusion

Classification models assign observations whose group membership is unknown to specified groups using a set of parameters dependent on each observation. Discriminant analysis is one such model, used to predict the membership of an observation. This study introduces a two-stage discriminant analysis model, where the first stage determines the presence of overlap and the second stage reclassifies observations within this area. A limitation of the discriminant analysis method is that it can only determine the category of a new observation. However, it is often necessary to have more information about the new observation. Therefore, this study introduces a degree of confidence based on the distance of the observation from the hyperplane.

The degree of confidence is then demonstrated using a numerical example of a new observation within the overlap. Next, sensitivity analysis using the Monte Carlo method is employed to calculate the effect of input indicators on output changes. The sensitivity analysis results showed that long-term deposit indicators have the most significant effect, while the number of personnel has the least effect, meaning that increasing the number of personnel has a minimal impact on changes in the degree of confidence. It is worth noting that, since the results of the sensitivity analysis depend on the value of the COV, appropriate studies must be conducted to determine the variance of the parameters in order to obtain an appropriate COV in the sensitivity analysis.

Author Contributaion

Joorbonyan designed the research framework, developed the two-stage discriminant analysis model, performed the Monte Carlo sensitivity analysis, analyzed and interpreted the results, and prepared the manuscript.

Funding

This research received no external funding.

Data Availability

The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

The author declares no conflict of interest.

References

[1] Coccorese, P. (2017). Banking competition and economic growth. In *Handbook of competition in banking and finance* (pp. 230–263). Edward Elgar Publishing. https://doi.org/10.4337/9781785363306.00020

[2] Fisher, R. A. (1936). The use of multiple measurements in taxonomic problems. *Annals of eugenics*, *7*(2), 179–188. https://doi.org/10.1111/j.1469-1809.1936.tb02137.x

[3] Peykani, P., Mohammadi, E., Jabbarzadeh, A., & Jandaghian, A. (2016). Utilizing robust data envelopment analysis model for measuring efficiency of stock, a case study: Tehran Stock Exchange, *1*, 15–24. https://b2n.ir/mm8921

[4] Karimi, T., Ahmadian, M., & Shahbazi, M. (2024). Performance evaluation of bank branches in the atmosphere of grey uncertainty. *Journal of modelling in management*, *19*(6), 2300–2319. https://doi.org/10.1108/JM2-09-2023-0206

[5] Mangasarian, O. L. (1965). Linear and nonlinear separation of patterns by linear programming. *Operations research*, *13*(3), 444–452. https://doi.org/10.1287/opre.13.3.444

[6] Freed, N., & Glover, F. (1981). Simple but powerful goal programming models for discriminant problems. *European journal of operational research*, *7*(1), 44–60. https://doi.org/10.1016/0377-2217(81)90048-5

[7] Olfati, M., Krömer, P., Fanati Rashidi, S., Mirjalili, S., & Snášel, V. (2025). A goal programming-based algorithm for solving multi objective optimization problems. *Annals of operations research*, 1–26. https://doi.org/10.1007/s10479-025-06646-0

[8] Banks, W. J., & Abad, P. L. (1991). An efficient optimal solution algorithm for the classification problem. *Decision sciences*, *22*(5), 1008–1023. https://doi.org/10.1111/j.1540-5915.1991.tb01904.x

[9] Glover, F. (1990). Improved linear programming models for discriminant analysis. *Decision sciences*, *21*(4), 771–785. https://doi.org/10.1111/j.1540-5915.1990.tb01249.x

[10] Retzlaff-Roberts, D. L. (1996). A ratio model for discriminant analysis using linear programming. *European journal of operational research*, *94*(1), 112–121. https://doi.org/10.1016/0377-2217(95)00196-4

[11] Sueyoshi, T. (1999). DEA-discriminant analysis in the view of goal programming. *European journal of operational research*, *115*(3), 564–582. https://doi.org/10.1016/S0377-2217(98)00014-9

[12] Sueyoshi, T. (2001). Extended DEA-discriminant analysis. *European journal of operational research*, *131*(2), 324–351. https://doi.org/10.1016/S0377-2217(00)00054-0

[13] Hasan, S., & Elwakil, E. (2021). Knowledge-driven stochastic reliable modeling for steel bridge deck condition rating prediction. *Journal of structural integrity and maintenance*, *6*(2), 91–98. https://doi.org/10.1080/24705314.2020.1862965

[14] Saltelli, A., Chan, K., & Scott, E. M. (2000). Sensitivity analysis: Gauging the worth of scientific models. John Wiley & Sons. https://B2n.ir/pn8336