
Evaluating loans using variable benchmark data envelopment analysis

Rashmi Malhotra*

Decision and Systems Sciences Department,
Saint Joseph's University,
5600 City Avenue, Philadelphia, PA 19131, USA
Email: rmalhotr@sju.edu
*Corresponding author

George Tsetsekos

Department of Finance,
Drexel University, USA
Email: tsetsekos@drexel.edu

Abstract: Loan officers use many business intelligence methods to screen consumer loan applications besides intuitive judgement and experience. They also use mathematical techniques such as credit-scoring models, traditional statistical models, and artificial intelligence methods such as expert systems, artificial neural systems, and fuzzy logic. This study illustrates the development of a decision support system using variable benchmark data envelopment analysis model to predicting bad loans. Further, the study also compares the performance of the DEA model with linear discriminant analysis model. The study illustrates the viability of the variable benchmark DEA model that outperforms the linear discriminant analysis model.

Keywords: benchmarking; data envelopment analysis; DEA; linear discriminant analysis; LDA: decision support system.

Reference to this paper should be made as follows: Malhotra, R. and Tsetsekos, G. (2016) 'Evaluating loans using variable benchmark data envelopment analysis', *Int. J. Business Intelligence and Systems Engineering*, Vol. 1, No. 1, pp.77–98.

Biographical notes: Rashmi Malhotra has published over 120 research articles in various journals and conference proceedings. She has presented about 85 papers. She is the Editor-in-Chief of the *International Journal of Business Intelligence and Systems Engineering*. She serves on the editorial boards of *International Journal of Applied Management Science*, *International Journal of Data Analysis Techniques and Strategies*, *International Journal of Strategic Decision Sciences*, *International Journal of Management Research*, *Journal of Management Systems and Collaborative Filtering Using Data Mining and Analysis*. She is the Editor of the Special Issue of *International Journal of Data Analysis Techniques and Strategies on Business Intelligence*.

George Tsetsekos serves as the Francis Professor of Finance at the LeBow College of Business at Drexel University. His area of research covers international finance and corporate finance. His work has been published in several journals including the *Journal of Financial Economics* and *Journal of*

International Business Studies. He has also published two books in finance and capital markets. In the past, he served as the Dean of LeBow College (2002–2012) and Vice Provost at Drexel University (1997–2012).

1 Introduction

Organisations use decision support systems that apply different business intelligence techniques such as statistical models, scoring models, neural networks, expert systems, neuro-fuzzy systems, case-based systems, or simply rules that have been developed through experience. Managers need a decision-making approach that is robust, competent, effective, efficient, and integrative to handle the multi-dimensional organisational entities. However, quite frequently, the decision maker cannot simultaneously handle data from different sources. Hence, we recommend that managers analyse different aspects of data from multiple models separately and integrate the results of the analysis. In addition, evaluating loans and differentiating good loans from bad loans has gained even more importance in the wake of subprime mortgage crisis of 2007–2010 that stemmed from the expansion of mortgage credit to potential homebuyers with below average credit histories. Thus, financial institutions should equip loan officers with appropriate business intelligence tools to assess the credit worthiness of a loan to manage risk. This study illustrates the design of a business intelligence and analytics tools, a multi-attribute-decision-support-system that uses the variable-benchmark data envelopment analysis (DEA) models to evaluate a loan application. DEA evaluates and measures the relative efficiency of decision making units that utilises multiple inputs and outputs to provide non-objective measures without making any specific assumptions about data. DEA's major advantage is to clearly identify the important factors contributing to the success of a decision. Further, to illustrate the viability of the proposed methodology, the study also compares the performance of the linear discriminant analysis (LDA) model with the variable benchmark DEA model. The rest of the paper is organised as follows; Section 2 we provide a literature review of previous studies on financial statement analysis, Section 3 discusses the methodology used, Section 4 describes the DEA model and LDA model, Section 5 provides an empirical analysis of our results, and Section 6 summarises and concludes our study.

2 Literature review

Recently, many studies have illustrated the use of DEA, a non-parametric methodology to analyse different aspects of business entities. The details of the DEA model are discussed in the next section. In contrast to other methodologies, DEA is one of the methods that have traditionally been used to assess the comparative efficiency of homogenous operating units such as schools, hospitals, utility companies, sales outlets, prisons, and military operations. This section evaluates the existing studies in literature using DEA and neuro-fuzzy and neural network models.

2.1 Previous studies on application of DEA to finance industry

More recently, many authors have applied DEA to banks (Haslem et al., 1999) and mutual funds (Haslem and Scheraga, 2003; Galagedera and Silvapulle, 2002; McMullen and Strong, 1998; Murthi et al., 1997). Murthi et al. (1997) examine the market efficiency of the mutual fund industry by different investment objectives. McMullen and Strong (1998) applied DEA to evaluate the relative performance of 135 US common stock funds using one, three, and five-year annualised returns, standard deviation of returns, sales charge, minimum initial investment, and expense ratio. Sedzro and Sardano (1999) analysed 58 US equity funds in Canada using DEA with annual return, expense ratio, minimum initial investment and a proxy for risk as factors associated with fund performance. Galagedera and Silvapulle (2002) use DEA to measure the relative efficiency of 257 Australian mutual funds. Haslem and Scheraga (2003) use DEA to identify efficiencies in the large-cap mutual funds in the 1999 Morningstar 500. Basso and Funari (2001) propose the use of DEA methodology to evaluate the performance of mutual funds. Morey and Morey (1999) present two basic quadratic programming approaches for identifying those funds that are strictly dominated, regardless of the weightings on different time horizons being considered, relative to their mean returns and risks. Edirisinghe and Zhang (2007) develop a DEA model to evaluate a firm's financial statements over time in order to determine a relative financial strength indicator that can predict firm's stock price returns.

Zhu (2000) uses DEA to develop a multi-factor financial performance model that recognises tradeoffs among various financial measures. Kao and Liu (2004) compute efficiency scores based on the data contained in the financial statements of Taiwanese banks. They use this data to make advanced predictions of the performances of 24 commercial banks in Taiwan. Pille and Paradi (2002) analyse the financial performance of Ontario credit unions. They develop models to detect weaknesses in Credit Unions in Ontario, Canada. Ozcan and McCue (1996) use DEA for measuring and assessing the financial performance for hospitals. Halkous and Salamouis (2004) explore the efficiency of Greek banks with the use of a number of suggested financial efficiency ratios for the time period 1997–1999. Neal (2004) investigates X-efficiency and productivity change in Australian banking between 1995 and 1999 using DEA and Malmquist productivity indexes. Paradi and Schaffnit (2004) evaluate the performance of the commercial branches of a large Canadian bank using DEA. Chen et al. (2005) study the efficiency and productivity growth of commercial banks in Taiwan before and after financial holding corporations' establishment. Howland and Rowse (2006) assess the efficiency of branches of a major Canadian bank by benchmarking them against the DEA model of US bank branch efficiency. Sufian (2007) uses DEA approach to evaluate trends in the efficiency of the Singapore banking sector. Sanjeev (2007) evaluates the efficiency of the public sector banks operating in India for a period of five years (1997–2001) using DEA. Lin et al. (2007) study the relative efficiency of management in the Taiwanese banking system through DEA. Bergendahl and Lindblom (2008) develop principles for an evaluation of the efficiency of a savings bank using DEA as a method to consider the service orientation of savings banks.

2.2 Previous studies on application of DEA to assess the financial performance of non-financial sector

Feroz et al. (2003) illustrate the use of DEA to evaluate the financial performance of oil and gas industry. Hoon and Chunyan (1994) analysed the productive efficiency of the railway services in 19 Organization for Economic Cooperation and Development (OECD) countries. They report that railway systems with high dependence on public subsidies are less efficient than similar railways with less dependence on subsidies. Cowie and Riddington (1996) evaluate the efficiency of the European railways through the use of a production frontier approach. Yu and Lin (2008) uses a multi-activity network DEA model to simultaneously estimate passenger and freight technical efficiency, service effectiveness, and technical effectiveness for 20 selected railways for the year 2002. Lozano and Gutiérrez (2011) illustrate the slacks based measure of efficiency of 39 Spanish airports using DEA. Liu and Liu (2010) illustrate the use of DEA in evaluating and ranking the research and redevelopment performance of Taiwan's government-supported research institutes. Saranga and Moser (2010) develop a comprehensive performance measurement framework using the classical and two-stage value chain DEA model.

2.3 Previous studies on evaluating loans

Fritz and Hosemann (2000) illustrate the use of pattern recognition, neural network, genetic algorithms, and decision trees to compute monthly credit standing of corporate clients on the basis of their current accounts.

Baesens et al. (2005) study the use of survival analysis methods for credit scoring. They compared the performance of neural networks with proportional hazards model to predict customer defaults as well as the customers paying off their loan early.

Martens et al. (2007) illustrate the use of AntMiner+ algorithm, which is a rule induction technique based on the principles of MAX.MINAnt System to build credit scoring model to evaluate credit risk.

Desai et al. (1996) illustrate the use of neural networks for credit scoring purposes. They concluded that neural networks outperform traditional statistical methods for loan evaluation. Malhotra and Malhotra (2003) also show the use of neural networks in evaluating consumer loans in credit union environment. They showed that neural network can outperform traditional statistical models in differentiating good loans from bad loans.

Malhotra and Malhotra (2003) illustrate the use of neurofuzzy systems to differentiate good loans from bad loans. They showed that artificial neurofuzzy systems can perform better in separating good loans from bad loans.

3 Methodology

This section illustrates the LDA model and the DEA model.

3.1 The LDA model

The LDA model is one of the most widely used data classification methods to discriminate between two or more categories of data. The objective of LDA is to find a linear combination of features that characterise or differentiate between two or more classes of objects or events. LDA uses a linear transformation ('discriminant function') of the predictor variables of an event to yield a new set of transformed values that provides a more accurate discrimination than either of the predictors alone. This method maximises the ratio of between-class variance to the within-class variance in any particular data set thereby guaranteeing maximal separation. The objective of the transformation is to rotate the axes so that when the categories are projected on the new axes, the differences between the groups are maximised. In the ideal case, a projection should be found that completely separates the categories. However, in most cases there is no transformation that provides complete separation, so the goal is to find the transformation that minimises the overlap of the transformed distributions. The classic discriminant method was originally developed by Fisher (1936). For a given set of C classes with mean μ_i and same covariance Σ , we can calculate the between class variability using the sample covariance of the class means as

$$\sum_b = \frac{1}{C} \sum_{i=1}^C (\mu_i - \mu)(\mu_i - \mu)^T$$

where μ is the class mean and T is a given threshold. Then, the class separation in a given direction $\bar{\mathbf{u}}$ can be calculated as

$$S = \frac{\bar{\mathbf{u}}^T \sum_b \bar{\mathbf{u}}}{\bar{\mathbf{u}} \sum_b \bar{\mathbf{u}}}$$

This implies that when $\bar{\mathbf{u}}$ is an eigenvector of $\sum_b^{-1} \sum_b$, and the separation is equal to the corresponding eigenvalue. Since \sum_b is of most rank $C - 1$, then these non-zero eigenvectors identify a vector subspace containing the variability between features.¹

3.2 The DEA model

The DEA (Charnes et al., 1978) is a widely used optimisation-based technique that measures the relative performance of decision-making units (DMUs) that are characterised by a multiple objectives and/or multiple inputs structure. The DEA methodology measures the performance efficiency of organisation units called DMUs. This technique aims to measure how efficiently a DMU uses the resources available to

generate a set of outputs. The performance of DMUs is assessed in DEA using the concept of efficiency or productivity defined as a ratio of total outputs to total inputs. Efficiencies estimated using DEA are relative, that is, relative to the best performing DMU or DMUs (if multiple DMUs are the most efficient). The most efficient DMU is assigned an efficiency score of unity or 100%, and the performance of other DMUs vary between 0 and 100% relative to the best performance.

The generalised version of various DEA models that we propose to use in this study is listed in Table 1.²

Table 1 Generalised DEA models

Frontier	Type	Input-oriented	Output-oriented
		$Min \theta - \varepsilon \left(\sum_{i=1}^m s_i^- + \sum_{r=1}^s s_r^+ \right)$	$Max \phi - \varepsilon \left(\sum_{i=1}^m s_i^- + \sum_{r=1}^s s_r^+ \right)$
		Subject to	
		$\sum_{j=1}^n \lambda_j x_{ij} + s_i^- = \theta x_{io} \quad i = 1, 2, \dots, m$	$\sum_{j=1}^n \lambda_j x_{ij} + s_i^- = x_{io} \quad i = 1, 2, \dots, m$
	CRS	$\sum_{j=1}^n \lambda_j x_{rj} + s_r^+ = y_{ro} \quad r = 1, 2, \dots, s;$	$\sum_{j=1}^n \lambda_j y_{rj} + s_r^+ = \phi y_{ro} \quad r = 1, 2, \dots, s$
		$VRS : Add \sum_{j=1}^n \lambda_j = 1; \quad NIRS : Add \sum_{j=1}^n \lambda_j \leq 1; \quad NDRS : Add \sum_{j=1}^n \lambda_j \geq 1$	

Source: Zhu (2003)

where s are the slack variables; x represents input variables; y represent output variables; λ is a scalar factor, and θ and ϕ represent efficiency score of a DMU. Also, CRS is constant returns to scale (CRS), VRS denotes variable returns to scale, NIRS stands for non-increasing returns to scale, and NDRS refers to non-decreasing returns to scale.³

The main objective of the DEA methodology is to define a valid measure of comparison among peer DMUs so as to determine the relative position of the peer DMUs. Thus, the DEA establishes an empirical standard of excellence or best practices. Therefore, after establishing the frontier, or best practices, for benchmarking, we can measure a set of new DMUs relative to the benchmark (frontier). However, on encountering a new DMU that outperforms the existing benchmarks, the DEA generates a new efficiency frontier. The model (1) uses all the DMUs under evaluation, including the best-practice frontier and the new DMUs under study. As a result, we do not have the same benchmark (frontier) for the new DMUs. Thus, the new best-practice frontier does not directly compare the new DMUs to the established standard. Zhu (2003) modifies and extends the original DEA method as a benchmarking tool so that the new DMUs are evaluated against a set of given benchmarks (standards). Let E^* represent the benchmarks or the best-practice identified by the DEA. Based upon the input-oriented CRS envelopment model, we have the following model:

$$Min \delta_o^{CRS} \tag{1}$$

Subject to

$$\begin{aligned}
\sum \lambda_j x_{ij} &\leq \delta_o^{CRS} x_{io}^{New}, i = 1, \dots, m, \\
\sum \lambda_j y_{rj} &\geq y_{ro}^{New}, r = 1, \dots, s, \\
\lambda_j &\geq 0, j \in E^*,
\end{aligned} \tag{2}$$

where a new observation is represented by DMU_o^{New} with inputs x_{io}^{New} ($i = 1, \dots, m$) and outputs y_{ro}^{New} ($r = 1, \dots, s$). The superscript of CRS indicates that the benchmark frontier composed by benchmark DMUs in set E^* exhibits CRS. Model (2) measures the performance of DMU_o^{New} with respect to benchmark DMUs in set E^* when outputs are fixed at current levels. Similarly, we can have an output-oriented CRS envelopment model that measures the performance of DMU_o^{New} in terms of outputs when inputs are fixed at their current levels.

$$Max \tau_o^{CRS}$$

subject to

$$\begin{aligned}
\sum \lambda_j x_{ij} &\leq x_{io}^{New}, i = 1, \dots, m, \\
\sum \lambda_j y_{rj} &\geq \tau_o^{CRS} y_{ro}^{New}, r = 1, \dots, s, \\
\lambda_j &\geq 0, j \in E^*,
\end{aligned} \tag{3}$$

Based upon models (2) and (3), we have

$\delta_o^{CRS^*} = 1 / \tau_o^{CRS^*}$, where $\delta_o^{CRS^*}$ is the optimal value for model (2) and $\tau_o^{CRS^*}$ is the optimal value for model (3). Further, models (2) and (3) yield a benchmark for DMU_o^{New} . The i^{th} input and the r^{th} output for the benchmark can be expressed as

$$\begin{aligned}
\sum_{j \in E^*} \lambda_j^* x_{ij}, & \text{ (} i^{\text{th}} \text{ input)}, \\
\sum_{j \in E^*} \lambda_j^* y_{rj}, & \text{ (} r^{\text{th}} \text{ output)}.
\end{aligned} \tag{4}$$

Further, although the DMUs identified as the best-practice benchmarks are given as set E^* , the benchmark for each DMU may be different as it is represented by a combination of DMUs associated with the set E^* (4). Therefore, models (2) and (3) represent a variable-benchmark scenario.

Thus, the performance of DMU_o^{New} , using model (4) can be interpreted as follows:

- a $\delta_o^{CRS^*} = 1$ or $\tau_o^{CRS^*} = 1$ implies that DMU_o^{New} achieves the same performance level as the benchmark in model (4).
- b $\delta_o^{CRS^*} > 1$ or $\tau_o^{CRS^*} < 1$ implies that DMU_o^{New} has input savings or output surpluses as compared to the benchmark in model (4).

If we allow scale inefficiency, models (2) and (3) can incorporate scale inefficiency by assuming VRS. Therefore, we have the following input-oriented VRS variable-benchmark model:

$$\text{Min } \delta_o^{VRS}$$

subject to

$$\begin{aligned} \sum \lambda_j x_{ij} &\leq \delta_o^{VRS} x_{io}^{New}, i = 1, \dots, m, \\ \sum \lambda_j y_{rj} &\geq y_{ro}^{New}, r = 1, \dots, s, \\ \sum \lambda_j &= 1 \\ \lambda_j &\geq 0, j \in E^* \end{aligned} \quad (5)$$

$$\text{Max } \tau_o^{VRS}$$

subject to

$$\begin{aligned} \sum \lambda_j x_{ij} &\leq x_{io}^{New}, i = 1, \dots, m, \\ \sum \lambda_j y_{rj} &\geq \tau_o^{VRS} y_{ro}^{New}, r = 1, \dots, s, \\ \sum \lambda_j &= 1 \\ \lambda_j &\geq 0, j \in E^* \end{aligned} \quad (6)$$

As we introduce scale inefficiency, unlike models (2) and (3), models (5) and (6) may be infeasible.⁴ Thus, we have following two scenarios:

- a If model (5) is infeasible, then the output vector of DMU_o^{New} dominates the output vector of the benchmark in (2).
- b If model (6) is infeasible, then the input vector of DMU_o^{New} dominates the input vector of the benchmark in (3).

Based on the above scenarios, we have four cases:

- Case 1 When both models (5) and (6) are infeasible, this implies that DMU_o^{New} has the smallest input level and the largest output level as compared to the best-practices benchmark. Thus, DMU_o^{New} offers both input savings and output surpluses.
- Case 2 When model (5) is infeasible and model (6) is feasible, this implies that DMU_o^{New} has the largest output level as compared to the best-practices benchmark to make model (5) infeasible. Thus, we use model (6) to calculate the output surplus offered by DMU_o^{New} .
- Case 3 When model (6) is infeasible and model (5) is feasible, this implies that DMU_o^{New} has the smallest input level as compared to the best-practices benchmark to make model (6) infeasible. Thus, we use model (5) to calculate the input savings offered by DMU_o^{New} .
- Case 4 When both models (5) and (6) are feasible, this implies that we use both the models to determine if DMU_o^{New} offers input savings and output surpluses.

Case 5 The underperforming DMUs belong to this category. We can use benchmark values to find the source of low performance.

4 The data source and LDA and DEA models

This study analyses a pooled data of four credit unions⁵ using DEA model and the LDA model. Table 2 displays the total number of applications processed by different credit unions.

Table 2 Description of the credit union consumer loan data used in this study

<i>Class</i>	<i>Description</i>	<i>Number of applications</i>
1	Accept and were good loans	317
2	Accept and were bad loans	329
3	Reject	104
	Total	750

Notes: The credit unions included in our data base are: Jefferson County Teachers Credit Union, Jefferson County Employees Credit Union, Family Security Credit Union, and Steering Credit Union.

There are three groups of applicants: applicants who were accepted and were good credits (group 1); applicants who were accepted, but were not good credits (group 2); and applicants who applied for a loan, but were rejected (group 3). Further, the dataset also includes information such as the applicant's age, housing, address time, total income,⁶ number of credit cards, number of dependents, job time, co-maker on other loans, total debt, monthly rent/mortgage payments and total payments.⁷ The credit unions use an algorithm to calculate different types of credit rating and a final rating, ranging from 1–4, with loan applicants divided into four credit groups – excellent (1), good (2), marginal (3), and poor (4). However, to analyse loans without any bias, we decided to discard the credit ratings generated by the credit unions. Table 3 displays the summary statistics of the variables used in this study. The data parameters such as the total debt, number of outstanding loans, and total income vary widely.

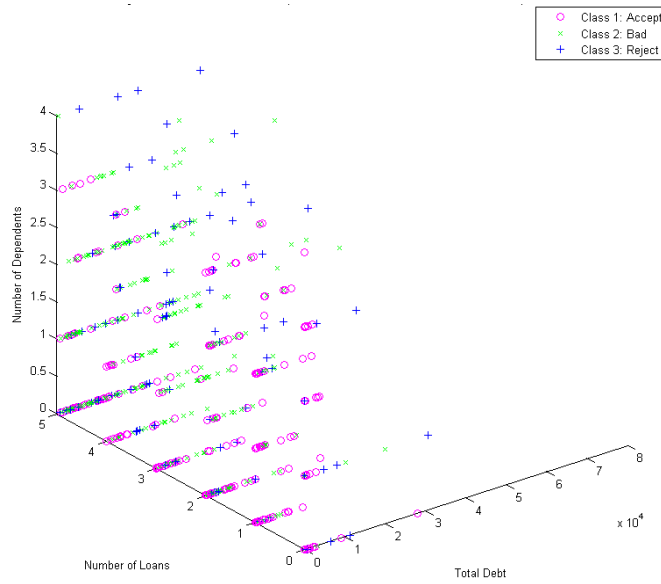
Table 3 Summary statistics of the variables used in this study to differentiate between good, bad, and outright reject loan applicants

<i>Statistics</i>	<i>Total debt</i>	<i>Number of loans</i>	<i>Number of dependents</i>	<i>Total payments</i>	<i>Total income</i>	<i>Job time (years)</i>
Mean	\$7,952.07	2.91	.97	\$606.99	\$2,160.53	8.5
Standard deviation	\$9,306.71	1.82	1.12	\$502.60	\$1,203.85	8.17
Minimum	\$0.00	0	0	\$0.00	\$0.00	0
Maximum	\$70,088	5	4	\$5,726	\$9,250	38
Median	\$4,565	3	1	\$511	\$1,900	6
Mode	\$0.00	5	0	\$0.00	\$1,900	1

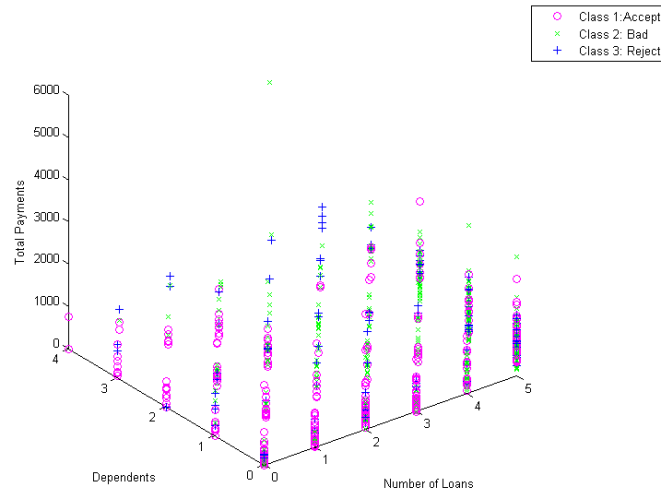
Notes: Total payments include payments for rent, automobile loan, and other payments. Total income includes gross income and other income.

The data set represents a cross section of information for 749 observations. There are three categories of applicants: applicants who were accepted and remained good, applicants who were originally accepted, but turned out to be bad credit, and applicants that were rejected. Clearly, the defaulting loan applicants should have been rejected by the loan officer in the initial screening. Figure 1 displays the plot of the data space showing the two categories.

Figure 1 Scatter plot of three input variables, (a) debt, number of loans, and dependents (b) number of loans, dependent and total payments (c) total payments, total income, and job time (see online version for colours)



(a)



(b)

Figure 1 Scatter plot of three input variables, (a) debt, number of loans, and dependents (b) number of loans, dependent and total payments (c) total payments, total income, and job time (continued) (see online version for colours)

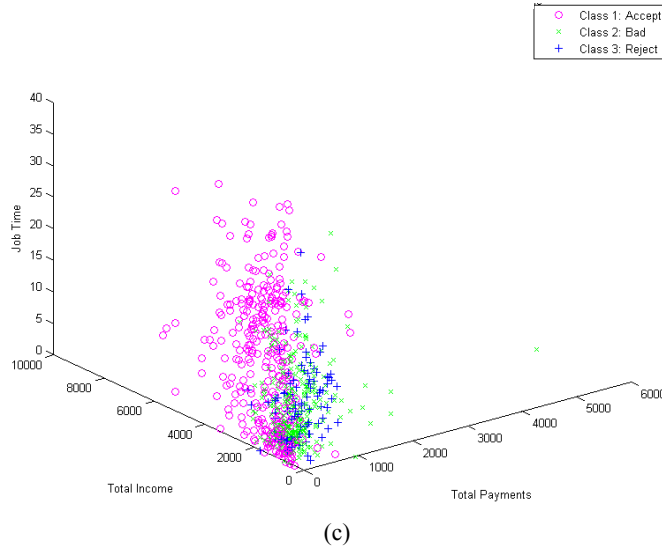


Figure 1(a) shows a scatter plot of the variables: total debt, number of outstanding loans, and total number of dependents. Figure 1(b) shows a scatter plot of the variables: number of outstanding loans, total number of dependents, and total payments. Figure 1(c) shows a scatter plot of the variables: total payments, total income, and time spent working (years). As illustrated in Figure 1(a), Figure 1(b), and Figure 1(c) the observations show overlapping clusters where each cluster includes data points from all classes. The three categories of the applicants do not form three separate clusters. As a result, the loan officer is unable to discriminate between good and bad loans. Therefore, the traditionally used, LDA model does not have a very high prediction rate with overlapping classes. The LDA model linearly divides the data space into three parts to correspond with the three classes, respectively. On the other hand, a DEA model uses the best policy loans as benchmark to compare the new loan applications. The next section illustrates the application of the discriminant analysis model and the DEA model (variable benchmark) to discriminate between good and bad loan applications. To illustrate the usefulness of the DEA model, we use the best case scenario for the LDA model that uses the entire data set for prediction. For the DEA model, we first identify 30 (100% efficient) loans using variable return to scale model. Further, we predict the efficiency of the 719 loan applications using variable benchmark DEA model that uses the 30 best policy loan applications as benchmarks.

Section 3 describes the computational details of the DEA model. In addition, to successfully apply DEA models, we should also consider many non-computational aspects to develop the specifications of the DEA model. The DEA specifications also include factors pertaining to selection of an extensive set of DMUs, suitable size of inputs and outputs, the most applicable DEA model (e.g., CRS, VRS, etc.), and an appropriate sensitivity analysis procedure for a given application (Ramanathan, 2003). There are no well-defined strategies for developing DEA model specifications due to DEAs

non-parametric nature. However, the results of DEA analysis depend on selection of appropriate DEA model. To begin with, we should consider two main factors: homogeneity and number of DMUs for DEA analysis. The applications of DEA methodology depends on using homogenous units that perform similar tasks and accomplish similar objectives. In our study, the loans are homogenous as they compete with each other to get sanctioned. Furthermore, the number of DMUs is also an important consideration. The number of DMUs should be reasonable so as to identify over-performing units, and clearly identify the relation between inputs and outputs. Many researchers provide some simple rules of thumb that guide the selection of inputs and outputs, and the number of participating DMUs.⁸

To study the outcome of a loan application, we consider six factors: total debt, number of outstanding loans, total number of dependents, total payments, total income, and time spent working (years). Out of these six factors, we specified total debt, number of outstanding loans, total number of dependents, and total payments as input, because the lower values of these parameters are an indicator of a good loan. On the other hand, higher total income and longer time spent working are considered as output factors as a higher value of these variables increases the credit-worthiness of a loan. Finally, the using the most appropriate DEA model for analysis is also an important consideration. We should select the DEA model with options such as input-maximising or output-minimising, multiplier or envelopment, and constant or variable returns to scale. DEA applications with inflexible inputs or not fully under control inputs should use output-based formulations. On the contrary, an application with flexible inputs and outputs that are an outcome of managerial goals, input-based DEA formulations are more appropriate. Furthermore, for an application that emphasises both inputs and outputs, we should use multiplier version. Similarly, for an application that considers relationship among DMUs, envelopment models are more suitable. Furthermore, the main factors involved in the DEA application dictate the use of constant or variable returns to scale. If the performance of DMUs depends heavily on the scale of operation, CRS is more applicable, otherwise variable returns to scale is a more appropriate assumption. In our study, the relationship among loans is an important consideration as all loans are competing with each other to be approved. Therefore, we select the envelopment models for our analysis. In addition, the focus of the credit union is to sanction loans that have lower input factors such as total debt, number of outstanding loans, total number of dependents, and total payments. Therefore, input-based formulation is recommended for our study. Furthermore, the credit-worthiness of these loans does not depend on the scale of operations, thus variable returns to scale is a safe assumption. Also, the structure of the DEA model (in envelopment form) uses an equation and separate calculation for every input and output. Therefore, all the input and output variables can be used simultaneously and measured in their own units. Further, as illustrated in Section 3, we use the variable-benchmark model to retain the best-performing loans on the efficiency frontier. The next section illustrates the empirical analysis of the application of the discriminant analysis model and the DEA model (variable benchmark) to discriminate between good and bad loan applications.

5 Empirical analysis

5.1 The linear discriminant analysis model

We use the *Class-dependent transformation* approach to the LDA model. This approach involves maximising the ratio of between class variance to within class variance. The main objective is to maximise this ratio to obtain adequate class separability. The class-specific type approach involves using multiple optimising criteria for transforming the data sets independently. Typically, the class-dependent LDA model aims for good discrimination between classes that coincides with the objective of the loan officer. The group variable takes three values:

- 1 the applicant should be accepted
- 2 the applicant that eventually default
- 3 the applicant should be rejected.

The predictor variables are: total debt, number of outstanding loans, total number of dependents, total payments, total income, and time spent working (years). For a classification problem having three distinct groups, space the maximum rank of the discriminant space is two. Thus, there are two canonical discriminant functions: function 1 and function 2. Figure 2 shows the scatter plot of the discriminant scores for the three groups using the unstandardised discriminant coefficients. Typically, the graph illustrates the group separation discerned by the two linear discriminant functions. The group centroid, the mean discriminant score in the discriminant space, visually illustrates the group separation in two-dimensional space. As evident in Figure 2, the three groups are not separable, as indicated by Figure 1(a), Figure 1(b), and Figure 1(c). Further, to test the performance of the discriminant model, we used the leave-one-out (LOO) procedure. The LOO estimator uses all the available data for training except one sample as the test set. Table 4 displays the results of the LDA model to classify the loan applications into the three categories with or without LOO estimator. Out of 317 good loans, LDA correctly classifies 264 loans (83.3%), without LOO estimator, and 262 loans (82.60%), with LOO estimator, in the accept-good-loans category. Likewise, out of 329 bad loans, LDA identifies 176 (53.5%), without LOO estimator, and 172 (52.3%), with LOO estimator, in the accept-bad-loans category. Similarly, out of 103 loans rejected, LDA identifies 51 (49.5%) in the reject category, and 48 (46.5%) with LOO estimator in this category. To compare the two models: LDA and variable-benchmark DEA models, we use the combined two categories (classes 2 and 3). Thus, using this classification, LDA rejects 375 (86.8%) out of 432 applications that should be rejected.

Figure 2 Discriminant score plot of the three-group sample on the two linear discriminant functions (function 1 and function2) illustrating the centroid of the three groups (1: accept, 2: accept turned bad, 3: reject) (see online version for colours)

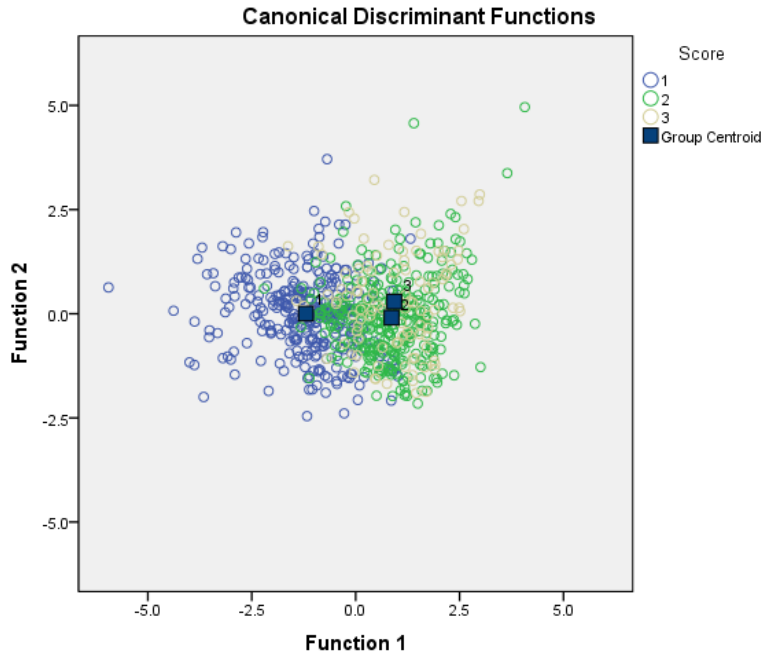


Table 4 Classification table reporting the number of cases for each group assigned using the original classification rule^a and the LOO estimator^b (see online version for colours)

Original group	Predicted groups				Total	Classification
	1 – Accept	2 – Bad	3 – Reject	2/3 – Bad/reject combined		
1 – Accept	264	37	16	53	317	a
	83.3%	11.7%	5.0%	16.70%	100.0%	a
2 – Bad	262	37	18	55	317	b
	82.6%	11.7%	5.7%	17.40%	100.0	b
	43	176	110	286	329	a
	13.1%	53.5%	33.4%	86.90%	100.0%	a
3 – Reject	44	172	113	285	329	b
	13.4%	52.3%	34.3%	86.60%	100.0%	b
	13	39	51	90	103	a
3 – Reject	12.6%	37.9%	49.5%	87.40%	100.0	a
	13	42	48	90	103	b
	12.6%	40.8%	46.6%	87.40%	100.0	b

Notes: a – Original grouped cases.
 b – Cross-validation grouped cases.

5.2 Variable benchmark DEA model

In this study, we use the variable benchmark data envelopment model. As illustrated in Section 3 and Section 5, a variable benchmark DEA model uses the best performing loans on the efficiency frontier. Further, the efficiency frontier does not change when a new loan outperforms the identified frontier. The envelopment DEA model modifies the efficiency frontier as new loans are presented, thereby the benchmark changes for the new incoming loans. In this study, we identified 30 best performing loans out of 749 loans. All of these loans are 100% efficient. We use Excel Solver and Visual Basic Application to solve the DEA model. The limitation of Solver for our application was 199 data points. Thus, we solved the DEA model using sets of 199 loan applications. The 30 most outperforming loans are 100% efficient compared to all the other loan applications. Table 5 illustrates the results of the input-minimising and output-maximising variable benchmark DEA models. We used these two models to compare the 719 loans with the 30 best practices loans. The DEA model uses a scale of 0%–100% (quality index of a loan) to relatively benchmark the loans with the best practices loans. We use 75% quality index to classify good loans, between 65% and 45% to classify potentially bad loans, and less than 45% to classify outright reject loans. The loan officer can determine the cut-off values of the quality index depending on the ability of the lending institution to undertake risk. Out of 287 good loans, the Variable Benchmark DEA model correctly classifies 257 loans (89.55%), using output-maximising model, and 249 loans (86.76%), using the input-maximising model, in the accept-good-loans category. Including the 30 best-practice loans, correctly identified by DEA, out of 317 good loans, the variable benchmark DEA model correctly classifies 287 loans (90.54%), using output-maximising model, and 279 loans (88.01%), using the input-maximising model, in the accept-good-loans category. Likewise, out of 329 bad loans, the variable benchmark DEA model rejects 304 loans (92.4%), using output-maximising model, and 307 loans (93.31%), using the input-maximising model, in the accept-bad-loans category. Similarly, out of 103 loans rejected, the Variable Benchmark DEA model rejects 95 loans (92.23%), using output-maximising model, and 97 loans (94.17%), using the input-maximising model, in this category. We use the above-specified assumptions on the quality of loan to classify in category 2 (high risk) and category 3 (outright reject) to assess the riskiness of the loans. For the bad loans category, the variable benchmark DEA model places 120 loans (36.47%), using output-maximising model, and 57 loans (17.33%), using the input-maximising model, in the high-risk-bad-loans category and 184 loans (55.93%), using output-maximising model, and 250 loans (75.99%), using the input-maximising model, in the outright-reject-loans category. For class 3 (reject loans), the variable benchmark DEA model places 35 loans (33.98%), using output-maximising model, and 18 loans (17.48%), using the input-maximising model, in the high-risk-bad-loans category and 60 loans (58.25%), using output-maximising model, and 79 loans (76.70%), using the input-maximising model, in the outright-reject-loans category. The input-minimising and output maximising models give different results as their goals are different. Thus, to compare the two models, we consider bad loans as reject and combine the two categories (classes 2 and 3). Thus, using this classification, out of 432 applications, DEA rejects 399 applications (92.36%) with the output-maximising model and 404 applications (93.52%) with the input-minimising model. Further, to compare the two models, we performed the Paired-wise student's t-test with 0.05 tolerance level. Table 6 displays the results of

pairwise statistical analysis with classes 2 and 3 treated separate, and Table 7 displays the results of the test with classes 2 and 3 treated as reject class. As is evident from Table 6, the input-maximising variable benchmark DEA model is statistically different from the LDA model for all categories, while the output-maximising model is not statistically different for bad loans and overall category. With classes 2 and 3 combined, the input-maximising variable benchmark DEA model is statistically different from the LDA model for all categories except the overall category. The output-maximising model is not statistically different for all categories. Thus, overall the input-maximising model statistically outperforms the LDA model.

The next step is to perform the case analysis of loans using the input-minimising and output-maximising efficiency scores. As illustrated in Section 3, based on the efficiency score of the input-minimising and output-maximising models a loan can be classified in one of the four cases. Table 8 illustrates the case analysis of 719 loans when compared to the 30 best-practices loans. The following is the case-wise breakup of the loans for good cases: total debt, number of outstanding loans, total number of dependents, total payments, total income, and time spent working (years).

- Case 1 Out of 287 given loan applications, 54 applications are of better quality than the best practices loans. These applications have the smallest level of total debt, number of outstanding loans, total number of dependents, total payments and the largest level of total income, and time spent working (years) as compared to the best-practices benchmark. These loan applications have lower input parameters and higher output factors, and therefore making infeasible both the input-minimising and output-maximising models. We can define the $\delta_o^{VRS^*} - 1$ or $\tau_o^{VRS^*} - 1$ as the performance gap between DMU_o^{New} and the benchmark as characterised by input-minimising and output-maximising models.
- Case 2 Twenty one accepted loan applications have largest level of total income, and time spent working (years) as compared to the best-practices benchmark to make the input-minimising model infeasible. Thus, we use model (6) to calculate the output surplus offered by DMU_o^{New} .
- Case 3 Forty nine accepted loan applications have lowest level of total debt, number of outstanding loans, total number of dependents, and total payments as compared to the best-practices benchmark to make the output-maximising model infeasible. Thus, we use model (5) to calculate the input savings offered by DMU_o^{New} .
- Case 4 Sixty eight accepted loans have the smallest level of total debt, number of outstanding loans, total number of dependents, total payments and the largest level of total income, and time spent working (years) as compared to the best-practices benchmark. As both models (5) and (6) are feasible, this implies that we use both the models to determine if DMU_o^{New} offers input savings and output surpluses.
- Case 5 Ninety four accepted loans underperform as compared to the best-practices benchmark. We can use benchmark values from models (5) or (6) to find the source of low performance.

Table 5 Classification table reporting the number of cases for each group assigned using the original VB-Max DEA^c and the VB-Min DEA^d (see online version for colours)

Original group	Predicted groups					Total	Classification
	1 – Accept	2 – Bad	3 – Reject	1 ^e – Accept	2/3 – Bad/reject combined		
1 – Accept	257 89.55%	24 8.36%	5 1.74%	287 90.54%	29 10.10%	287 + 30 = 317 ^e 100.00%	c
2 – Bad	249 86.76%	29 10.10%	9 3.14%	279 88.01%	38 13.24%	287 + 30 = 317 ^e 100.0	d
3 – Reject	24 7.29%	120 36.47%	184 55.93%	2/3 – Bad/reject combined	304 92.40%	329 100.0%	c
	23 6.99%	57 17.33%	250 75.99%	(Out of total – 432)	307 93.31%	329 100.0%	d
	6 5.83%	35 33.98%	60 58.25%	399 92.36%	95 92.23%	103 100.0	c
	6 5.83%	18 17.48%	79 76.70%	404 93.52%	97 94.17%	103 100.0	d

Notes: c – Variable benchmark output maximising DEA model.

d – Variable benchmark input minimising DEA model.

e – Includes the 30 best practices loans.

Table 6 Paired-wise student's t-test of LDA model and variable benchmark DEA models

		<i>Original</i>		<i>Cross-validation</i>	
		<i>One-tailed</i>	<i>Two-tailed</i>	<i>One-tailed</i>	<i>Two-tailed</i>
Accept	DEA-Max	0.0187	0.0374	0.0108	0.0217
	DEA-Min	0.0061	0.0121	0.0032	0.0065
Bad	DEA-Max	0.0000	0.0001	0.0001	0.0002
	DEA-Min	0.0000	0.0000	0.0000	0.0000
Reject	DEA-Max	<i>0.4468</i>	<i>0.8935</i>	0.3040	<i>0.6080</i>
	DEA-Min	0.0000	0.0000	0.0000	0.0000
Bad/reject	DEA-Max	0.0003	0.0006	0.0002	0.0005
	DEA-Min	0.0000	0.0000	0.0000	0.0000
All	DEA-Max	<i>0.0566</i>	<i>0.1132</i>	<i>0.0725</i>	<i>0.1450</i>
	DEA-Min	0.0000	0.0000	0.0000	0.0000

Table 7 Paired-wise student's t-test of LDA model and variable benchmark DEA models with bad/reject categories combined as a single category

		<i>Original</i>		<i>Cross-validation</i>	
		<i>One-tailed</i>	<i>Two-tailed</i>	<i>One-tailed</i>	<i>Two-tailed</i>
Accept	DEA-Max	0.0121	0.0241	0.0089	0.0179
	DEA-Min	0.0196	0.0392	0.0146	0.0292
Bad	DEA-Max	<i>0.5000</i>	<i>1.0000</i>	<i>0.4311</i>	<i>0.8621</i>
	DEA-Min	0.0002	0.0004	0.0001	0.0002
Reject	DEA-Max	<i>0.5000</i>	<i>1.0000</i>	<i>0.3823</i>	<i>0.7647</i>
	DEA-Min	0.0094	0.0189	0.0054	0.0107
Bad/reject	DEA-Max	<i>0.5000</i>	<i>1.0000</i>	<i>0.3817</i>	<i>0.7634</i>
	DEA-Min	0.0000	0.0000	0.0000	0.0000
All	DEA-Max	0.0454	<i>0.0907</i>	<i>0.0566</i>	<i>0.1132</i>
	DEA-Min	<i>0.1080</i>	<i>0.2161</i>	<i>0.0935</i>	<i>0.1870</i>

Table 8 Classification results using variable benchmark DEA model

<i>Loans</i>	<i>Model</i>	<i>Percentage of classification</i>	<i>Case 1</i>	<i>Case 2</i>	<i>Case 3</i>	<i>Case 4</i>	<i>Case 5</i>	<i>Benchmark</i>	<i>Total</i>
Good	Max	90%	54	21	49	68	94	1	287
	Min	87%							
Bad	Max	92%	16		2		311		329
	Min	93%							
Reject	Max	92%	3	2			98		103
	Min	94%							

Finally, one accepted loan is on the best-practices benchmark, indicating that the loan has performance comparable to the best-practices benchmark. Similarly, we can analyse the bad and reject loans categories. As expected, 95% of loans in these categories are underperforming, further validating the adequacy of best-practices benchmark.

6 Summary and conclusions

Besides intuitive judgement and experience, loan officers use mathematical modelling techniques such as statistical analysis and credit scoring models to differentiate good loans from bad loans. However, in more recent years, loan officers also use other artificial intelligence tools such as expert systems, neural networks, fuzzy logic, and genetic algorithms. DEA is a relatively new technique to capture the attention of the finance community. This study illustrates the use of the variable benchmark DEA model to assess the credit-worthiness of a potential loan. Further, we also compare the performance of the DEA model with LDA model. In addition, to aid the loan officer to make an informed decision with respect to a given loan, we evaluate and benchmark a loan against a set of best practice loans.

To analyse the efficiency of the two models, we used a pooled data of nine credit unions. An analysis of the data space plot indicates that the good loans and bad loans do not form two distinct clusters. Instead, the two classes form highly overlapped clusters. We use a set of 719 observations to implement the LDA and DEA models. We find that the DEA model outperforms the LDA model, especially in screening out potential defaulters. Thus, we can use DEA as a useful analytical modelling tool to screen consumer loan applications. Further, we illustrated the use of two DEA models: input-minimising variable benchmark DEA model and output-maximising variable benchmark to perform a case wise analysis of a new loan application as opposed to benchmark of best practices loan. This analysis allows a loan officer to examine all the possible scenarios, and gain an understanding of the potential risk involved in approving a less than perfect loan application. Thus, DEA can serve as an effective method for performance evaluation and benchmarking against best-practices. The DEA model provides a composite efficiency score that is indicative of an 'overall quality measure' for a loan application. We also illustrate the use of variable benchmark DEA model in eliciting 'high performance/quality' loans that can aid loan officers to better benchmark a given loan using quality indicators with the given high-performing loans in the 'loan base'.

References

- Baesens, B., Van Gestel, T., Stepanova, M., Van de Poel, D. and Vanthienen, J. (2005) 'Neural network survival analysis for personal loan data', *The Journal of the Operational Research Society*, Vol. 56, No. 9, pp.1089–1098.
- Basso, A. and Funari, S. (2001). 'A data envelopment analysis approach to measure the mutual fund performance', *European Journal of Operational Research*, Vol. 135, No. 3, p.477, Amsterdam.
- Bergendahl, G. and Lindblom, T. (2008) 'Evaluating the performance of Swedish savings banks according to service efficiency', *European Journal of Operational Research*, Vol. 185, No. 3, pp.1663–1673.

- Charnes, A., Cooper, W.W. and Rhodes, E. (1978) 'Measuring the efficiency of decision making units', *European Journal of Operational Research*, Vol. 2, No. 6, p.429, Amsterdam.
- Chen, Y., Sun, L. and Peng, C. (2005) 'Commercial banks' performance in Taiwan', *International Journal of Performance Management*, Vol. 7, No. 4, p.444.
- Cowie, J. and Riddington, G. (1996) 'Measuring the efficiency of European railways', *Applied Economics*, Vol. 28, No. 8, pp.1027–1035.
- Darrat, A.F., Topuz, C. and Yousef, T. (2003) 'Assessing bank efficiency in an emerging market: the Kuwaiti experience in the 1990s', *Studies in Economics and Finance*, Vol. 21, No. 2, pp.1–21.
- Desai, V.S., Crook, J.N. and Overstreet Jr., G.A. (1996) 'A comparison of neural networks and linear scoring models in the credit union environment', *European Journal of Operational Research*, Vol. 95, No. 1, p.24.
- Edirisinghe, N.C.P. and Zhang, X. (2007) 'Generalized DEA model of fundamental analysis and its application to portfolio optimization', *Journal of Banking & Finance*, Vol. 31, No. 11, pp.3311–3335.
- Feroz, E.H., Kim, S. and Raab, R.L. (2003) 'Financial statement analysis: a data envelopment analysis approach', *Journal of the Operational Research Society*, Vol. 54, No. 1, pp.48–58.
- Fisher, R.A. (1936) 'The use of multiple measurements in taxonomic problems', *Annals of Eugenics*, Vol. 7, No. 2, pp.179–188, doi: 10.1111/j.1469-1809.1936.tb02137.x.hdl:2440/15227.
- Fritz, S. and Hosemann, D. (2000) 'Restructuring the credit process: behaviour scoring for German corporates', *International Journal of Intelligent Systems in Accounting, Finance and Management*, Vol. 9, No. 1, pp.9–21.
- Galagedera, D.U.A. and Silvapulle, P. (2002) 'Australian mutual fund performance appraisal using data envelopment analysis', *Managerial Finance*, Vol. 28, No. 9, p.60, Patrinton.
- Halkos, G. and Salamouris, D. (2004) 'Efficiency measurement of the Greek commercial banks with the use of financial ratios: a data envelope analysis approach', *Management Accounting Research*, Vol. 15, No. 2, p.201.
- Haslem, J.H. and Scheraga, C.A. (2003) 'Data envelopment analysis of Morningstar's large-cap mutual funds', *Journal of Investing*, Vol. 12, No. 4, p.41, New York.
- Haslem, J.H., Scheraga, C.A. and Bedingfield, J.P. (1999) 'DEA efficiency profiles of US banks operating internationally', *International Review of Economics & Finance*, Vol. 8, No. 2, p.165, Greenwich.
- Hoon, O. and Chunyan, Y. (1994) 'Economic efficiency of railways and implications for public policy', *Journal of Transport Economics and Policy*, Vol. 28, No. 2, pp.121–139.
- Howland, M. and Rowse, J. (2006) 'Measuring bank branch efficiency using data envelopment analysis: managerial and implementation issues', *INFOR*, Vol. 44, No. 1, pp.49–63.
- Kao, C. and Liu, S. (2004) 'Predicting bank performance with financial forecasts: a case of Taiwan commercial banks', *Journal of Banking & Finance*, Vol. 28, No. 10, p.2353.
- Lin, H., Hsu, G.J. and Hsiao, C. (2007) 'Measuring efficiency of domestic banks in Taiwan: application of data envelopment analysis and Malmquist index', *Applied Economics Letters*, Vol. 14, No. 11, pp.821–827.
- Liu, J.S. and Liu, W. (2010). 'DEA and ranking with the network-based approach: a case of R&D performance', *Omega*, December, Vol. 38, No. 6, p.453, Oxford.
- Lozano, S. and Gutiérrez, E. (2011) 'Slacks-based measure of efficiency of airports with airplanes delays as undesirable outputs', *Computers & Operations Research*, January, Vol. 38, No. 1, p.131, New York.
- Malhotra, R. and Malhotra, D.K. (2003) 'Differentiating between good credits and bad credits using neuro-fuzzy systems', *European Journal of Operational Research*, Vol. 136, No. 1, pp.190–231.

- Martens, B.B., Van Gestel, T. and Vanthienen, J. (2007) 'Comprehensible credit scoring models using rule extraction from support vector machines', *Eur. J. Oper. Res.*, Vol. 183, No. 3, pp.1466–1476.
- McLachlan, G.J. (2004) *Discriminant Analysis and Statistical Pattern Recognition*, ISBN 0471691151, MR1190469, Wiley Interscience.
- McMullen, P.R. and Strong, R.A. (1998) 'Selection of mutual funds using data envelopment analysis', *The Journal of Business and Economic Studies*, Vol. 4, No. 1, p.1, Oakdale.
- Morey, M. and Morey, R.C. (1999) 'Mutual fund performance appraisals: a multi-horizon perspective with endogenous benchmarking', *Omega*, Vol. 27, No. 2, p.241, Oxford.
- Murthi, B.P.S., Choi, Y.K. and Desai, P. (1997) 'Efficiency of mutual funds and portfolio performance measurement: a non-parametric approach', *European Journal of Operational Research*, Vol. 98, No. 2, p.408, Amsterdam.
- Neal, P. (2004) 'X-efficiency and productivity change in Australian banking', *Australian Economic Papers*, Vol. 43, No. 2, pp.174–191.
- Ozcan, Y.A. and McCue, M.J. (1996) 'Development of a financial performance index for hospitals: DEA approach', *The Journal of the Operational Research Society*, Vol. 47, No. 1, pp.18–26.
- Paradi, J. and Schaffnit, C. (2004) 'Commercial branch performance evaluation and results communication in a Canadian bank – a DEA application', *European Journal of Operational Research*, Vol. 156, No. 3, pp.719–734.
- Pille, P. and Paradi, J. (2002) 'Financial performance analysis of Ontario (Canada) credit unions: an application of DEA in the regulatory environment', *European Journal of Operational Research*, Vol. 139, No. 2, pp.339–350.
- Ramanathan, R. (2003) *An Introduction to Data Envelopment Analysis – A Tool for Performance Measurement*, Sage Publications, New Delhi, India.
- Sanjeev, G. (2007) 'Does bank's size matter in India?', *Journal of Financial Services Research*, Vol. 6, No. 2, pp.135–144.
- Saranga, H. and Moser, R. (2010) 'Performance evaluation of purchasing and supply management using value chain DEA approach', *European Journal of Operational Research*, 16 November, Vol. 207, No. 1, p.197, Amsterdam.
- Sedzro, K. and Sardano, D. (1999) *Mutual Funds Performance Evaluation Using Data Envelopment Analysis*, Working paper, School of Business, University of Quebec, Montreal, Canada.
- Sufian, F. (2007) 'Trends in the efficiency of Singapore's commercial banking groups A non-stochastic frontier DEA window analysis approach', *International Journal of Productivity and Performance Management*, Vol. 56, No. 2, pp.99–135.
- Yu, M. and Lin, E. (2008) 'Efficiency and effectiveness in railway performance using a multi-activity network DEA model', *Omega*, Vol. 36, No. 6, pp.1005–1017.
- Zhu, J. (2000) 'Multi-factor performance measure model with an application to Fortune 500 companies', *European Journal of Operational Research*, Vol. 123, No. 1, pp.105–124.
- Zhu, J. (2003) *Quantitative Models for Performance Evaluation and Benchmarking*, p.13, Kluwer's International Series.

Notes

1 For more details on LDA, please refer to

- Fisher (1936)
- McLachlan (2004).

- 2 There are multiple versions of the DEA model. For more details on the two-stage DEA model refer to Zhu (2003).
- 3 The original Charnes, Cooper, and Rhodes (CCR) model is CRS. It is built on the assumption that if an activity (x, y) is feasible, then, for every positive scalar t , the activity (tx, ty) is also feasible. Thus, we can scale the inputs and outputs linearly without increasing or decreasing efficiency. Non-decreasing returns to scale implies that changing all inputs by the same proportion changes the output by a greater extent than the proportional value. Similarly, in non-increasing returns to scale scaling up is interdicted, while scaling down is permitted. Variable returns to scale combine these three possibilities – NDRS, CRS, and NIRS for different ranges of output.
- 4 For more details on the proof of propositions for variable benchmark DEA model refer to Zhu (2003).
- 5 The credit unions included in our database are: Jefferson County Teachers Credit Union, Jefferson County Employees Credit Union, Family Security Credit Union, and Steering Credit Union.
- 6 Total income includes gross income and other income.
- 7 Total payments include payments for rent, automobile loan, and other payments.
- 8 The following are the guidelines for DMU model selection:
 - a The number of DMUs is expected to be larger than the product of number of inputs and outputs (Darrat et al., 2003) to discriminate effectively between efficient and inefficient DMUs. The sample size should be at least two or three times larger than the sum of the number of inputs and outputs (Ramanathan, 2003).
 - b The criteria for selection of inputs and outputs are also quite subjective. A DEA study should start with an exhaustive, mutual list of inputs and outputs that are considered relevant for the study. Screening inputs and outputs can be quite quantitative (e.g., statistical) or qualitative that are simply judgemental, use expert advice, or use methods such as analytical hierarchy process. Typically, inputs are the resources utilised by the DMUs or condition affecting the performance of DMUs. On the other hand, outputs are the benefits generated as a result of the operation of the DMUs, and records higher performance in terms of efficiency. Typically, we should restrict the total number of inputs and outputs to a reasonable level. As the number of inputs and outputs to a reasonable level. As the number of inputs and outputs increases, more number of DMUs get an efficiency rate of 1, as they become too specialised to be evaluated with respect to other units (Ramanathan, 2003).