

## **A Comparative Analysis of Two-Stage Distress Prediction Models**

Mohammad Mahdi Mousavi  
University of Bradford School of Management, UK  
[m.mousavi@bradford.ac.uk](mailto:m.mousavi@bradford.ac.uk)

Jamal Ouenniche  
University of Edinburgh, Business School, UK  
[jamal.ouenniche@ed.ac.uk](mailto:jamal.ouenniche@ed.ac.uk)

Kaoru Tone  
National Graduate Institute for Policy Studies, JAPAN  
[tone@grips.ac.jp](mailto:tone@grips.ac.jp)

## **A Comparative Analysis of Two-Stage Distress Prediction Models**

**Abstract:** On feature selection, as one of the critical steps to develop a distress prediction model (DPM), a variety of expert systems and machine learning approaches have analytically supported developers. Data envelopment analysis (DEA) has provided this support by estimating the novel feature of managerial efficiency, which has frequently been used in recent two-stage DPMs. As key contributions, this study extends the application of expert system in credit scoring and distress prediction through applying diverse DEA models to compute corporate market efficiency in addition to the prevailing managerial efficiency, and to estimate the decomposed measure of mix efficiency and investigate its contribution compared to Pure Technical Efficiency and Scale Efficiency in the performance of DPMs. Further, this paper provides a comprehensive comparison between two-stage DPMs through estimating a variety of DEA efficiency measures in the first stage and employing static and dynamic classifiers in the second stage. Based on experimental results, guidelines are provided to help practitioners develop two-stage DPMs; to be more specific, guidelines are provided to assist with the choice of the proper DEA models to use in the first stage, and the choice of the best corporate efficiency measures and classifiers to use in the second stage.

**Keywords:** Corporate Two-stage Distress Prediction; Efficiency; Data Envelopment Analysis; Malmquist Index

## 1. Introduction

Recent studies on corporate failure prediction aim to use broadly recognised sources and indicators of financial distress such as difficulties in operating and financing activities, and poor performance in management and leadership of the company in developing an early distress warning system to take proper preventive action against bankruptcy and immune the firm (see, for example, Altman, Iwanicz-Drozdowska, Laitinen, & Suvas, 2017; Baudraerts, 2016; Bauer & Agarwal, 2014; Laitinen & Suvas, 2016; Liang, Lu, Tsai, & Shih, 2016; Wu, Gaunt, & Gray, 2010; Yeh, Chi, & Hsu, 2010).

According to Zhou (2013), distress and bankruptcy prediction models are data-fitting based empirical research consisting of four steps: sampling, features selection, choice or design of classifier, and performance evaluation. Regarding the design of classifiers, related studies have used different techniques from a variety of fields such as statistics and probability, machine learning and expert systems. Statistical models, however, are based on assumptions which might not be valid for a given data set; e.g., linearity, multivariate normality, independence among predictor or input variables, and equal within-group variance-covariate matrices. The less vulnerable classifiers to the underlying statistical assumptions are the ones from the field of artificially intelligent and expert systems (AIES) such as recursively partitioned decision trees (e.g., Frydman, Altman, & Kao, 1985), case-based reasoning models (e.g., H. Li & Sun, 2009, 2011), neural networks (e.g., Du Jardin & Séverin, 2012; Kim & Kang, 2010), rough set theory (e.g., McKee & Lensberg, 2002; Yeh et al., 2010), genetic programming (e.g., Back, Laitinen, Sere, & Wezel, 1995; Alfaro-Cid, Sharman, & Esparcia-Alcazar, 2007; Etemadi, Anvary Rostamy, & Dehkordi, 2009), as well as the ones from field of operations research (OR), such as multi-criteria decision making analysis (MCDA) (e.g., Zopounidis & Doumpos, 2002) and Data Envelopment Analysis (DEA) (e.g., Sueyoshi & Goto, 2009; Sueyoshi, Goto, & Omi, 2010; Z. Li, Crook, & Andreeva, 2014, Ouenniche & Tone, 2017) – for a detailed classification of failure prediction models, the reader is referred to Balcaen and Ooghe (2006), Aziz and Dar (2006), Bellovary, Giacomino, & Akers (2017), Baharammirzaee (2010), Abdou and Pointon (2011), and Chen, Ribeiro, & Chen (2016).

In addition to the application of expert systems and machine learning as classifiers of distressed and healthy firms, a variety of these techniques have been applied in the feature selection stage

(Chandrashekar & Sahin, 2014; Guenther & Schonlau, 2016; Jain, 1997; Sun et al., 2013; Tang & Chi, 2005; Wang, Ma, & Yang, 2014). The prevailing features used in DPMs are accounting, market, and macroeconomic indicators (see, for example, Xu & Wang, 2009; Yeh et al., 2010; Z. Li et al., 2014; Z. Li, Crook, & Andreeva, 2017), yet, recent studies commonly acknowledged the explanatory power of management performance in corporate distress (Seballos & Thomson, 1990; Gestel et al., 2006; Yeh et al., 2010). In practice, newly developed DPMs have incorporated business efficiency, i.e. the ratio of weighted outputs (e.g., sales, profit, and net income) to weighted inputs (e.g., equity, asset, and employees), as an effective reflection of corporate management performance.

Direct estimation of a company's efficiency (i.e., technical and productivity efficiency) using financial statements is problematic. Although, most machine learning techniques and expert systems, e.g. neural networks, cannot provide efficiency estimations (Z. Li et al., 2017), DEA as a type of machine learning technique (Cielen, Peeters, & Vanhoof, 2004) can incorporate multiple inputs and outputs to estimate a measure of the relative efficiency of a company. The estimated DEA efficiency measures have been applied both directly for classification to identify distressed and healthy companies (e.g., Cielen et al., 2004; Paradi, Asmild, & Simak, 2004; Sueyoshi, 2006; Premachandra, Chen, & Watson, 2011; Ouenniche & Tone, 2017), or indirectly as features in developing DPMs (Psillaki, Tsolas, & Margaritis, 2010; Xu & Wang, 2009; Sueyoshi et al., 2010; Yeh et al., 2010; Z. Li et al., 2014, 2017). Note that the latter application of DEA efficiency measures in the literature of bankruptcy and distress prediction is called two-stage modelling, whereby a DEA model estimates a corporate efficiency measure in the first stage; then, the estimated measure is retained as a feature to develop DPM in the second stage. In practice, however, there are no guidelines regarding the choice of the proper DEA model (i.e. static vs. dynamic), type of returns-to-scale (i.e. constant returns-to-scale (CRS) vs. variable returns-to-scale (VRS)), orientation of DEA analysis (i.e. input-oriented, output-oriented, or non-oriented), and appropriate DEA inputs and outputs to be employed in the first stage, as well as the type of efficiency measures and the classifier to be used in the second stage.

Our survey of the literature reveals several gaps in the literature on two-stage distress prediction modelling. First, to the best of our knowledge, no study provides a comprehensive comparison between two-stage DPMs; neither considering different DEA models that are used to estimate company efficiency in the first stage nor using different classifiers in the second stage. Second,

the literature on employing DEA efficiency measures as new features in model building (see, for example, Z. Li et al., 2014) is limited to employing Pure Technical Efficiency (PTE)<sup>1</sup> and Scale Efficiency (SE)<sup>2</sup>, thus ignoring Mix Efficiency (ME) or equivalently input and output slacks. Third, our survey indicates that the choice of DEA inputs and outputs to estimate company's efficiency measures is restricted to accounting variables (Barr & Siems, 1997; Z. Li et al., 2014, 2017; Psillaki et al., 2010; Xu & Wang, 2009; Yeh et al., 2010) thus ignoring market variables (see, Table 1). There are several issues or criticisms regarding the use of accounting-based information, e.g., accounting-based information only present a firm's historical performance and may not be informative in predicting the future; the "true" asset values may be very different from the book values; and accounting numbers can be manipulated by Management (Agarwal & Taffler, 2008; Balcaen & Ooghe, 2006; Hernandez Tinoco & Wilson, 2013; Trujillo-Ponce, Samaniego-Medina, & Cardone-Riportella, 2014). To overcome these drawbacks, Shumway (2001) as the pioneer study used firms' market value, past stock returns, and the idiosyncratic standard deviation of stock returns as market-driven variables to improve the performance of bankruptcy prediction models. The rationale behind the use of market-based variables is that, in an efficient market, stock prices will reflect both the information contained in accounting statements and the information contained in the future expected cash-flows. Furthermore, market variables are unlikely to be influenced by the firm's accounting policies.

This study adds to the current literature of two-stage DPMs in several respects. First, building on the work of Li et al. (2014) in analysing the contribution of decomposed scores of Technical efficiency (TE), i.e., PTE and SE, to predict financial distress, we suggest decomposing the non-radial technical efficiency score, i.e., Slack-Based Measure (SBM) of efficiency (Tone, 2011), into PTE, SE and ME, and investigating how each of these measures individually contributes to DPMs. Second, considering the frequent use of market-driven information as features of DPMs in recent studies (see, for example, Xu & Wang, 2009; Yeh et al., 2010; Z. Li et al., 2014, 2017) and inspired by the prevailing application of Sharpe ratio (1966, 1994) to examine the performance of an investment, we suggest using Shumway's market-driven information as input and output of DEA models to estimate the market efficiency performance of a firm (hereafter

---

<sup>1</sup> refers to the ability to improve the effectiveness by prudently allocating resources and using new technology

<sup>2</sup> indicates the capacity to attain better efficiency by adjusting to its optimal scale

called market efficiency measure). Third, as an empirical contribution, this study provides a comprehensive comparison of two-stage DPMs that apply different DEA models, say, input-oriented versus output-oriented, radial versus non-radial, static versus dynamic, to compute the measures of management efficiency and market efficiency of companies at the first stage of two-stage analyses and use static and dynamic classifiers at the second stage. In sum, this paper extends the application of expert systems and machine learning in feature selection of distress prediction modelling through employing diverse DEA models to estimate market and managerial efficiency of firms and investigating the effectiveness of decomposing DEA measures on the performance of DPMs.

The remainder of this paper is organised as follows. Section 2 reviews the literature on DEA in distress prediction. Section 3 describes the details of the experimental design including data, sampling, and hybrid two-stage models of distress prediction to be assessed and the proposed evaluation technique. Section 4 describes the empirical results and the findings. Finally, section 5 presents the conclusion of this study.

## **2. Literature review**

DEA is a non-parametric technique, which was introduced to measure the relative efficiency of a group of decision-making units (DMUs), e.g. firms, hospitals, products, prediction models, cities, and others, based on their respective inputs and outputs (Charnes, Cooper, & Rhodes, 1978). DEA has been one of the most successfully used techniques in the research activities related to performance evaluation of banking and other financial institutions – for a comprehensive survey on DEA in banking; the reader is referred to Emrouznejad and Yang (2018), Paradi and Zhu (2013) and Fethi and Pasiouras (2010). The rational association between the company's efficiency (as a proxy of management efficiency) and the probability of distress is commonly recognised in recent distress prediction studies (Seballos & Thomson, 1990; Gestel et al., 2006; Yeh et al., 2010).

More relevant to this research, distress prediction studies applied DEA in two different ways. First, DEA is used as a classifier to discriminate between distressed and healthy groups of firms (Cielen et al., 2004; Paradi et al., 2004; Sueyoshi, 2006; Premachandra et al., 2011; Ouenniche & Tone, 2017). Second, in hybrid two-stage prediction frameworks, DEA is used to measure the relative efficiency of companies at the first stage. Then, the estimated DEA efficiency score is

used as an input to or explanatory variable in the prediction model at the second stage (Xu & Wang, 2009; Sueyoshi et al., 2010; Psillaki et al., 2010; Yeh et al., 2010; Z. Li et al., 2014, 2017). The next two sections provide a concise review of the application of DEA in distress prediction as a classifier (see section 2.12.1) and as a predictor (see section 2.22.2). Table 2 provides a summary of these studies.

## 2.1. DEA as a Classifier

Compared to the conventional statistical models, DEA as a non-parametric classifier has some methodological advantages. For example, DEA is a distribution-free framework and does not require specifying the distribution of features. Also, DEA relaxes the assumption of equality of variance-covariance matrices among all groups. Further, it does not incorporate *a priori* probabilities to account for the relative occurrence of observations in different populations and does not require *a priori* specification of a functional form for the input-output relationship (Paradi et al., 2004; Premachandra, Bhabra, & Sueyoshi, 2009).

[Insert *Table 1* Here]

In the literature, the application of DEA scores for classification consists of using such scores to discriminate between two groups of *Good* (e.g. bankrupt) and *Bad* (e.g. non-bankrupt) entities or firms using a cut-off point or a statistical test (Simak, 1997; Pille & Paradi, 2002; Tsai, Lin, Cheng, & Lin, 2009; Shetty, Pakkala, & Mallikarjunappa, 2012; Paradi et al., 2004; Cielen et al., 2004; Ouenniche & Tone, 2017) – see Table 2 for more details.

As the pioneer study, Barr et al. (1993) used CCR model (Charnes, Cooper and Rhodes, 1978) under the constant returns-to-scale (CRS) assumption to measure the management efficiency of US banks. They found that there is a gap between efficiency scores of non-failed and failed banks, which is both significant and increasing as the failure date approaches. However, in bankruptcy and distress prediction, the returns-to-scale (RTS) regime is increasing or decreasing, therefore, the CRS assumption is unrealistic.

Pille and Paradi (2002) developed four input-oriented BCC models (Banker, Charnes, & Cooper, 1984) with different combinations of inputs and outputs, under the variable returns-to-scale (VRS) assumption, to predict financial failure of Credit Unions. The performance of DEA efficiency scores was statistically compared with a government modified “Z-score” model and

“equity to asset” ratio. Overall, consistent with Barr et al. (1993), they found that failed Credit Unions, especially one year before the failure time, have lower scores than healthy ones.

Paradi et al. (2004) proposed the worse practice DEA analysis under BCC - aimed at finding the companies that are efficient at being bad - in combination with a layering technique, rather than a fixed cut-off point, to classify manufacturing firms into bankrupt and non-bankrupt. Also, they employed a different combination of inputs (the drivers of bad performance like current liabilities, interest expense, and bad debt) and outputs (the drivers of good performance like total asset, sales, profit) to identify the best set of inputs/outputs. The results suggested that combining the first three layers of worst practice efficiency frontiers improves the classification accuracy of identifying bankrupt and non-bankrupt firms up to 100 and 67 percent, respectively.

Cielen et al. (2004) used CCR scores to predict bankruptcy and compared such a DEA classifier with a linear programming model (minimised sum of deviations (MSD)) and a rule induction (C5.0) model. They suggested using financial ratios with a positive correlation as inputs and those with a negative correlation as outputs. Regarding prediction accuracy, the result indicated that DEA outperforms both C5.0 and MSD models. However, the main methodological issue is that CCR cannot deal with negative values of financial ratios.

[Insert *Table 2* Here]

Bowlin (2004) analysed cross-sectional and longitudinal differences in DEA scores under the BCC model, over a 10-year period, 1988 -1997, to compare the financial stability of different groups of firms using the statistical approach of Banker (1993).

Emel et al. (2003) and Min and Lee (2008) applied an input-oriented CCR model to measure financial performance, namely, creditability scores, which were then used to classify firms with scores equal to one and less than one as companies with a good and relatively worse financial performance, respectively. To validate the discriminatory power of DEA, they used the DEA score as the dependent variable and financial ratios as independent variables in regression and discriminant analysis. The results suggested that DEA is a valid method for estimating the creditworthiness of companies.

Premachandra et al. (2009) employed the additive DEA model of Charnes, Cooper, Golany, Seiford, and Stutz (1985) for bankruptcy prediction to take advantage of its specific features.



First, the additive DEA allows for negative values of inputs and outputs, as a result of its translation invariance property. Second, in contrast to radial models (i.e. CCR and BCC models), which require examination of both a DEA efficiency score and slacks to estimate the efficiency of a DMU, the additive model requires the consideration of slacks only. Third, while the radial DEA model based on an input-oriented or an output-oriented measurement results in different efficiency scores, the additive model includes both input and output slacks in the efficiency analysis, then it avoids the problem related to ratio form. The comparison of the additive DEA model with Logistic regression (LR) indicates that the DEA model (respectively LR) outperforms (respectively underperforms) in predicting non-bankrupt (respectively bankrupt) firms.

However, additive DEA for bankruptcy prediction has some drawbacks. First, the selection of input and output variables are in reverse order to conventional DEA model that would usually cause different results (Shetty et al., 2012). Second, the additive model does not provide an efficiency score between 0 and 1. In other words, although the estimated measure can discriminate between bankrupt and non-bankrupt firms, it fails to evaluate the depth of bankruptcy (Premachandra et al., 2011; Shetty et al., 2012).

To overcome the above drawbacks, Premachandra et al. (2011) applied super-efficiency additive DEA model (Fang, Lee, Hwang, & Chung, 2013) to develop a discriminant index based upon two frontiers, namely failure and success. Switching input-output classification identifies these two frontiers. Therefore, for determining failure (respectively success) frontier, the smaller (respectively larger) values in the financial ratios are considered as input (respectively output), and the larger (respectively lower) values in those ratios are considered as output (respectively input). The results indicate that the super-efficiency additive DEA model is relatively weaker in predicting failed firms compared to non-failed ones. However, the discriminant index based on two frontiers improves this weakness by giving the practitioners the option to choose different accuracy levels of failure, non-failure, and total prediction.

Also, to overcome the shortcomings of the application of the additive DEA by Premachandra et al. (2009) in bankruptcy prediction, Shetty et al. (2012) proposed a modified efficiency measure using orientation-free non-radial directional distance formulation of DEA. Contrary to the additive DEA model, this approach measures the worst relative efficiency within the range of

zero to one. Further, contrary to the conventional DEA methods, this method identifies the worst performers and locates an inefficient frontier.

Sueyoshi (1999) proposed a new type of discriminant analysis, namely “DEA-Discriminant Analysis (DEA-DA)” that incorporates the methodological advantages of DEA (for example, nonparametric and distribution-free features) into the discriminant analysis. This two-stage approach is designed to identify the existence of an overlap between the two groups at the first stage and to determine a group classification function for new observation samples at the second stage. Sueyoshi (2001) proposed the “extended DEA-DA” approach, which has two important features; 1) it can deal with negative values, and 2) it can estimate the weights of a DA function by minimising the total distance of misclassified observations. However, the drawback of the “extended DEA-DA” model is that it does not reduce the number of misclassified observations (as explained in accuracy performance evaluation), but the total distance of misclassified observations. To overcome this methodological issue, Sueyoshi (2004) proposed a mixed integer programming (MIP) version of DEA-DA to estimate the weights of the linear discrimination function by minimising the total number of misclassified observation.

Furthermore, Sueyoshi (2006) compared the performance of two advanced versions of DEA-DA classifiers, namely standard MIP and two-stage MIP models with six other bankruptcy prediction models; logit, probit, Fisher’s linear DA, Smith’s quadratic DA, neural network, and decision tree. Tsai et al. (2009) also used the MIP version of DEA-DA (Sueyoshi, 2004) as a predictor of loan default and compared its accuracy with DA, LR and NN models. The result suggests that DEA-DA and NN have a better-classifying capability.

Further, in a proposed two-stage model, Sueyoshi et al. (2010) applied RAM (range-adjusted measure: Aida, Cooper, Pastor, & Sueyoshi, 1998; Cooper, Park, & Pastor, 1999) as a DEA model to measure the operational efficiency scores of Japanese companies, in the first step. In the second step, the RAM efficiency score is used as the dependent variable in a Tobit regression to investigate whether corporate governance variables influence the operational efficiency of firms.

Mukhopadhyay et al. (2012) proposed a combination of DEA and Multi-Layer Perceptron (MLP) to predict failure. First, they used super-efficiency negative DEA to identify the worst performers amongst the non-failed firms (i.e. companies with an efficiency score greater than 1).

The recognised worst non-failed firms in combination with failed firms are labelled as a failed group and then used to train the MLP. The developed MLP was then used for failure prediction. The proposed technique, therefore, recognises firms that have a high likelihood of facing failure along with those that have filed for bankruptcy.

Avkiran and Cai (2014) applied a super-efficiency SBM model (Fang et al., 2013; Tone, 2011) as a forward-looking approach to predict distressed bank holding companies. Results suggested that DEA could identify distressed banks up to 2 years ahead.

More recently, Ouenniche and Tone (2017) used a BCC (respectively SBM) model to estimate efficiency scores of the London Stock Exchange (LSE) listed companies and proposed a novel rule to classify firms in the training sample into bankrupt and non-bankrupt ones, which involves solving a non-linear programme to determine a DEA score-based cut-off point so as to optimise a given performance measure. Then, a k-Nearest Neighbour (k-NN) algorithm is trained on the in-sample classification and used to classify out-of-sample firms into bankrupt and non-bankrupt ones. The performance of the proposed DEA-CBR framework for in-sample and out-of-sample classification was tested by reworking the DA model of Taffler in their framework, which delivered an outstanding performance.

## 2.2. DEA Score as a Predictor

In the recent trend of DEA application in distress prediction, the use of DEA efficiency score as a feature in developing prediction models is becoming more prevalent. In the earliest study, Barr and Siems (1997) used the CCR model to measure the managerial efficiency of US banks at the first stage and then used the CCR efficiency score as a predictor in a Probit model at the second stage. Their findings suggest that removing the management efficiency variable from the Probit model decreases the model's fit and classification accuracy.

Xu and Wang (2009) used the BCC model to estimate the efficiency scores of Chinese firms in the first step. The second step compares the prediction accuracy of three failure prediction models, namely, SVM, MDA and logistic regression, with and without DEA efficiency. The results indicate that using an efficiency score improves the performance of prediction models effectively.

Yeh et al. (2010) used the CCR model to measure the efficiency of Taiwanese information and electronic manufacturing firms. The estimated CCR efficiency score and a list of frequently used financial ratios are employed as inputs of the second stage, namely rough set theory (RST), to select the most significant features. Finally, the selected features from RST are used as inputs of support vector machines (SVM) and back-propagation neural networks (BPN) to predict business failures. The results suggest that for both RTS-SVM and RST-BPN, the models using both financial ratios and DEA scores deliver better classification results than the corresponding models only employing financial ratios.

Psillaki et al. (2010) proposed a two-stage model of credit risk prediction. In the first stage, they used a directional distance DEA model under VRS to determine the efficiency scores of a sample of French manufacturing firms. The firm efficiency score measures the company's distance from the industry's best practice frontier. In the second stage, they used logistic regression to evaluate the effect of a company's efficiency in predicting failure over and above that explained by financial features. The findings suggest that more efficient firms are less likely to fail.

Li et al. (2014) proposed a new application of DEA in bankruptcy prediction through using SBM-VRS to estimate Technical Efficiency (TE) and decomposing TE into Pure Technical Efficiency (PTE) and Scale Efficiency (SE) for a sample of Chinese companies, at the first stage. In the second stage, these efficiency measures along with other financial ratios are used in a Logistic regression model to predict the probability of failure. The authors introduced an interaction term into the model to allow for the impact of a variety of efficiency scores, across all industries, on the probability of failure to be identified.

There are several limitations to the application of DEA in corporate failure and distress prediction studies, which could be summarised as follows. First, most of the studies have used cross-sectional or static DEA models that fail to consider the changes in efficiency over time. To the best of our knowledge, the only exception is Li et al. (2017) who applied time-varying Malmquist DEA to estimate dynamic efficiency scores and used them in a dynamic prediction model. Second, many of these studies have applied DEA under a constant returns-to-scale (CRS) regime (Paradi et al., 2004; Xu & Wang, 2009; Yeh et al., 2010; Avkiran & Cai, 2014; Mukhopadhyay et al., 2012) rather than VRS ones (Psillaki et al., 2010; Z. Li et al., 2014), which is the typical case in bankruptcy and distress prediction. Third, on DEA analysis orientation,

three studies (Cielen et al., 2004; Psillaki et al., 2010; Yeh et al., 2010) are output-oriented, and the remaining majority are input-oriented with a rather limited justification of the choice of orientation. Forth, while a large number of studies have estimated the DEA efficiency scores of firms using financial accounting variables (e.g. total assets, total liabilities, total sales, employees, cash flow), as far as we are aware only one (Avkiran & Cai, 2014) has estimated efficiency using market variables (e.g., market capitalisation, annual stock return, liquid asset) as inputs and outputs of DEA models. Finally, the effect of mix efficiency, which indicates the capacity to improve the effectiveness of firms by managing input- or output-slacks, on the probability of distress has never been investigated.

### **3. Research Methodology**

This section provides the details of our research methodology, where we compare the performance of two-stage distress prediction models. To this end, we provide the details on our dataset (see section 3.1), the static and dynamic DEA models used in the first stage of two-stage DPMs (see section 3.2), and the static and dynamic models specification in the second stage of two-stage DPMs (see section 3.3).

#### **3.1. Data**

We took the following steps to select our dataset. First, we considered all non-financial and non-utility UK companies listed on LSE at any time during an 8-year period from 2007 through 2014 - Financial and utility companies are excluded because they are regulated. Second, we excluded the firms, which are listed less than two years in LSE, as historical information is a requirement for some modelling frameworks. Third, we excluded the firms with missing values for the principal accounting items (e.g., sales, total assets) and market information (e.g., price), which are necessary for calculating many financial ratios (Lyandres & Zhdanov, 2013). We replaced the remaining missing values with the recently observed ones for each firm (Shumway, 2001; Zhou, 2013). Fourth, we winsorised the outlier values by replacing the values higher (respectively lower) than 99<sup>th</sup> (respectively 1<sup>st</sup>) percentile of each variable with the 99<sup>th</sup> (respectively 1<sup>st</sup>) percentile value (Shumway, 2001). Fifth, we lagged all data including market-driven variables described below to ensure that the data are available at the beginning of the year in which bankruptcy is observed (Bauer & Agarwal, 2014; Shumway, 2001).

[Insert *Table 3* Here]

Regarding the classification of firms into distress and non-distress classes, we followed the proposed definition of financial distress by Pindado, Rodrigues, & de la Torre (2008), where a company is classified as distressed if it experiences both of the following conditions for two consecutive years. First, the company's earnings before interest, taxes, depreciation and amortisation (EBITDA) is lower than its interest expenses, and second, the company shows negative growth in market value. To be more specific, the distress variable, say  $y$ , equals 1 for financially distressed companies and equals 0 otherwise. In sum, our dataset consists of 2,096 firms and 11,943 firm-year observations. Among the total number of observations, there are 676 firm-year observations classified as distressed resulting in a distress rate average of 5.66 percent per year. The models are developed using the training sample period ranging from 2007 to 2011 and tested using the holdout sample period ranging from 2012 to 2014. Table 3 presents the sample sizes.

### 3.2. Stage One: Estimating Efficiency Measures Using DEA Models

In this section, we explain cross-sectional (static) DEA models (see section 3.2.1), and Malmquist DEA model (see section 3.2.2) applied in the first stage of two-stage DPMs. Then, we describe the choice of inputs and outputs for DEA models (see section 3.2.3).

#### 3.2.1. Static DEA Models

Several types of DEA models can be used depending on the conditions of the problem. Further, the types of DEA model can be identified based on the scale and orientation of the model. In this study, to compute the cross-sectional efficiency measures of companies, we use CCR (Charnes, Cooper, & Rhodes 1978), BCC (Bankert, Charnes & Cooper, 1984) and SBM (Tone, 2011) DEA models and perform both input-oriented (IO) and output-oriented (OO) analyses – See Table 4 and Table 5 for details about DEA models. Also, we use SBM model under both constant returns-to-scale (CRS) and variable returns-to-scale (VRS) regimes, separately.

[Insert *Table 4* Here]

Note that the CCR and BCC scores are called the (global) technical efficiency (TE) and the (local) pure technical efficiency (PTE), respectively. The BCC model estimates the efficiency of

DMUs when returns-to-scale (RTS) is not necessarily constant, i.e., it takes account of scale effect and postulates that convex combinations of the observed DMUs form the production possibility set (Cooper, Seiford, & Tone, 2007, p. 153). If a DMU has full BCC efficiency but a low CCR efficiency, then it is operating locally efficient but not globally efficient, because of the scale size of the DMU. Considering these concepts and denoting CCR and BCC scores as  $\theta_{CCR}^*$  and  $\theta_{BCC}^*$ , respectively, the *scale efficiency* (SE) is defined as follows (Charnes et al., 1978) :

$$SE = \frac{\theta_{CCR}^*}{\theta_{BCC}^*} \quad \text{Eq. 1}$$

Therefore, the technical efficiency could be decomposed as

$$TE = PTE \times SE \quad \text{Eq. 2}$$

The advantage of this decomposition is that it determines the sources of inefficiency, i.e., whether it is due to inefficient operation (PTE) or due to detrimental conditions displayed by the scale efficiency (SE) or by both.

[Insert *Table 5* Here]

Moreover, radial DEA models, i.e., CCR and BCC scores, overlook possible slacks in inputs and outputs, and therefore, would possibly over-estimate the efficiency scores by ignoring mix efficiency. The SBM model is a non-radial model that considers slacks in inputs and outputs. Note that the equality of optimal input-oriented (respectively, output-oriented) SBM measure, i.e.,  $\rho_{in}^*$  (respectively,  $\rho_{out}^*$ ), and optimal input-oriented (respectively, output-oriented) CCR measure, i.e.,  $\theta_{CCR-in}^*$  (respectively,  $\theta_{CCR-out}^*$ ) holds, i.e.,  $\rho^* = \theta_{CCR}^*$ , if the input-oriented (respectively, output-oriented) CCR model has zero input-slacks (respectively, output-slacks) for every optimal solution. In other words, the strict inequality, i.e.,  $\rho_{in}^* < \theta_{CCR-in}^*$  (respectively,  $\rho_{out}^* < \theta_{CCR-out}^*$ ) holds if and only if the CCR measure indicates an input (respectively, output) mix inefficiency. Considering these concepts, the input and output “mix efficiency” (ME) are defined by Cooper et al. (2006: P.156) as follows.

$$ME_{in} = \frac{\rho_{in}^*}{\theta_{CCR-in}^*} \quad \text{and} \quad ME_{out} = \frac{\rho_{out}^*}{\theta_{CCR-out}^*} \quad \text{Eq. 3}$$

Considering equation 1 (the decomposition of TE), the non-radial input- or output-oriented technical efficiency (SBM) could be decomposed into mixed efficiency (ME), pure technical efficiency (PTE) and scale efficiency (SE) as follows:

$$SBM = PTE \times SE \times ME$$

Eq. 4

In this study, we use CCR-IO, CCR-OO, BCC-IO, BCC-OO, SBM-CRS-IO, SBM-CRS-OO, SBM-VRS-IO and SBM-VRS-OO models to measure the cross-sectional managerial efficiency and market efficiency of companies. Also, we decompose the SBM measure of each company into ME, PTE and SE, and incorporate them in developing distress prediction models in the second stage.

### 3.2.2. Dynamic DEA Model

To estimate the efficiency measures of companies over time, we use the Malmquist DEA productivity index (Färe, Grosskopf, Norris, & Zhang, 1994; Färe, Lindgren, & Roos, 1992). Malmquist productivity index (MPI) is a multi-criteria assessment framework for comparing the performance of DMUs over time. Färe et al. (1992, 1994) used DEA to extend the original Malmquist Index proposed by Malmquist (1953) and constructed the DEA-based Malmquist productivity index as the product of two components; (1) catching-up to the frontier, which refers to the efficiency change (EC) of DMU with respect to the efficiency possibilities defined by the frontier in each period, and (2) efficient frontier-shift (EFS), which refers to the shift of efficient frontier between the two time periods  $t$  and  $t + 1$  (see, Table 6 for details about Malmquist productivity index).

Caves et al. (1982) introduced a distance function,  $\Delta(\cdot)$ , to measure technical efficiency with the basic CCR model (Charnes et al., 1978). Though, in the non-parametric framework, instead of using a distance function, DEA models are implemented. For example, Färe et al. (1994) used input (or output) oriented radial DEA model to measure the MPI. However, the radial model faces a lack of attention to slacks that could be overcome using Slacks-based non-radial oriented (or orientation-free) DEA model (Tone, 2011, 2002). Along with measuring cross-sectional DEA scores (section 3.2.1), we incorporate CCR-IO, CCR-OO, BCC-IO, BCC-OO, SBM-CRS-IO, SBM-CRS-OO, SBM-VRS-IO and SBM-VRS-OO models to measure the MPI. Also, we decompose the SBM measure of each company into ME, PTE and SE, and incorporate them in developing dynamic distress prediction models in the second stage.

[Insert *Table 6* Here]



### *Global Malmquist Productivity Index*

The primary objective of this study is to estimate the relative efficiency of *DMUs* (companies) in each period. However, the estimated Malmquist productive index (MPI), say,  $MPI_0^{t,t+1}$ , indicates the change of efficiency score between period  $t$  and  $t + 1$ , and not relative efficiency scores of *DMUs* at each period, then should be modified for our purpose. Further, referring to Pastor and Lovell (2005), the contemporaneous MPI is not circular, its adjacent period components can give conflicting signals, and it is sensitive to LP infeasibility.

The adjacent reference index, proposed by Färe et al. (1992), suggests multiplying  $MPI_0^{t,t+1}$  by  $\Delta_0^t(x_0^t, y_0^t)$ , which results in the relative efficiency of  $DMU_0$  at period  $t + 1$  compared to period  $t$ . However, the main drawback of this index is that it cannot estimate the relative efficiency score of non-adjacent periods, e.g., period  $t$  and  $t + 2$  or  $t + 1$  and  $t + 3$ . To overcome this drawback, Berg, Forsund, & Jansen (1992) used a fixed reference index, which compares and refers the relative efficiencies of all periods (say,  $t$  ( $t \geq 2$ )) to the first period (say,  $t = 1$ ). Therefore, it is possible that the efficiency scores of the periods later than the first one are more than 1 since the technology develops over time. Although, the fixed reference index acquires the circularity property with a base period dependence, it remains sensitive to LP infeasibility. More recently, Pastor and Lovell (2005) suggested a global MPI that contains circular competent, is not susceptible to LP infeasibility and provides a single measure of productivity change. Further, in a situation where efficient frontiers of multiple periods cross each other, the global index can be measured by the best practices in all periods. As Figure 2 presents, the relative efficiency of  $DMU_0$  can be measured in terms of either the frontier of period 1 (consists of four *DMUs*: 1,2,3,4 and 5) or the frontier of period 2 (consist of four *DMUs*: 6,7,8,9 and 10). An alternative is the global frontier, which is the combination of the best *DMUs* in the history, i.e. five *DMUs*: 6,7,3,4 and 5.

[Insert *Figure 1* and *Figure 2* Here]

It is argued that if the length of the observation period is long enough, the current *DMUs* would be covered by the best historical *DMUs*, probably themselves. Thus, the relative efficiency to the global frontier could be considered as an absolute efficiency with the scores less than or equal to 1 (Pastor & Lovell, 2005). In this paper, we use the global MPI proposed by Pastor and Lovell (2005), which overcome the above mentioned issues with previously proposed MPIs.

### 3.2.3. Choice of Inputs and Outputs for the First Stage

To select suitable inputs and outputs for DEA models, we considered the following issues. First, the survey on the application of DEA in bankruptcy and distress prediction indicates that there is no standard choice of inputs and outputs for DEA – see Table 1 above. In practice, different DEA applications use different inputs and outputs, which is one of the drawbacks of DEA application (Premachandra et al., 2009). However, the choice of inputs and outputs should be related to the competitive environment (Oral & Yolalan, 1990). Second, regarding most of two-stage prediction models, since financial ratios are used as features in the second stage, the monetary items of financial statements are used as inputs and outputs of DEA models in the first stage (Z. Li et al., 2014, 2017; Psillaki et al., 2010; Xu & Wang, 2009). Third, to the best of our knowledge, two-stage studies only used accounting items as inputs and outputs of DEA models to compute the managerial efficiency of companies (see, Table 8 for details). Fourth, to deal with negative values in inputs and outputs of DEA, the following popular approaches have been proposed: The Range Directional Measure introduced by Portela, Thanassoulis, and Simpson (2004), the Modified Slack-Based Measure introduced by Sharp, Meng and Liu (2007), the Semi-Oriented Radial Measure introduced by Emrouznejad, Anouze, and Thanassoulis (2010), and Variant of Radial Measure introduced by Cheng, Zervopoulos, and Qian (2013).

In this study, for estimating management efficiency, we selected three inputs (Total Liabilities, Total Shareholders' Equity and Number of Employees) and one output (Total Sales). Also, for estimating the market efficiency of firms, we implement the simple concept of Sharpe ratio, i.e., the ratio of excess return over the volatility of return, into DEA to estimate the market efficiency of firms. We use the market-driven variables of Shumway (2001) as inputs and outputs of DEA models. Traders discount the equity of firms that are close to distress then a firm's last year excess returns and market value are associated with the market performance and probability of distress. Further, the idiosyncratic standard deviation of each firm's stock returns, denoted sigma, is associated with market performance and the probability of distress logically. For a firm with high market value and excess return (respectively, high variable stock return), the market performance is high (respectively, low) and the probability of distress is low (respectively, high). Following the logic of maximising outputs and minimising inputs, we used the lag of volatility as input and the lag of excess return and market value as outputs of DEA models.

The firm's market value is calculated as the number of outstanding shares multiplied by the share price at the end of the year before the observation. The firm's past excess return, denoted lag of excess return, in year  $t$  is estimated as the return of firm in year  $t - 1$  minus the value-weighted FTSE index return in year  $t - 1$ . The excess returns of a firm are the cumulative monthly returns of that firm. The firm's sigma is estimated as the standard deviation of the residual derived from regressing monthly stock return on market return in year  $t - 1$ .

To estimate efficiency measures using different DEA models, we employed MaxDEA that deals with negative values in inputs (such as shareholders' equity) and outputs (such as lag of excess return) using the variant radial measure approach (G. Cheng et al., 2013). Table 7 summarises the descriptive statistics of winsorised inputs and outputs.

[Insert *Table 7* Here]

### 3.3. Stage Two: Developing Distress Prediction Model

In this stage, we fed logistic regression with static DEA scores and selected features to develop static DPMs. Also, we fed multi-period logistic regression with dynamic DEA scores and selected features to develop dynamic DPMs.

#### 3.3.1. Static Logit Model

Since the seminal work of Ohlson (1980), Logit has become a frequently used static model in distress and bankruptcy prediction (Duda & Schmidt, 2010; Martin, 1977; Ohlson, 1980; Back et al., 1995). In the field of financial distress prediction, the dependent variable is a binary variable, which takes on two values, zero or one. The generic model for binary variables could be stated as follows:

$$\begin{cases} P(\text{distress}) = P(y_i = 1|x_i) \\ P(\text{distress}) = G(\beta, X) \end{cases} \quad \text{Eq. 5}$$

where  $Y$  denotes the binary response variable,  $X$  denotes the vector of covariates,  $\beta$  denotes the vector of coefficients of covariates in the model, and  $G(\cdot)$  is a link function that maps the scores,  $\beta^t x$ , onto a probability. In practice, depending on the choice of the link function, the type of probability model is determined. As for the logit regression model, the link function is the cumulative logistic distribution function, say  $\theta$ .

$$G(\beta, X) = \theta^{-1}(\beta^t X) \quad \text{Eq. 6}$$

which is between zero and one for all real numbers  $\beta^t X$ . For our analysis, we specified logistic regression to be

$$\text{logit}(p_i) = \alpha + \sum_{i=1}^n \sum_{r=1}^m \gamma_r S_{ri} + \sum_{i=1}^n \sum_{j=1}^l \beta_j x_{ji} \quad \text{Eq. 7}$$

where  $p_i$  denotes the probability of facing distress for company  $i$ ;  $S_{ri}$  denotes the  $r$ -th static efficiency score for company  $i$ ;  $\gamma_r$  denotes a parameter for the static efficiency score  $r$  to be estimated;  $x_{ji}$  denotes a measure of feature  $j$  for company  $i$ , and  $\beta_j$  is a parameter for feature  $j$  to be estimated.

### 3.3.2. Dynamic Discrete-Time Hazard Model

Shumway (2001) proposed a discrete time hazard model using an estimation procedure similar to the one used for determining the parameters of a multi-period (dynamic) logit model. Many studies have applied this approach for computing the probability of a hazard occurrence (see, for example, K. F. Cheng, Chu, & Hwang, 2010; Nam, Kim, Park, & Lee, 2008; El Kalak & Hudson, 2016; Shumway, 2001). A general description of a discrete time hazard model could be presented as follows:

$$P(y_{i,t} = 1 | x_{i,t}) = h(t | x_{i,t}) = \frac{e^{(\alpha_t + x_{i,t}\beta)}}{1 + e^{(\alpha_t + x_{i,t}\beta)}} = h_0(t) \cdot e^{x_{i,t}\beta} \quad \text{Eq. 8}$$

where  $h(t | x_{i,t})$  represent the individual hazard rate of firm  $i$  at time  $t$ ,  $x_{i,t}$  is the vector of covariates of each firm  $i$  at time  $t$ ;  $\beta$  denotes the vector of coefficients;  $\alpha_t$  is the time-variant baseline hazard function related, which could be related to the firm, e.g.  $\ln(\text{age})$ , or related to macroeconomic variables, e.g. volatility of exchange rate (Nam et al., 2008). We followed Shumway (2001) in using a constant time variant term,  $\ln(\text{age})$ , as a proxy of baseline rate. A firm's age is defined as the number of calendar years it has been traded on the LSE. For our analysis, we modified the discrete-time hazard model as follows:

$$\text{logit}(h_{i,d=1}(t)) = \alpha + \beta_0 h_0(t) + \sum_{i=1}^n \sum_{r=1}^m \gamma_r D_{rit} + \sum_{i=1}^n \sum_{j=1}^l \beta_j x_{jit} \quad \text{Eq. 9}$$

where  $h_{i,d=1}$  denotes the probability of facing distress for company  $i$  at time  $t$ ;  $h_0(t)$  denotes the baseline hazard function;  $D_{rit}$  denotes the dynamic efficiency score  $r$  for company  $i$  at time  $t$ ;  $x_{jit}$  denotes a measure of feature  $j$  for company  $i$  at time  $t$ ;  $\beta_0$  is the coefficient of the baseline hazard rate to be estimated;  $\gamma_r$  is a parameter for the dynamic efficiency score  $r$  at time  $t$  to be estimated; and  $\beta_j$  is a parameter for feature  $j$  at time  $t$  to be estimated.

### 3.3.3. Choice of Features for the Second Stage

To select suitable features for prediction models, we applied the following steps. First, we reviewed the literature (e.g., Ravi Kumar & Ravi, 2007; Zhou, 2013; Zhou, Lu, & Fujita, 2015) to select the most commonly used features in other studies including 83 accounting-based ratios and 7 market-based information. Second, we used *t-test* method to choose features which show a significant difference between two group's means (Shin & Lee, 2002; Huang, Chen, Hsu, Chen, & Wu, 2004; Shin, Lee, & Kim, 2005).

[Insert *Table 8* Here]

Third, for further reduction of features, we applied factor analysis, and principal component analysis with VARIMAX technique (M. Y. Chen, 2011; Mousavi & Ouenniche, 2018). To be more specific, we used factors analysis to select the variables that both the absolute values of their loadings and communities are greater than 0.5 and 0.8, respectively. Fourth, 34 variables which presented high factor loadings and high communality values, were retained as input features into the stepwise procedure in the second stage of two-stage distress prediction models (see, Table 8), where a stepwise procedure for each framework is used to select the most significant features.

### 3.3.4. Choice of Efficiency Scores for the Second Stage

Table 9 presents the descriptive statistics of static and dynamic managerial efficiency measures for two groups of distressed and healthy companies. The results of *F-test* suggest that, in most cases, input-oriented DEA scores discriminate better between the two groups of distressed and healthy firms. Therefore, we select input oriented managerial efficiency measures, i.e., CCR-IO, BCC-IO, SBM-CRS-IO and SBM-VRS-IO, and use equations 1 and 3 to compute SE-IO and ME-IO for the second stage. Table 10 shows the descriptive statistics of static and dynamic

market efficiency measures for two groups of distressed and healthy companies. The results of *F-test* indicate that, in most cases, output-oriented DEA models discriminate better between the two groups of distressed and healthy firms. Then, we chose output-oriented market efficiency measures, i.e., CCR-OO, BCC-OO, SBM-CRS-OO, SBM-VRS-OO, SE-OO and ME-OO for the second stage. We retain the selected static and dynamic scores and relate them to the probability of distress using equations 7 and 9, respectively.

[Insert *Table 9* Here]

[Insert *Table 10* Here]

#### **4. Empirical Results**

The objective of this study is to evaluate the relative performance of two-stage distress prediction models using UK data. Section 4.1 provides the assessment of models using the conventional unidimensional-ranking framework. Section 4.2 assesses the models using a multi-criteria evaluation framework.

##### **4.1. Unidimensional Ranking of Distress Prediction Models**

For a unidimensional ranking of different models, we use the commonly used performance criteria in the literature; i.e., the discriminatory power, the calibration accuracy, the information content, and the correctness of categorical prediction. Regarding the discriminatory power criterion that measures how much a prediction model can discriminate between distressed firms and healthy ones, we use Receivable Operating Characteristic (ROC), Kolmogorov-Smirnov (KS) statistics, Gini Index (GI), and Information Value (IV) as measures. Regarding the calibration accuracy criterion that measures how much a model is qualified in estimating the probability of distress (PD), we use Brier Score (BS) as a measure. Regarding the information content criterion that measures the extent to which the output of a model (e.g., PD, scores) carries enough information for prediction, we follow Agarwal and Taffler (2008) and use a log-likelihood statistic (LL) and pseudo- $R^2$  as measures. Finally, with respect to the correctness of the categorical prediction criterion that measures how often a model can predict distressed firms (respectively, healthy firms) as distressed (respectively, healthy) ones, we use Type I error (T1), Type II error (T2), misclassification rate (MR), sensitivity (Sen), specificity (Spe), and overall

correct classification (OCC) as measures (for more details about performance criteria and measures, the reader is referred to Mousavi, Ouenniche and Xu (2015)). Table 11, Table 12, Table 13 and Table 14 present the estimated distress prediction models. The  $\chi^2$  tests indicate that all 34 models explain a significant amount of variation in the probability of distress. Tables 11 and 12 present the estimated static models using the stepwise procedure in a logit framework using managerial efficiency scores and market efficiency scores, respectively.

[Insert *Table 11* Here]

[Insert *Table 12* Here]

Also, Tables 13 and 14 indicate the estimated dynamic models using the stepwise procedure in a multi-period logit framework using managerial efficiency scores and market efficiency scores, respectively. Retained earnings to total assets, negative net income for last two years, the lag of excess return, log (total asset to GNP index), real size, current liabilities over current assets, log (price) and inventory turnover are amongst the selected variables using the stepwise procedure.

[Insert *Table 13* Here]

[Insert *Table 14* Here]

Table 15 summarises the performance measures of the 34 developed distress prediction models. The results could be summarised as follows. First, considering the performance of models without efficiency measures, i.e., the one-stage static model 1 and the one-stage dynamic model 18, and models fed with efficiency measures, i.e., two-stage models, the results suggest that incorporating efficiency measures improves the performance of models. This result is consistent with the findings of Li et al. (2014, 2017), Psillaki et al. (2010), Yeh et al. (2010) and Xu and Wang (2009).

Second, comparing the performance of dynamic models with static models in our study, for most of the performance measures, the dynamic models outperform static ones. To be more specific, on most performance measures – see, for example, T1, ROC, Gini, KS, IV, CIER, BS, LL and  $R^2$ , the two-stage dynamic models are superior to static ones. However, considering T2, MR and OCC as performance measures of correctness of categorical prediction, static models 16, 11 and 10 are amongst the best performers. In general, the density of dynamic models amongst the top-ranking performers suggests their superiority in performance. The superiority of dynamic to static models could be related to their ability in incorporating time-varying features of the firms. This finding indicates that taking into account the multi-period performance of companies over

time as an explanatory variable in a dynamic framework is an appropriate technique to improve the performance of prediction models.

Third, considering the performance of two-stage models with different types of company efficiency measures, i.e., market efficiency and managerial efficiency, for most of the performance measures, the models with management efficiency outperform the models with market efficiency. The reason is that as the *F*-tests in Table 9 and Table 10 indicate, the discriminatory power of the management efficiency scores is more than market efficiency measures.

[Insert *Table 15* Here]

However, the findings suggest that taking into account T2, MR and OCC as measures of correctness of categorical prediction and BS as a measure of calibration accuracy, the models with market efficiency score outperform others.

Fourth, on the type of DEA scores that models are fed with, i.e., decomposed DEA scores and original DEA scores, the following findings are notable. For static models, models 8, 5 and 4 that use decomposed managerial DEA scores, i.e., PTE, SE and ME, outperform the models that use original DEA scores, i.e., TE and SBM, considering most of the performance criteria. Also, model 16 that use decomposed market DEA scores, i.e., PTE, SE and ME, outperform all models on T2, MR, and OCC. For dynamic models, models with decomposed managerial DEA scores, i.e., models 25, 22 and 21 are superior regarding most of the performance criteria. Further, model 32 with market ME score is the best performer considering BS, LL,  $R^2$ , T2, MR, and OCC. This finding suggests that using decomposed efficiency DEA scores improve the performance of prediction models. These results are consistent with Li et al. (2014, 2017) that suggest models with decomposed measures are superior.

#### 4.2. Multi-criteria Ranking of Distress Prediction Models

For multi-criteria evaluation of DPMs, we followed Mousavi et al. (2015) in using the super-efficiency orientation-free SBM-DEA framework. We perform two rounds of evaluation using four different measures. In the first round, we use T1 error (as a measure of correctness of categorical prediction), BS (as a measure of calibration accuracy) as inputs and ROC (as a measure of discriminatory power) and  $R^2$  (as a measure of information content) as outputs of



DEA model. Also, in the second round, we replace T1 with T2 error as a measure of correctness of categorical prediction.

From Table 16 the following results of the multi-criteria assessment of DPMs are noteworthy. First, comparing the performance of models without efficiency measures with models fed with efficiency measures as predictors, numerical results indicate that using efficiency measures improve the performance of models.

Second, comparing the performance of dynamic models with static models in our study, taking into account T1 error (Panel A of Table 16) as a measure of correctness of categorical prediction, numerical results show that the dynamic models outperform the static ones. However, concerning T2 error (Panel B of Table 16), the results suggest that the static models are comparable to the dynamic ones.

Third, comparing the performance of two-stage models with different types of company efficiency measures, i.e., market efficiency and managerial efficiency, under T1 error as a measure of correctness of categorical prediction (Panel A of Table 16), the results suggest that model 21 with managerial efficiency is the best model; though, model 32 that uses market efficiency is the third in ranking. However, choosing T2 error (Panel B of Table 16), models 16, 31 and 11 that use market efficiency are among the top five models. This result is consistent with a unidimensional ranking of models that suggest models with market efficiency scores outperform others under T2, MR and OCC as measures of correctness of categorical prediction. These results could be linked to the efficient market hypothesis theory, which claims that, in an efficient market, shares prices contain all available information, i.e. past, present and insider information, about the company.

Fourth, considering the type of DEA scores that models are fed with, i.e., decomposed DEA scores and original DEA scores, the following findings are notable. In panel A (respectively, panel B) of multi-criteria assessment (Table 16), the dynamic model 25 that uses decomposed dynamic managerial DEA score, i.e., PTE, SE and ME (respectively, the static model 16 that uses decomposed static market DEA score, i.e., PTE, SE and ME) are the best performers. Also, model 32 with market ME score is one of the best performers in both Panel A and Panel B of multi-criteria assessment. These results are consistent with unidimensional assessment and suggest that the models with decomposed measures are superior in performance. In practice,

decomposing efficiency scores and using them in building prediction models provide more effective drivers of failure and therefore improve the performance of failure prediction models.

[Insert *Table 16* Here]

## **5. Conclusion**

Following the extended application of DEA in providing analytical support for business decision making in the field of credit scoring and distress prediction, this study extends previous research on expert systems (Z. Li et al., 2017; Min & Lee, 2008; Shetty et al., 2012; Xu & Wang, 2009) by demonstrating the potential effectiveness of managing input- and output-slacks, i.e. mix efficiency (ME), in enhancing the performance of prediction models; by investigating the association of companies' market efficiency and probability of distress; and by providing an empirical comparative analysis between several developed two-stage DPMs using different measures of efficiency as features and different dynamic and static frameworks.

This study uses CCR, BCC and SBM-DEA models to estimate cross-sectional efficiency measures and applies Malmquist-DEA models to estimate dynamic efficiency measures. Also, it decomposes overall static and dynamic SBM efficiency scores into PTE, SE and ME scores, and overall static and dynamic TE efficiency score into PTE and SE scores and combine them with accounting, market, and macroeconomic ratios to develop DPMs.

The empirical results suggest that taking account of the efficiency measures of companies, e.g., managerial, and market efficiency, improves the performance of DPMs. Further, findings suggest that the measures of managerial efficiency contribute more to enhancing distress prediction, although the measures of market efficiency, especially mix efficiency, improve the performance of prediction models as well. The lower contribution of the firms' measures of market efficiency compared to managerial efficiency could be explained by the choices of inputs and outputs of DEA models. Also, the findings indicate that incorporating dynamic efficiency measures in a dynamic distress framework is the best approach to improve the accuracy of DPMs. This is because dynamic models by design could take account of changes in the conditions of firms over time. Moreover, the results show that the decomposition of TE (respectively, SBM) efficiency scores into PTE and SE (respectively, PTE, SE and ME) improves the performance of prediction models. This is because incorporating decomposed

measures of efficiency in the model would provide more detailed information on the firm and therefore would enhance distress prediction.

The main strengths of the applied methodology are as follows: firstly, the two-stage modelling uses DEA in the first stage to estimate performance measures of a company, the obtained efficiency scores are then used as features of distress prediction models in the second stage. Second, using market information as input/output of DEA models leads to a new performance measure of a company, namely market efficiency. Third, using DEA models in the first stage benefits developers by incorporating several important inputs/outputs of performance at the same time. Forth, decomposing efficiency measures helps developers to analyse the effect of different perspectives of company performance, say, allocating resources and using new technology, adjusting optimal scale, and managing input- and output-slacks, on distress and bankruptcy. However, the main shortcoming of using DEA models in this study is the lack of a standard tool to select the most substantial inputs and outputs for DEA models; then, they are selected based on popularity in other studies or arbitrary.

The implication of this study is that developers of two-stage DPMs can now make an informed decision regarding the selection of the best DEA models to evaluate companies' efficiency in the first stage, as well as choosing appropriate decomposed efficiency measures to feed the chosen classifiers in the second stage. Another implication of this study is that, in addition to managerial efficiency, it offers insights into another efficiency feature of a company, i.e. market efficiency, that could result in distress.

The main limitations of this research are the time and space and as such this study is restricted to specific DEA models and classifiers in the first and second stages, respectively. Also, as most failure prediction studies (Balcaen & Ooghe, 2006), the criterion of distress is chosen arbitrarily, which may have adverse consequences for the resulting DPMs.

Several future research directions emerge from this study. First, larger and more diversified datasets for experiments and applications, especially with more variety of distress data structures, should be collected to validate the empirical findings of this study further. Second, considering failure-related events, other events such as bankruptcy, capital restructuring, takeover and liquidation should be used to validate the accuracy of this study conclusions. Third, concerning the choice of DEA models to evaluate the managerial and market efficiency of firms, future

studies should incorporate other types of DEA models, e.g. dynamic DEA or Dynamic-Network DEA, and other sources of information as inputs and outputs of DEA models. Forth, regarding corporate efficiency, in addition to managerial and market efficiency, other features of the firm such as corporate governance, financial and operational performance could be considered. Fifth, regarding the second stage of modelling, other statistical techniques, including static frameworks, e.g. discriminant analysis, mixed logit analysis, linear probability analysis, and dynamic ones, e.g. Cox-proportional hazard model, are applicable. Further, a variety of non-parametric techniques such as k-nearest neighbour, cluster analysis, neural networks, support vector machine and Rough set approach should be investigated in future research.

## References

- Abdou, H. A., & Pointon, J. (2011). Credit Scoring, Statistical Techniques and Evaluation Criteria: a Review of the Literature. *Intelligent Systems in Accounting, Finance and Management*, 18(2–3), 59–88. <https://doi.org/10.1002/isaf.325>
- Agarwal, V., & Taffler, R. (2008). Comparing the performance of market-based and accounting-based bankruptcy prediction models. *Journal of Banking and Finance*, 32(8), 1541–1551. <https://doi.org/10.1016/j.jbankfin.2007.07.014>
- Aida, K., Cooper, W. W., Pastor, J. T., & Sueyoshi, T. (1998). Evaluating water supply services in Japan with RAM: A range-adjusted measure of inefficiency. *Omega*, 26(2), 207–232. [https://doi.org/10.1016/S0305-0483\(97\)00072-8](https://doi.org/10.1016/S0305-0483(97)00072-8)
- Alfaro-Cid, E., Sharman, K., & Esparcia-Alcazar, A. (2007). A genetic programming approach for bankruptcy prediction using a highly unbalanced database. In M. Giacobini (Ed.), *Workshops on Applications of Evolutionary Computation* (pp. 169–178). Springer Berlin Heidelberg. [https://doi.org/doi:10.1007/978-3-540-71805-5\\_19](https://doi.org/doi:10.1007/978-3-540-71805-5_19)
- Altman, E. I., Iwanicz-Drozowska, M., Laitinen, E. K., & Suvas, A. (2017). Financial Distress Prediction in an International Context: A Review and Empirical Analysis of Altman's Z-Score Model. *Journal of International Financial Management and Accounting*, 28(2), 131–171. <https://doi.org/10.1111/jifm.12053>
- Avkiran, N. K., & Cai, L. (2014). Identifying distress among banks prior to a major crisis using non-oriented super-SBM. *Annals of Operations Research*, 217(1), 31–53. <https://doi.org/10.1007/s10479-014-1568-8>
- Aziz, M. A., & Dar, H. A. (2006). Predicting corporate bankruptcy: Where we stand? *Corporate Governance*, 6(1), 18–33. <https://doi.org/10.1108/14720700610649436>
- Back, B., Laitinen, T., Sere, K., & Wezel, van M. (1995). Choosing Bankruptcy Predictors Using Discriminant Analysis, Logit Analysis, and Genetic Algorithms. *Proceedings of the 1st International Meeting on Artificial Intelligence in Accounting, Finance and Tax*, 40(40), 337–356.
- Bahrammirzaee, A. (2010). A comparative survey of artificial intelligence applications in finance: Artificial neural networks, expert system and hybrid intelligent systems. *Neural Computing and Applications*, 19(8), 1165–1195. <https://doi.org/10.1007/s00521-010-0362-z>
- Balcaen, S., & Ooghe, H. (2006). 35 Years of Studies on Business Failure: an Overview of the Classical Statistical Methodologies and Their Related Problems. *Working Papers of Faculty of Economics and Business Administration, Ghent University, Belgium*, 32(0), 63–93. Retrieved from <http://ideas.repec.org/p/rug/rugwps/04-248.html>
- Banker, R. D. (1993). Maximum Likelihood, Consistency and Data Envelopment Analysis: A Statistical Foundation. *Management Science*, 39(10), 1265–1273. <https://doi.org/10.1287/mnsc.39.10.1265>
- Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis. *Management Science*, 30(9), 1078–1092. <https://doi.org/10.1287/mnsc.30.9.1078>

- Barr, R. S., Seiford, L. M., & Siems, T. F. (1993). An envelopment-analysis approach to measuring the managerial efficiency of banks. *Annals of Operations Research*, 45(1), 1–19. <https://doi.org/10.1007/BF02282039>
- Barr, R. S., & Siems, T. F. (1997). Bank Failure Prediction Using Dea to Measure Management Quality. In *Interfaces in Computer Science and Operations Research* (pp. 341–365). Springer. [https://doi.org/10.1007/978-1-4615-4102-8\\_15](https://doi.org/10.1007/978-1-4615-4102-8_15)
- Baudraerts, J. (2016). Predicting Bankruptcy in Private Firms: Towards a Stepwise Regression Procedure. *International Journal of Financial Research*, 7(2), 147–153. <https://doi.org/10.5430/ijfr.v7n2p147>
- Bauer, J., & Agarwal, V. (2014). Are hazard models superior to traditional bankruptcy prediction approaches? A comprehensive test. *Journal of Banking and Finance*, 40(1), 432–442. <https://doi.org/10.1016/j.jbankfin.2013.12.013>
- Berg, S. A., Forsund, F. R., & Jansen, E. S. (1992). Malmquist Indices of Productivity Growth during the Deregulation of Norwegian Banking, 1980-89. *Scandinavian Journal of Economics*, 94(1991), S211-28. <https://doi.org/10.2307/3440261>
- Bowlin, W. F. (2004). Financial analysis of civil reserve air fleet participants using data envelopment analysis. *European Journal of Operational Research*, 154(3), 691–709. [https://doi.org/10.1016/S0377-2217\(02\)00814-7](https://doi.org/10.1016/S0377-2217(02)00814-7)
- Caves, D. W., Christensen, L. R., & Diewert, W. E. (1982). The Economic Theory of Index Numbers and the Measurement of Input, Output, and Productivity. *Econometrica*, 50(6), 1393. <https://doi.org/10.2307/1913388>
- Chandrashekar, G., & Sahin, F. (2014). A survey on feature selection methods. *Computers and Electrical Engineering*, 40(1), 16–28. <https://doi.org/10.1016/j.compeleceng.2013.11.024>
- Charnes, A., Cooper, W. W., Golany, B., Seiford, L., & Stutz, J. (1985). Foundations of data envelopment analysis for Pareto-Koopmans efficient empirical production functions. *Journal of Econometrics*, 30(1–2), 91–107. [https://doi.org/10.1016/0304-4076\(85\)90133-2](https://doi.org/10.1016/0304-4076(85)90133-2)
- Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2(6), 429–444. [https://doi.org/10.1016/0377-2217\(78\)90138-8](https://doi.org/10.1016/0377-2217(78)90138-8)
- Chen, M. Y. (2011). Bankruptcy prediction in firms with statistical and intelligent techniques and a comparison of evolutionary computation approaches. *Computers and Mathematics with Applications*, 62(12), 4514–4524. <https://doi.org/10.1016/j.camwa.2011.10.030>
- Chen, N., Ribeiro, B., & Chen, A. (2016). Financial credit risk assessment: a recent review. *Artificial Intelligence Review*, 45(1), 1–23. <https://doi.org/10.1007/s10462-015-9434-x>
- Cheng, G., Zervopoulos, P., & Qian, Z. (2013). A variant of radial measure capable of dealing with negative inputs and outputs in data envelopment analysis. *European Journal of Operational Research*, 225(1), 100–105. <https://doi.org/10.1016/j.ejor.2012.09.031>
- Cheng, K. F., Chu, C. K., & Hwang, R. C. (2010). Predicting bankruptcy using the discrete-time semiparametric hazard model. *Quantitative Finance*, 10(9), 1055–1066.

<https://doi.org/10.1080/14697680902814274>

- Cielen, A., Peeters, L., & Vanhoof, K. (2004). Bankruptcy prediction using a data envelopment analysis. *European Journal of Operational Research*, *154*(2), 526–532.  
[https://doi.org/10.1016/S0377-2217\(03\)00186-3](https://doi.org/10.1016/S0377-2217(03)00186-3)
- Cooper, W. W., Park, K. S., & Pastor, J. T. (1999). RAM: A Range Adjusted Measure of Inefficiency for Use with Additive Models, and Relations to Other Models and Measures in DEA. *Journal of Productivity Analysis*, *11*(1), 5–42.  
<https://doi.org/10.1023/A:1007701304281>
- Cooper, W. W., Seiford, L. M., & Tone, K. (2006). Data envelopment analysis: A comprehensive text with models, applications, references and DEA-Solver Software. Second editions. *Springer, ISBN, 387452818*, 490.
- Cooper, W. W., Seiford, L. M., & Tone, K. (2007). *Data envelopment analysis: A comprehensive text with models, applications, references and DEA-solver software: Second edition. Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software: Second Edition.* <https://doi.org/10.1007/978-0-387-45283-8>
- Du Jardin, P., & Séverin, E. (2012). Forecasting financial failure using a Kohonen map: A comparative study to improve model stability over time. *European Journal of Operational Research*, *221*(2), 378–396. <https://doi.org/10.1016/j.ejor.2012.04.006>
- Duda, M., & Schmidt, H. (2010). *Bankruptcy Prediction: Static Logit Model versus Discrete Hazard Models Incorporating Macroeconomic Dependencies*. Lund University. Retrieved from <http://lup.lub.lu.se/luur/download?func=downloadFile&recordOId=1614029&fileOId=1614031>
- El Kalak, I., & Hudson, R. (2016). The effect of size on the failure probabilities of SMEs: An empirical study on the US market using discrete hazard model. *International Review of Financial Analysis*, *43*, 135–145. <https://doi.org/10.1016/j.irfa.2015.11.009>
- Emel, A. B., Oral, M., Reisman, A., & Yolalan, R. (2003). A credit scoring approach for the commercial banking sector. *Socio-Economic Planning Sciences*, *37*(2), 103–123.  
[https://doi.org/10.1016/S0038-0121\(02\)00044-7](https://doi.org/10.1016/S0038-0121(02)00044-7)
- Emrouznejad, A., Anouze, A. L., & Thanassoulis, E. (2010). A semi-oriented radial measure for measuring the efficiency of decision making units with negative data, using DEA. *European Journal of Operational Research*, *200*(1), 297–304.  
<https://doi.org/10.1016/j.ejor.2009.01.001>
- Emrouznejad, A., & Yang, G. L. (2018). A survey and analysis of the first 40 years of scholarly literature in DEA: 1978–2016. *Socio-Economic Planning Sciences*, *61*, 4–8.  
<https://doi.org/10.1016/j.seps.2017.01.008>
- Etemadi, H., Anvary Rostamy, A. A., & Dehkordi, H. F. (2009). A genetic programming model for bankruptcy prediction: Empirical evidence from Iran. *Expert Systems with Applications*, *36*(2), 3199–3207. <https://doi.org/10.1016/j.eswa.2008.01.012>

- Fang, H. H., Lee, H. S., Hwang, S. N., & Chung, C. C. (2013). A slacks-based measure of super-efficiency in data envelopment analysis: An alternative approach. *Omega*, *41*(4), 731–734. <https://doi.org/10.1016/j.omega.2012.10.004>
- Färe, R., Grosskopf, S., Norris, M., & Zhang, Z. (1994). American Economic Association Productivity Growth, Technical Progress, and Efficiency Change in Industrialized Countries. *Source: The American Economic Review*, *84*(1), 66–83. <https://doi.org/10.2307/2951340>
- Färe, R., Lindgren, B., & Roos, P. (1992). Productivity changes in Swedish pharmacies 1980–1989: A non-parametric malmquist approach. In T. R. G. Jr & C. A. K. Lovell (Eds.), *The Journal of Productivity Analysis* (Vol. 3, pp. 85–101). Springer Netherlands.
- Fethi, M. D., & Pasiouras, F. (2010). Assessing bank efficiency and performance with operational research and artificial intelligence techniques: A survey. *European Journal of Operational Research*, *204*(2), 189–198. <https://doi.org/10.1016/j.ejor.2009.08.003>
- Frydman, H., Altman, E. I., & Kao, D. L. (1985). Introducing recursive partitioning for financial classification: the case of financial distress. *Journal of Finance*, *40*(1), 269–291.
- Gestel, T. Van, Baesens, B., Suykens, J. A. K., Van den Poel, D., Baestaens, D. E., & Willekens, M. (2006). Bayesian kernel based classification for financial distress detection. *European Journal of Operational Research*, *172*(3), 979–1003. <https://doi.org/10.1016/j.ejor.2004.11.009>
- Guenther, N., & Schonlau, M. (2016). Support vector machines. *Stata Journal*, *16*(4), 917–937. <https://doi.org/10.1016/j.aca.2011.07.027>
- Hernandez Tinoco, M., & Wilson, N. (2013). Financial distress and bankruptcy prediction among listed companies using accounting, market and macroeconomic variables. *International Review of Financial Analysis*, *30*, 394–419. <https://doi.org/10.1016/j.irfa.2013.02.013>
- Huang, Z., Chen, H., Hsu, C. J., Chen, W. H., & Wu, S. (2004). Credit rating analysis with support vector machines and neural networks: A market comparative study. *Decision Support Systems*, *37*(4), 543–558. [https://doi.org/10.1016/S0167-9236\(03\)00086-1](https://doi.org/10.1016/S0167-9236(03)00086-1)
- Jain, A. (1997). Feature selection: evaluation, application, and small sample performance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *19*(2), 153–158. <https://doi.org/10.1109/34.574797>
- Kim, M. J., & Kang, D. K. (2010). Ensemble with neural networks for bankruptcy prediction. *Expert Systems with Applications*, *37*(4), 3373–3379. <https://doi.org/10.1016/j.eswa.2009.10.012>
- Laitinen, E. K., & Suvas, A. (2016). The effect of national culture on financial distress prediction modelling: evidence from European countries. *International Journal of Accounting and Finance*, *6*(4), 299. <https://doi.org/10.1504/IJAF.2016.082310>
- Li, H., & Sun, J. (2009). Predicting business failure using multiple case-based reasoning combined with support vector machine. *Expert Systems with Applications*, *36*(6), 10085–10096. <https://doi.org/10.1016/j.eswa.2009.01.013>



- Li, H., & Sun, J. (2011). Predicting business failure using forward ranking-order case-based reasoning. *Expert Systems with Applications*, 38(4), 3075–3084. <https://doi.org/10.1016/j.eswa.2010.08.098>
- Li, Z., Crook, J., & Andreeva, G. (2014). Chinese companies distress prediction: an application of data envelopment analysis. *Journal of the Operational Research Society*, 65(3), 466–479. <https://doi.org/10.1057/jors.2013.67>
- Li, Z., Crook, J., & Andreeva, G. (2017). Dynamic prediction of financial distress using Malmquist DEA. *Expert Systems with Applications*, 80, 94–106. <https://doi.org/10.1016/j.eswa.2017.03.017>
- Liang, D., Lu, C. C., Tsai, C. F., & Shih, G. A. (2016). Financial ratios and corporate governance indicators in bankruptcy prediction: A comprehensive study. *European Journal of Operational Research*, 252(2), 561–572. <https://doi.org/10.1016/j.ejor.2016.01.012>
- Lyandres, E., & Zhdanov, A. (2013). Investment opportunities and bankruptcy prediction. *Journal of Financial Markets*, 16(3), 439–476. <https://doi.org/10.1016/j.finmar.2012.10.003>
- Malmquist, S. (1953). Index numbers and indifference surfaces. *Trabajos de Estadística*, 4(2), 209–242. <https://doi.org/10.1007/BF03006863>
- Martin, D. (1977). Early warning of bank failure. A logit regression approach. *Journal of Banking and Finance*, 1(3), 249–276. [https://doi.org/10.1016/0378-4266\(77\)90022-X](https://doi.org/10.1016/0378-4266(77)90022-X)
- McKee, T. E., & Lensberg, T. (2002). Genetic programming and rough sets: A hybrid approach to bankruptcy classification. *European Journal of Operational Research*, 138(2), 436–451. [https://doi.org/10.1016/S0377-2217\(01\)00130-8](https://doi.org/10.1016/S0377-2217(01)00130-8)
- Min, J. H., & Lee, Y. C. (2008). A practical approach to credit scoring. *Expert Systems with Applications*, 35(4), 1762–1770. <https://doi.org/10.1016/j.eswa.2007.08.070>
- Mousavi, M. M., & Ouenniche, J. (2018, March 19). Multi-criteria ranking of corporate distress prediction models: empirical evaluation and methodological contributions. *Annals of Operations Research*, pp. 1–34. <https://doi.org/10.1007/s10479-018-2814-2>
- Mousavi, M. M., Ouenniche, J., & Xu, B. (2015). Performance evaluation of bankruptcy prediction models: An orientation-free super-efficiency DEA-based framework. *International Review of Financial Analysis*, 42, 64–75. <https://doi.org/10.1016/j.irfa.2015.01.006>
- Mukhopadhyay, A., Tiwari, S., Narsaria, A., & Karmaker, B. R. (2012). A New Approach to Predicting Bankruptcy: Combining DEA and New Approach to Predicting Bankruptcy: Combining DEA and Multi-Layer Perceptron Layer Perceptron Layer Perceptron. *International Journal of Computer Science Issues (IJCSI)*, 9(4), 71–78. Retrieved from <http://ijcsi.org/papers/IJCSI-9-4-2-71-78.pdf>
- Nam, C. W., Kim, T. S., Park, N. J., & Lee, H. K. (2008). Bankruptcy prediction using a discrete-time duration model incorporating temporal and macroeconomic dependencies. *Journal of Forecasting*, 27(6), 493–506. <https://doi.org/10.1002/for.985>
- Ohlson, J. A. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of*

- Accounting Research*, 18(1), 109. <https://doi.org/10.2307/2490395>
- Oral, M., & Yolalan, R. (1990). An empirical study on measuring operating efficiency and profitability of bank branches. *European Journal of Operational Research*, 46(3), 282–294. [https://doi.org/10.1016/0377-2217\(90\)90002-S](https://doi.org/10.1016/0377-2217(90)90002-S)
- Ouenniche, J., & Tone, K. (2017). An out-of-sample evaluation framework for DEA with application in bankruptcy prediction. *Annals of Operations Research*, 254(1–2), 235–250. <https://doi.org/10.1007/s10479-017-2431-5>
- Paradi, J. C., Asmild, M., & Simak, P. C. (2004). Using DEA and worst practice DEA in credit risk evaluation. *Journal of Productivity Analysis*, 21(2), 153–165. <https://doi.org/10.1023/B:PROD.0000016870.47060.0b>
- Paradi, J. C., & Zhu, H. (2013). A survey on bank branch efficiency and performance research with data envelopment analysis. *Omega (United Kingdom)*, 41(1), 61–79. <https://doi.org/10.1016/j.omega.2011.08.010>
- Pastor, J. T., & Lovell, C. A. K. (2005). A global Malmquist productivity index. *Economics Letters*, 88(2), 266–271. <https://doi.org/10.1016/j.econlet.2005.02.013>
- Pille, P., & Paradi, J. C. (2002). Financial performance analysis of Ontario (Canada) Credit Unions: An application of DEA in the regulatory environment. *European Journal of Operational Research*, 139(2), 339–350. [https://doi.org/10.1016/S0377-2217\(01\)00359-9](https://doi.org/10.1016/S0377-2217(01)00359-9)
- Pindado, J., Rodrigues, L., & de la Torre, C. (2008). Estimating financial distress likelihood. *Journal of Business Research*, 61(9), 995–1003. <https://doi.org/10.1016/j.jbusres.2007.10.006>
- Portela, M. C. A. S., Thanassoulis, E., & Simpson, G. (2004). Negative data in DEA: A directional distance approach applied to bank branches. In *Journal of the Operational Research Society* (Vol. 55, pp. 1111–1121). <https://doi.org/10.1057/palgrave.jors.2601768>
- Premachandra, I. M., Bhabra, G. S., & Sueyoshi, T. (2009). DEA as a tool for bankruptcy assessment: A comparative study with logistic regression technique. *European Journal of Operational Research*, 193(2), 412–424. <https://doi.org/10.1016/j.ejor.2007.11.036>
- Premachandra, I. M., Chen, Y., & Watson, J. (2011). DEA as a tool for predicting corporate failure and success: A case of bankruptcy assessment. *Omega*, 39(6), 620–626. <https://doi.org/10.1016/j.omega.2011.01.002>
- Psillaki, M., Tsolas, I. E., & Margaritis, D. (2010). Evaluation of credit risk based on firm performance. *European Journal of Operational Research*, 201(3), 873–881. <https://doi.org/10.1016/j.ejor.2009.03.032>
- Ravi Kumar, P., & Ravi, V. (2007). Bankruptcy prediction in banks and firms via statistical and intelligent techniques - A review. *European Journal of Operational Research*, 180(1), 1–28. <https://doi.org/10.1016/j.ejor.2006.08.043>
- Seballos, L. D., & Thomson, J. B. (1990). Underlying Causes of Commercial Bank Failures in the 1980s. *Economic Commentary*, (Sep).

- Sharp, J. A., Meng, W., & Liu, W. (2007). A modified slacks-based measure model for data envelopment analysis with “natural” negative outputs and inputs. *Journal of the Operational Research Society*, 58(12), 1672–1677. <https://doi.org/10.1057/palgrave.jors.2602318>
- Sharpe, W. F. (1966). Mutual Fund Performance. *The Journal of Business*. <https://doi.org/10.1086/294846>
- Sharpe, W. F. (1994). The Sharpe Ratio. *The Journal of Portfolio Management*. <https://doi.org/10.3905/jpm.1994.409501>
- Shetty, U., Pakkala, T. P. M., & Mallikarjunappa, T. (2012). A modified directional distance formulation of DEA to assess bankruptcy: An application to IT/ITES companies in India. *Expert Systems with Applications*, 39(2), 1988–1997. <https://doi.org/10.1016/j.eswa.2011.08.043>
- Shin, K. S., Lee, T. S., & Kim, H. J. (2005). An application of support vector machines in bankruptcy prediction model. *Expert Systems with Applications*, 28(1), 127–135. <https://doi.org/10.1016/j.eswa.2004.08.009>
- Shin, K. S., & Lee, Y. J. (2002). A genetic algorithm application in bankruptcy prediction modeling. *Expert Systems with Applications*, 23(3), 321–328. [https://doi.org/10.1016/S0957-4174\(02\)00051-9](https://doi.org/10.1016/S0957-4174(02)00051-9)
- Shumway, T. (2001). Forecasting Bankruptcy More Accurately: A Simple Hazard Model. *The Journal of Business*, 74(1), 101–124. <https://doi.org/10.1086/209665>
- Simak, P. C. (1997). *DEA Based Analysis of Corporate Failure*. University of Toronto.
- 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](https://doi.org/10.1016/S0377-2217(98)00014-9)
- 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](https://doi.org/10.1016/S0377-2217(00)00054-0)
- Sueyoshi, T. (2004). Mixed integer programming approach of extended DEA-discriminant analysis. *European Journal of Operational Research*, 152(1), 45–55. [https://doi.org/10.1016/S0377-2217\(02\)00657-4](https://doi.org/10.1016/S0377-2217(02)00657-4)
- Sueyoshi, T. (2006). DEA-Discriminant Analysis: Methodological comparison among eight discriminant analysis approaches. *European Journal of Operational Research*, 169(1), 247–272. <https://doi.org/10.1016/j.ejor.2004.05.025>
- Sueyoshi, T., & Goto, M. (2009). DEA-DA for bankruptcy-based performance assessment: Misclassification analysis of Japanese construction industry. *European Journal of Operational Research*, 199(2), 576–594. <https://doi.org/10.1016/j.ejor.2008.11.039>
- Sueyoshi, T., Goto, M., & Omi, Y. (2010). Corporate governance and firm performance: Evidence from Japanese manufacturing industries after the lost decade. *European Journal of Operational Research*, 203(3), 724–736. <https://doi.org/10.1016/j.ejor.2009.09.021>

- Sun, X., Liu, Y., Xu, M., Chen, H., Han, J., & Wang, K. (2013). Feature selection using dynamic weights for classification. *Knowledge-Based Systems*, 37, 541–549. <https://doi.org/10.1016/j.knosys.2012.10.001>
- Tang, T. C., & Chi, L. C. (2005). Neural networks analysis in business failure prediction of Chinese importers: A between-countries approach. *Expert Systems with Applications*, 29(2), 244–255. <https://doi.org/10.1016/j.eswa.2005.03.003>
- Tone, K. (2002). A slacks-based measure of super-efficiency in data envelopment analysis. *European Journal of Operational Research*, 143, 32–41.
- Tone, K. (2011). Slacks-Based measure of efficiency. *International Series in Operations Research and Management Science*, 164(June), 195–209. [https://doi.org/10.1007/978-1-4419-6151-8\\_8](https://doi.org/10.1007/978-1-4419-6151-8_8)
- Trujillo-Ponce, A., Samaniego-Medina, R., & Cardone-Riportella, C. (2014). Examining what best explains corporate credit risk: accounting-based versus market-based models. *Journal of Business Economics and Management*, 15(2), 253–276. <https://doi.org/10.3846/16111699.2012.720598>
- Tsai, M. C., Lin, S. P., Cheng, C. C., & Lin, Y. P. (2009). The consumer loan default predicting model - An application of DEA-DA and neural network. *Expert Systems with Applications*, 36(9), 11682–11690. <https://doi.org/10.1016/j.eswa.2009.03.009>
- Wahid, H., Ahmad, S., Nor, M. A. M., & Rashid, M. A. (2017). Prestasi kecekapan pengurusan kewangan dan agihan zakat: perbandingan antara majlis agama islam negeri di Malaysia. *Jurnal Ekonomi Malaysia*, 51(2), 39–54. <https://doi.org/10.1017/CBO9781107415324.004>
- Wang, G., Ma, J., & Yang, S. (2014). An improved boosting based on feature selection for corporate bankruptcy prediction. *Expert Systems with Applications*, 41(5). <https://doi.org/10.1016/j.eswa.2013.09.033>
- Wu, Y., Gaunt, C., & Gray, S. (2010). A comparison of alternative bankruptcy prediction models. *Journal of Contemporary Accounting and Economics*, 6(1), 34–45. <https://doi.org/10.1016/j.jcae.2010.04.002>
- Xu, X., & Wang, Y. (2009). Financial failure prediction using efficiency as a predictor. *Expert Systems with Applications*, 36(1), 366–373. <https://doi.org/10.1016/j.eswa.2007.09.040>
- Yeh, C. C., Chi, D. J., & Hsu, M. F. (2010). A hybrid approach of DEA, rough set and support vector machines for business failure prediction. *Expert Systems with Applications*, 37(2), 1535–1541. <https://doi.org/10.1016/j.eswa.2009.06.088>
- Zhou, L. (2013). Performance of corporate bankruptcy prediction models on imbalanced dataset: The effect of sampling methods. *Knowledge-Based Systems*, 41(0), 16–25. <https://doi.org/10.1016/j.knosys.2012.12.007>
- Zhou, L., Lu, D., & Fujita, H. (2015). The performance of corporate financial distress prediction models with features selection guided by domain knowledge and data mining approaches. *Knowledge-Based Systems*, 85, 52–61. <https://doi.org/10.1016/j.knosys.2015.04.017>
- Zopounidis, C., & Doumpos, M. (2002). Multi-criteria decision aid in financial decision making:

Methodologies and literature review. *Journal of Multi-Criteria Decision Analysis*, 11(4–5), 167–186. <https://doi.org/10.1002/mcda.333>