Expert Systems with Applications 40 (2013) 5932–5944

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

The application of brute force logistic regression to corporate credit scoring models: Evidence from Serbian financial statements

Nebojsa Nikolic *, Nevenka Zarkic-Joksimovic, Djordje Stojanovski, Iva Joksimovic

University of Belgrade, Faculty of Organizational Sciences, Jove Ilica 154, Belgrade, Serbia

ARTICLE INFO

Keywords:
Credit scoring
Logistic regression
Probability of default
Weight of evidence approach
Corporate entities

ABSTRACT

In this paper a brute force logistic regression (LR) modeling approach is proposed and used to develop predictive credit scoring model for corporate entities. The modeling is based on 5 years of data from end-of-year financial statements of Serbian corporate entities, as well as, default event data. To the best of our knowledge, so far no relevant research about predictive power of financial ratios derived from Serbian financial statements has been published. This is also the first paper that generated 350 financial ratios to represent independent variables for 7590 corporate entities default predictions'. Many of derived financial ratios are new and were not discussed in literature before. Weight of evidence (WOE) method has been applied to transform and prepare financial ratios for brute force LR fitting simulations. Clustering method has been utilized to reduce long list of variables and to remove highly correlated financial ratios from partitioned training and validation datasets. The clustering results have revealed that number of variables can be reduced to short list of 24 financial ratios which are then analyzed in terms of default event predictive power. In this paper we propose the most predictive financial ratios from financial statements of Serbian corporate entities. The obtained short list of financial ratios has been used as a main input for brute force LR model simulations. According to literature, common practice to select variables in final model is to run stepwise, forward or backward LR. However, this research has been conducted in a way that the brute force LR simulations have to obtain all possible combinations of models that comprise of 5–14 independent variables from the short list of 24 financial ratios. The total number of simulated resulting LR models is around 14 million. Each model has been fitted through extensive and time consuming brute force LR simulations using SAS® code written by the authors. The total number of 342,016 simulated models ("well-founded" models) has satisfied the established credit scoring model validity conditions. The well-founded models have been ranked according to GINI performance on validation dataset. After all well-founded models have been ranked, the model with highest predictive power and consisting of 8 financial ratios has been selected and analyzed in terms of receiver-operating characteristic curve (ROC), GINI, AIC, SC, LR fitting statistics and correlation coefficients. The financial ratio constituents of that model have been discussed and benchmarked with several models from relevant literature.

1. Introduction

Credit scoring models play an important role in contemporary banking risk management practice. They contribute to the key requirement in loan approval process, which is to accurately and efficiently quantify the level of credit risk associated with a customer. The credit scoring models objective is to predict future behavior in terms of credit risk by relying on past experience of customers with similar characteristics. The level of credit risk of a borrower is associated with probability that it will default on approved loan over given time horizon, usually 1 year. The main task of credit scoring model is to provide discrimination between the ones who do default and the ones who do not, i.e. between good and bad corporate entities in terms of their creditworthiness. Discrimination ability is the key indicator of model successfulness. The higher the discrimination power the more precise the credit scoring model will be.

The models can be established on judgmental basis or with support of statistical tools. Judgmental or expert-based models are established through set of formal 'rule-of-thumb' quantitative criteria. It is an easiest way to incorporate the best practices and the knowledge of credit managers into formal automated decision rules. On the contrast, statistical scoring models are built upon optimization algorithm which is applied on historical data of credit

* Corresponding author. Address: Svetozara Markovica 71/13, Belgrade, Serbia. Tel.: +381 603584674.

E-mail address: nebojsa.nikolic.bg@gmail.com (N. Nikolic).

0957-4174/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.eswa.2013.05.022

performance of both good and bad customers. For an extensive review of statistical methods and their credit scoring application we refer to study of Crook, Edelman, and Thomas (2007).

Contemporary risk management practice emphasizes and promotes the use of credit scoring models for various asset classes of bank's credit portfolio (BCBS., 2006). Retail banking practice uses application and behavioral credit scoring models for automation of loan approval process for individuals (Kennedy, Namee, Delany, O'Sullivan, & Watson, 2013; Sustersic, Mramor, & Zupan, 2009). By employing process automation, the bank's staff costs are reduced, loan approval process is simplified, speeded up and more control on approval decision making process is attained (Blochlinger & Leippold, n.d.). In retail banking decision to grant a loan based on fundamental analysis and credit analyst assessment is left to be applied only for high amount or non-standard loans.

Latest Serbian credit bureau data report states that about 68% of banking loan exposures belongs to corporate entities (ASB., 2013). In process of financial statement analysis, in order to evaluate the financial health of a corporate entity the financial ratios are commonly used as a part of fundamental analysis. Granting loans to corporate entities based solely on credit scoring models is generally performed only for smaller loan amounts and for particular standardized loan products. More often, credit scoring models are used as an additional tool or decision making criteria. Credit scoring models based on logistic regression (LR) modeling technique provide the results in form of probability of default (PD) level for particular corporate entity. The PD level represents a quantitative estimate of credit risk inherited into corporate entity and it plays a valuable role in credit risk assessment within bank. The PD estimate has been referred to as one of the main and most widely used risk factor in Basel II era (Pluto & Tasche, 2010). The main business utilization of estimated PD serves for calculating: expected and unexpected losses, statistical rating of corporate entities, loan loss provisions and cost of risk component of interest rate (Altman & Sabato, 2007; Ruthenberg & Landskroner, 2008). Applying a credit scoring model with a higher discrimination power could result in lower capital requirements and more accurate PD estimates (Altman & Sabato, 2007).

The Basel II standards recommend portfolio segmentation[1] of corporate entities based on sales (BCBS, 2006). Many banks already follow this segmentation practice when modeling credit risk, but in academic literature according to Altman and Sabato (2007) a study that reveals significant benefit of such choice is lacking. The segmentation study of Bijak and Thomas (2012) has shown that segmentation does not always improve credit scoring model performance. Governed with this idea, the underlying principle of our study was to treat all corporate entities as one segment. The goal was to try to build "one-size-fits-all" credit scoring model for whole population of corporate entities.

According to relevant literature, two mostly used statistical methods for building credit scoring models for corporate entities are discriminant analysis (DA) and logistic regression (LR). The first attempt to link financial ratios of corporate entities to risk in terms of probability of bankruptcy was done by Beaver (1967). The well-known *Z*-score model, developed on small corporate entities sample by Altman (1968) and afterwards enhanced by Altman, Haldeman, and Narayanan (1977), was the first credit risk model to predict default probabilities of corporate entities using DA technique. For the first time LR was applied in default prediction study of (Ohlson (1980). The main benefits of LR over DA were emphasized in terms of less restrictive modeling assumptions. The linearity, normality conditions, as well as, independence among independent variables is not assumed in LR approach which leaves more flexibility in working with real-life data. The first reported LR prediction results were of less predictive power than the ones reported in DA studies. Later on, studies have shown that LR is a sound and powerful statistical approach for modeling credit risk. Further researches of models for predicting business failures using LR are discussed and implemented by Johnsen and Melicher (1994), Dimitras, Zanakis, and Zopounidis (1996), Laitinen and Laitinen (2000), Becchetti and Sierra (2003) Westgaard and Wijst (2001), Altman and Sabato (2007), Kumar and Ravi (2007) and Chen (2011).

In the last decade the extensive development of credit scoring models has been done. Default prediction as a classification problem entails forecast of corporate entity failure likelihood given a number of independent variables in terms of financial ratios (Altman & Sabato, 2007; Fantazzini & Figini, 2009; Westgaard & Wijst, 2001). Credit scoring models were first built on data from developed world economies and only later they started to utilize data from different emerging markets. The study of Zekic-Susac, Sarlija, and Bensic (2004) compared LR results with other different estimation methodologies on Croatian bank dataset. The paper of Hermanto and Gunawidjaja (2010) tested the performance of LR model on Indonesian SME data over the period of 2005–2007. The LR study performed on 700 SME loans in Slovakia between 2000 and 2005 pointed out that liquidity and profitability factors are important determinants of SME defaults (Fidrmuc & Hainz, 2010). The recent research of Louzada, Ferreira-Silva, and Diniz (2012) tried to reveal the LR models performance on state-dependent sample extracted from a portfolio of a Brazilian bank. Furthermore, the research of Jain, Gupta, and Sanjiv (2011) examined the behavior of default risk measures and explored the most significant financial variables for SMEs using LR technique. For the purpose of mentioned research, the Indian database of about 3000 SMEs has been used, covering years from 2007 to 2009. Another research, based on Korean dataset (Sohn & Kim, 2012) tried to reveal the best behavioral credit scoring model for technology-based SMEs. The behavioral scoring results have been revealed and compared to its application credit scoring counterpart. Finally, in the most recent study of Blanco, Pino-Mejias, Lara, and Rayo (2013) compared LR results with other non-parametric techniques, based on a sample of almost 5500 microfinance borrowers from Peru.

Recent studies for corporate entities show that beside financial ratios there is potential value added, in prediction power terms, when economic, environmental and non-financial information are included in the model as a default predictors (Blanco et al., 2013; Moon & Sohn, 2010).

Even with the existence of more sophisticated classification models for credit scoring, such as neural networks (Derelioglu & Gurgen, 2011; Lee, Han, & Kwon, 1996; Leshno & Spector, 1996), support vector machines (Kim & Ahn, 2012) and case based reasoning (Vukovic, Delibasic, Uzelac, & Suknovic, 2012) the popularity and usage of LR has continued mostly due to its practicality and theoretical soundness.

To the best of our knowledge, this is the first study that has examined all possible combinations of the models given the short list of financial ratios as input variables. In comparison to other studies we have generated long list of 350 financial variables and then tailored the principal component clustering technique in order to reduce this long list to a short list of 24 variables. We examined in details the predictive power of financial ratios as standalone variables, as well as, the all possible combinations of models that include 5–14 financial ratio variables.[2] The total num-

[1] Small medium enterprises (SME) segment is for sales less than €50 million, corporate (CO) over €50 million and large corporate (LC) entities over €500 million.

[2] Due to computational reasons the maximum number of variables considered in model has been set to 14. Analysis that follows shows this to be more than sufficient for identifying the most suitable model.