Knowledge-Based Systems 70 (2014) 78–87

Contents lists available at ScienceDirect

Knowledge-Based Systems

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

A financial early warning logit model and its efficiency verification approach

Shuangjie Li [a,*,1], Shao Wang [b]

[a] Economics & Management School, Beijing University of Technology, Pingleyuan 100, Chaoyang District, Beijing 100124, China
[b] China Security Co., Ltd., Beijing, China

ARTICLE INFO

Article history:
Available online 3 April 2014

Keywords:
Financial early warning
Logit model
Prediction
Data envelopment analysis
Efficiency

ABSTRACT

Financial early warning (FEW) models aim to help companies recognize possible financial crises and reduce financial risks through generating FEW messages. The current FEW models are mainly constructed by a set of financial indicators and the predictive accuracy has only been verified by these financial indicators rather than non-financial indicators. The issue in such a situation is that these financial indicators can be controlled or manipulated by related senior managerial personnel of companies, and therefore, using only financial indicators to verify FEW models cannot ensure the reliability of predictive accuracy of the models. To handle this issue, this paper develops a new FEW logit model which has better predictive accuracy than existing ones. More importantly, we propose a new approach which verifies the predictive accuracy of the logit model by using non-financial efficiency indicators of data envelopment analysis. An empirical study on Chinese company datasets revealed that the accuracy rates of predictions of the proposed model, for in-sample and out-of-sample companies, are 97.1% and 94.1% respectively, higher than existing results. Using non-financial efficiency indicators, the verification rates for the prediction results of the logit model for in-sample and out-of-sample companies are 95.8% and 96.2%, respectively. The findings show that the proposed FEW logit model has improved the accuracy of prediction and stability; the approach which uses non-financial efficiency indicators to verify the results of FEW logit model has significantly ensured the reliability of the FEW models.

1. Introduction

The financial crises which occurred frequently needed related financial early warning (FEW) models to recognize possible crises and therefore reduce financial risks. The Asian financial crisis of 1997 remains fresh in one's memory, as well as the global financial crisis of 2007–2008, which has not yet come to an end. The companies all over the world that find their financial indicators have deteriorated greatly before taking financial saving measure, and are often caught off guard and find themselves heading towards bankruptcy. Nowadays, more people realize that the operation of the whole company must build a sound financial risk prediction system in order to recognize the pending financial crisis, to take actions at the inception of a financial crisis, and to prevent further deterioration. A main component of building a financial risk prediction system is to build financial risk prediction models [34,48].

Various statistical FEW techniques including the single-variable FEW model [11,12,15,56], multivariable linear FEW model [2,4,59], linear probability model [24,42,45,52], logit model [14,16,20,23,39,43,44,53] and Probit model [7,25,46,54,59] have been developed and well applied in real world cases. Particularly, using the logit models to solve FEW problems has received much attention from researchers worldwide [30,40,50] and is strongly recommended by many researchers [20,21,32,50]. Those logit models use a variety of key financial indicators, such as asset liability ratio, financial leverage, and flow rate to forecast the financial situation of companies in the future.

In general, artificial intelligence methods such as artificial neural networks (ANN) have higher predictive accuracy than those traditional statistical techniques [49], but artificial intelligence methods have some shortcomings. Altman [3] pointed out the "black box" problems, such as illogical weightings of the indicators and over fitting in the training stage which resulted in a negative impact on predictive accuracy. Those choice-based sampling techniques may result in non-random samples and hence produce biased estimates [60]. Boritz [13] examined predictive accuracy of two different ANN learning algorithms to a number of statistical

* Corresponding author. Tel.: +86 13910657085; fax: +86 1067391993.
E-mail addresses: lishuangjie@bjut.edu.cn (S. Li), wangshao@csc.com.cn (S. Wang).
[1] Supported by Beijing Social Science Fund (12JGA005).

http://dx.doi.org/10.1016/j.knosys.2014.03.017
0950-7051/

techniques. They concluded that ANNs appeared to have a higher predictive accuracy but they were sensitive to the proportion of bankrupt firms in the training samples. ANN studies were based on the back propagation algorithm. There are typically two difficulties with the use of the back propagation algorithm: one is that the algorithm requires much time to obtain a good approximation, the other is that the gradient-based searching method of the back propagation algorithm is easy to trap in a local minimum. Kim and Han [35] found that the genetic algorithm approach was limited by the validation of the training examples, which were obtained through experts' subjective ratings for companies. Such ratings could not be consistent from expert to expert, and hence affected the validation of the learning result.

In the past 20 years, many studies have investigated the performance of intelligent models on FEW. However, the two statistical models (logit model and multivariate discriminant analysis) are still very popular (particularly for industrial use) because they are well-known models for FEW predictions and are easy to model, interpret, and explain. Various logit models are used more frequently and more widely because they are less demanding than multivariate discriminant analysis [33]. This paper proposed a logit model to predict financial risk because (1) logit model has been widely used and taught, it is a non-linear probability model which is suitable for FEW since finance prediction is always non-linear [8]; (2) logit model is relatively easy to understand and readily available in virtually all software packages; it does not assume multivariate normality, but gives a crisp relationship between explanatory and response variables based on the given data [56]. Logit models easily explain the reasons why a financial crisis will or will not occur [22]; (3) logit models have a higher accuracy and stability than many other models; (4) logit models don't require rich, quality data and, therefore, are more applicable than intelligent techniques in the situations which do not have high quality data, such as in the finance industry of China and other developing countries.

In this paper, we develop a new FEW logit model by using the forward screening Wald method which makes variables enter sequentially into the logit model in a stepwise manner. In the case of a default value, variables with a significant level of 0.05 or smaller are to be retained, otherwise removed. Therefore, the Wald method can ultimately ensure that variables remaining in the model have a significant impact on the predictive results and ensure the simplicity of the model, since only the most important variables are selected. The empirical study reveals that the accuracy rates of prediction of the logit model for in-sample and out-of-sample companies are 97.1% and 94.1% respectively. The model has a highly predictive accuracy for in-sample companies. The high accuracy rate of prediction for out-of-sample companies shows the stability of the model.

The current FEW logit models are constructed by financial indicators. The accuracy of these models has not been verified by studies from the perspective of non-financial indicators. Although our model has better predictive accuracy, it cannot guarantee reliable prediction for other data sets. Therefore, we then propose a new verifying approach which uses non-financial efficiency indicators from a data envelopment analysis (DEA) model to verify the prediction of the proposed FEW logit models. The DEA efficiency is measured by the ratio of physical output and input. It is a physical indicator (for example: input 3 employees, 1 bread machine, output 80 boxes of bread) rather than a financial indicator (for example: total profit $1000, return on assets = total profit/total assets = 9%). The efficiency indicator is closer to the operation of companies than the financial indicator. It is truer than the financial indicator because it is harder to control and manipulate. As far as we know, as a non-financial indicator, efficiency indicator of DEA is rarely used to verify the prediction of the financial crisis for companies except the studies of Huang [31] and Lin [37]. In the study of Huang [31], the author used the financial indicators and efficiency indicator to predict the business crisis. The author found that the logit model was better than the ANN approach. Lin [37] used the methods of DEA-discriminate analysis, neural network and logistic regression to establish the prediction models of financial crises. The results showed that the DEA-discriminate analysis approach was the highest prediction approach. Our paper aims to use DEA efficiency as a key indicator to make some supplementary verification to the prediction of the logit model. Using efficiency indicators, the results of the logit model for in-sample and out-of-sample companies are verified 95.8% and 96.2% respectively. This indicates that the study has valuable and innovative novelty.

The main contributions of this paper include: (1) it develops a new FEW logit model which has a higher predictive accuracy and stability; (2) it proposes a new approach that uses non-financial efficiency indicators to verify the prediction of the logit model, this ensures the reliability of the predictive accuracy of the logit model.

The remaining sections of the paper are organized as follows: Section 2 gives a literature review of FEW logit models. In Section 3, a FEW logit model is constructed to test the in-sample and out-of-sample companies. Section 4 verifies the reliability of prediction of the logit model from the perspective of efficiency. In Section 5, we summarize the results and provide concluding remarks.

2. Literature review

Logit models are developed from many other statistical FEW models, which include the single-variable FEW model, multivariable linear FEW model, linear probability model, and Probit model.

The single variable model refers to the FEW model based on a single financial indicator, such as Return on equity or financial leverage. The model judges the financial difficulties of a company by the deterioration of the single indicator [9]. Chen [15] collected 27 matched pairs of ST (special treatment, negative net profits for two consecutive years) and non-ST Chinese companies on the same industry and of similar size, and then used a dichotomy with different financial indicators to predict ST. Randomly extracting 27 companies from a total of 54, he selected different values for each financial indicator, compared the total number of samples of the Type-I error and Type-II error, committing the least number of errors, as one determined the critical point. For each financial indicator, he obtained an annual rate of misjudgment in the last three years before special treatment: as the announcement of ST drew closer, the misjudgment rate was lower. The author concluded that the Asset-liability ratio, Current ratio and Return on assets were better financial indicators in a single variable FEW model. Furthermore, Wang and Zhang [56] analyzed the defects of single financial indicators. They found that the single financial indicators were easy to fake and whitewash for defects in the accounting system itself (historical cost, accounting method diversity of choice, accrual recognition of revenue and expenses). The authors added indicators of debt-paying ability, operational efficiency, and earnings quality in the financial indicators, and each area also included those three specific indicators, analyzing the case of Yin Guangxia (a public company of China). They found that those financial indicators could not reflect the company's operating situations. We think that although the advantage of the model of single financial indicator is simple, the drawbacks are more evident. Firstly, a single financial indicator only reflects one aspect of the company's entire production and operation change, such as asset utilization and liquidity. Hence, applying a single indicator to judge the financial situation may create defects in theory; secondly, single financial indicator are vulnerable to manipulation by business executives, especially some key indicators like net income and