

Construction of the Logistic Regression Estimation Model in Early Warning on Pure Financial Indicators

Zhang Lixin* and Wang Long

School of Economics, Wuhan University of Technology, Wuhan, China

Abstract: In order to establish a reasonable and effective financial crisis early warning model, the article chooses some financial indicators and uses factor analysis to get the common factors to conduct the most salient financial indicators. In the case that multirole linearity is not significant, the article uses the logistic regression to analyse eligible financial data and obtains the financial crisis warning model. Then the article found that this model has a high predictive accuracy.

Keywords: Logistic regression estimation, factor analysis, early warning, financial indicators, experimental analysis.

1. INTRODUCTION

In the 20th century, foreign scholars began to study the forecasting model of pure financial indicators [1]. The basic methods used to forecast the pure financial indicator is divided into qualitative and quantitative analysis [2]. Qualitative methods are analysis methods which mainly rely on subjective judgments of financial early warning. The representative quantitative analysis of early warning model are multivariate and univariate forecasting model which is based on the financial information and uses mathematical models or statistical methods to process and analyse on various financial data [2], and then establishes a model which reflects an enterprise's financial position. From the development course of studies, the academic uses the corresponding indicators of financial data and conducts a series of studies on the financial early-warning from the aspects of model construction and indicator choice, and it is divided into three main parts: (1) Single variable models. These models use the representative single financial indicator to predict the financial crisis [1]. Foreign research scholar Fitzpatrick (1932) was the first to use such methods. He found the two financial ratios had the lowest error rate: net profit/shareholders' equity and shareholders' equity/liabilities. Following Fitzpatrick's research, William Beaver studied 30 financial indicators of 158 firms (of which 79 companies are irregular enterprise, and another 79 companies are normal business) and finally got that the two indicators with highest forecast accuracy are debt guarantee rate and the asset-liability ratio [3]. (2) Multivariate model. These models use more than one indicator to conduct comprehensive analysis. Multivariable model is

Altman used multiple discriminant analysis on the 33 failed companies and 33 normal companies in America in the 1945-1965. He selected five indicators from the total 22

financial ratios indicators based on a large number of empirical observation and analysis and weighted totals to forecast financial situation of enterprises which has the accuracy up to 95% [2]. Finally, these findings formed the famous z-value model. (3) Multivariate conditional probability models. These models mainly seek the conditional probability of the observed object, and then determine financial status and operating conditions of the observed object including logistic and probity models [2]. Martin is the first to use the model to predict the bankruptcy of the company as well as the probability of default [5]. Olson used the multivariate logistic regression analysis in his study and concluded that corporate size, capital structure; debt size and current liquidity have more greatly effect on model predictive accuracy and use these four criteria to build the model prediction accuracy rate up to 96.12% [4].

The rest of the paper is organized as follows. In Section 2, early warning financial indicators is summarized briefly. In Section 3, the logistic regression estimation method is described. In Section 4, experiments are presented. And the results are discussed. Finally, a conclusion is provided in Section 5.

2. OVERVIEW OF EARLY WARNING FINANCIAL INDICATORS

Perfect financial crisis early warning system consists of the following four components: (1) financial crisis early warning organization mechanism. (2) Financial crisis early-warning information collection, collation and delivery mechanisms. (3) Financial crisis management mechanisms. (4) Accountability mechanism of the financial crisis.

Financial indicator analysis refers to a summary and assessment analysis of an enterprise's financial position and operating results of the index including solvency indicators, profitability indicators, operational capability and capacity development indicators [5, 6]. Solvency refers to the enterprise's ability to pay its debts. Solvency analysis

*Address correspondence to this author at the Wuhan University of Technology, Wuhan, China; Tel: + 86-136-2406-9677; Fax: + 86-136-2406-9677; E-mail: fanleiyang@126.com

includes analysis of short-term solvency and long-term solvency analysis. Operational capability analysis is to calculate the enterprise cash flow analysis of the relevant indicators of efficiency in the use of its assets which is the analysis of enterprise management level and asset management capabilities. Profitability indicators are the ability of funding increasing for enterprise. It is often presented by amounts the size and level of business income. We can analyse enterprise profitability from two aspects of profitability and social contribution [7]. Analysis can be obtained from the general analysis of two studies. Capacity development is the strength potential for expanding scale based on the existence of enterprises [8].

3. THE LOGISTIC REGRESSION ESTIMATION METHOD

3.1. Factor Analysis Method

Factor analysis is an important branch of multivariate statistical analysis. Its main purpose is to concentrate data. We can use the imaginary a few variables to represent the main information of the original variables through the study of correlation between the variables. Factor analysis was originally proposed by psychologist Spearman. Factor analysis has two main functions: one is seeking the basic structure, another is data reduction. We can extract the useful data and reject those useful data to significantly reduce the workload of data analysis through factor analysis [9].

In the relevant set of indicators, each indicator is determined by common factors and specific factors. Factor analysis is to identify the common factors of a set of indicators. Its basic model is as follows:

$$\begin{aligned}
 x_1 &= a_{11}f_1 + a_{12}f_2 + \dots + a_{1m}f_m + e_1 \\
 &\dots \dots \dots \\
 x_j &= a_{j1}f_1 + a_{j2}f_2 + \dots + a_{jm}f_m + e_j
 \end{aligned}
 \tag{1}$$

Indicates indicator, indicates common factor, and they are orthogonal to each other. Not all variables can be used for factor analysis. Factor analysis assumes that the variable should be relevant. If the variables are orthogonal, then there are no common factors between these variables and factor analysis would be unable to do. So we must test the correlation between variables before factor analysis. If they have high degree of relevance, factor analysis can be done. There are two variable correlation test methods: KMO sample measurement and Bartlett ball test [10].

3.2. Logistic Regression Analysis Method

Regression analysis is to study the impact between a variable and a set of variables to another variable. It is a mathematical tool for studying relationships between variables. In the past, most samples for establishment of financial early-warning model must be subject to multivariate normal distributions and it is not reasonable for the fact. Logistic regression model can avoid this problem which is a more useful model. It is broadly consistent with

the enterprise and the sample's assumptions are not so high, especially even when the sample is not normally distributed, we can also build the analysis model. It is a very effective way to tackle 0-1 return. Logistic regression is a qualitative variable regression analysis. Logistic regression model is as follows:

$$P(y = 1 / x) = \frac{1}{1 + e^{-\varepsilon}}
 \tag{2}$$

Where $\varepsilon = \alpha + \beta x$ is a series of linear function of the factors affecting the probability of an event occurring. Logit transformation of Logistic regression is also known as the natural logarithmic transformations which can convert the nonlinear function to a linear function.

$$\text{Logit}(y_j) = \text{Logit } p_j = \ln\left(\frac{p_j}{1 - p_j}\right) = \alpha + \beta x_j
 \tag{3}$$

For k independent variables, we can get:

$$\text{Logit } p = \alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_j x_j
 \tag{4}$$

The corresponding logistic regression model is as follows:

$$P = \frac{e^{\alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_j x_j}}{1 + e^{\alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_j x_j}}
 \tag{5}$$

Statistical inference of the logistic regression includes two aspects: significance test and parameter estimation. (1) Significance test. Significance test discusses significance testing for independent variables in the model on Loge including the Wald test, L.R. test and so on. (2) Parameters estimation.

Logistic regression models estimation is usually used maximum likelihood estimation method which uses the overall density, expressions of probability distribution and other information provided by samples to seek a method for estimating unknown parameters. It is the typical method of parameter estimation in statistical analysis.

4. EMPIRICAL ANALYSIS

In this section, the article selects 18 representatives' indicators of the financial health of the company as alternate early warning indicators and uses the logistic regression method in the process to pick 3 financial indicators which have significant effect on the dependent variable as shown in Fig. (1) and Table 1 to 5.

4.1. Choice of Financial Indicators

Pure financial ratios are hard to reveal why companies get into financial crises or special treatment. The usual practice is to select the different variables that represent different aspects of the financial characteristics of listed companies. Therefore, the article initially selects one or several indicators in the categories feature which reflect the company's financial health and creates a financial crisis early-warning index system and then uses regression models to select the highest accuracy combinations.

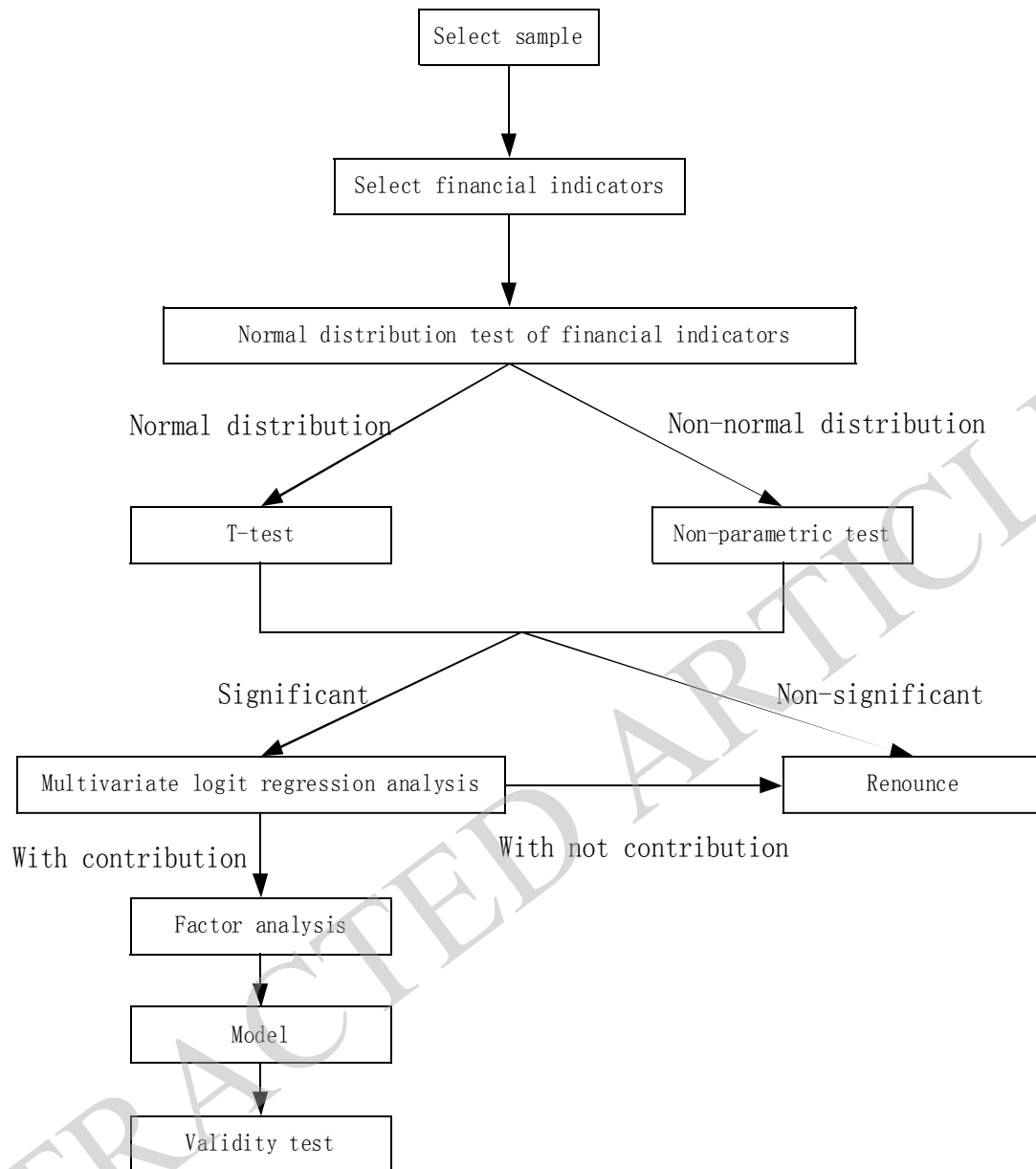


Fig.(1). Flow chart.

Table 1. Alternative financial early-warning indicators.

| Financial Analysis | Financial Indicators |
|----------------------------------|--|
| A: asset liquidity | X1 flow ratio X2 quick ratio |
| B: debt ratio | X3 working capital/total assets X4 assets and liabilities X5 debt ratio X6 interest coverage ratio X7 EPS |
| C: asset management capabilities | X8 total asset turnover X9 receivable accounts turnover X10 inventory turnover |
| D: profitability | X11 return on equity X12 return on asset X13 operating income net profit margin |
| E: company growth capabilities | X14 profit margin X15 net profit growth rate of assets X16 growth rate of capital X17 capital accumulation rate X18 main business revenue growth rate |

Table 2. Stepwise variable selection results.

| Step | Variable | | |
|------|----------|---------------------|--------------------|
| 1 | X12 | Secondary Education | Tertiary Education |
| 2 | X3 | 0.2 | 0.0 |
| 3 | X11 | 0.5 | 0.0 |

Table 3. Multi-collinearity test.

| Variable | X3 | | | | X12 | | | |
|----------|--------|--------|-------------|-------|----------------------------|----------------------------|---------------|-------|
| TOL | 0.5212 | 0.3276 | Blue-Collar | Other | Upper- Level White- Collar | Lower- Level White- Collar | Blue - Collar | Other |
| VIF | 1.9186 | 3.0525 | | | | | | |

Table.4. Maximum likelihood estimation.

| Variable | Estimated Value | | Standard Deviation | | Chi-square Score | | Probability | | Variable | |
|-----------|-----------------|---------|--------------------|--------|------------------|-------------|-------------|-------------|----------|-------------|
| Intercept | 4.8485 | 7.5958 | 0.4074 | 0.5233 | intercept | | | | | |
| X3 | - 36.0411 | 51.3619 | 0.4924 | 0.4829 | X3 | 1 | | 1 | | |
| X12 | 113.4 | 127.1 | 0.7957 | 0.3724 | X12 | (0.64-1.24) | 0.77 | (0.52-1.14) | 1.10 | (0.69-1.77) |
| X15 | - 142.8 | 153.3 | 0.8672 | 0.3517 | X15 | (1.31-2.47) | 1.89 | (1.33-2.70) | 2.55 | (1.66-3.91) |

Table 5. Logistic regression model results.

| Original value | | Predicted value | | Total | Error rate |
|----------------|---|-----------------|----|-------|------------|
| | | 0 | 1 | | |
| Count | 0 | 59 | 7 | 66 | 7.58% |
| | 1 | 3 | 63 | 66 | |

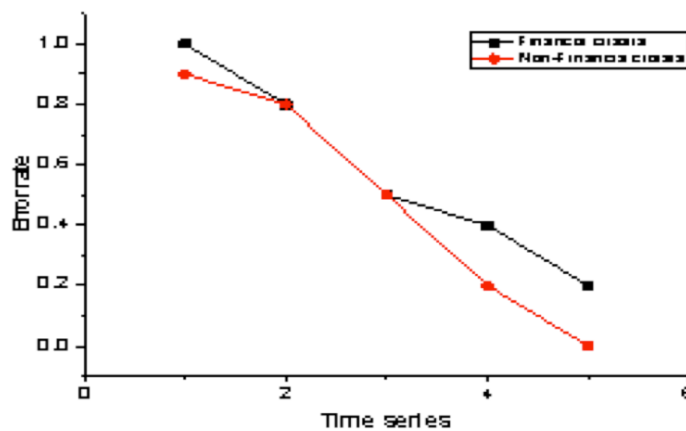


Fig. (2). Results on logistic regression model.

Before the financial crisis, we can know that there are 4 wrongly convicted of 72 non-financial crisis and 3 wrongly convicted of 72 financial crises. On the whole there are 7

wrongly convicted of 144 companies, and the error rate was 4.86%. It indicates the logistic model is accurate for predicting the financial crisis warning.

CONCLUSION

In this paper, a logistic regression model is developed for early warning on pure financial indicators as shown in Fig. (2). The paper applies the factor analysis and logistic regression analysis and uses listed company financial statements to disclose financial data. And then it conducts empirical research of financial risk early-warning and has achieved good results. The study results showed that logistic regression model for early-warning based on the financial data can provide a high level of discriminant accuracy and able to reflect the true financial position and operating results.

It is desirable to further apply logistic regression model to solving those more complex real-world optimization problems and it will be our further work.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflicts of interest.

ACKNOWLEDGEMENTS

This work is supported by the Humanities and Social Sciences Research Projects of Ministry of Education of the P.R.C. (13YJCZH060) and the Education Research Foundation of Hubei Province (2012111).

REFERENCES

- [1] E.I. Altman, "Financial ratios, discriminant analysis and prediction of corporate bankruptcy", *Journal of Finance*, vol. 9, pp. 589-609, 1968.
- [2] A. Albert and J.A. Anderson, "On the existence of maximum likelihood estimates in logistic regression models", *Biometrika*, vol. 71, no. 1, pp. 1-10, 1984.
- [3] M.Y. Chen, "Predicting corporate financial distress based on integration of decision tree classification and logistic regression", *Expert Systems with Applications*, vol. 38, no. 9, pp. 11261-11272, 2011.
- [4] J.D. Groarke, J. Gallagher, J. Stack, A. Aftab, C. Dwyer, R. McGovern and G. Courtney, "Use of an admission early warning score to predict patient morbidity and mortality and treatment success", *Emergency Medicine Journal*, vol. 25, no. 12, pp. 803-806, 2008.
- [5] Z. Hua, Y. Wang, X. Xu, B. Zhang and L. Liang "Predicting corporate financial distress based on integration of support vector machine and logistic regression", *Expert Systems with Applications*, vol. 33, no. 2, pp. 434-440, 2007.
- [6] H. Li and J. Sun, "Empirical research of hybridizing principal component analysis with multivariate discriminant analysis and logistic regression for business failure prediction", *Expert Systems with Applications*, vol. 38, no. 5, pp. 6244-6253, 2011.
- [7] D. Martin, "Early warning of bank failure: A logit regression approach", *Journal Of Banking & Finance*, vol. 1, no. 3, pp. 249-276, 1977.
- [8] J.S. Ohlson, "Financial Ratios and the Probabilistic Prediction of Bankruptcy", *Journal of Accounting Research*, vol. 18, no. 1, pp. 109-131, 1980.
- [9] B. Silke, J. Kellett, T. Rooney, K. Bennett and D. O'Riordan, "An improved medical admissions risk system using multivariable fractional polynomial logistic regression modelling", *QJM: An International Journal of Medicine*, vol. 103, no. 1, pp. 23-32, 2010.
- [10] A.B. Wilson, B.J. Kerr, N.D. Bastian and L.V. Fulton, "Financial performance monitoring of the technical efficiency of critical access hospitals: a data envelopment analysis and logistic regression modeling approach", *Journal of Healthcare Management*, vol. 57, no. 3, pp. 200-212, 2012.

Received: June 10, 2015

Revised: July 29, 2015

Accepted: August 15, 2015

© Lixin and Long; Licensee Bentham Open.

This is an open access article licensed under the terms of the (<https://creativecommons.org/licenses/by/4.0/legalcode>), which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.