Comparison of Early-warning Methods on Both Logistic Regression Analysis Model and Factor Analysis Model

Yu-mei HE¹,a, Xuan CHENG²,b,* , Xin-yi XU³,c and Xiao-wei ZHU⁴,d
¹,²,⁴College of Business, Chengdu University of Technology, Chengdu, Sichuan, China
³Southwestern University of Finance and Economics, Chengdu, Sichuan, China
¹hym_1@sina.com, b1318623426@qq.com, c1318623426@qq.com, d792347230@qq.com
*Corresponding author

Keywords: Factor Analysis Method, Logistic Regression Analysis Method, Financial Index, Comparison of Early-warning Methods

Abstract. Risk prevention is vital for scientific decision-making and sustainable development of enterprises. It has been proved practically to be effective that the operating situation of enterprises is diagnosed based on the financial indexes. This paper is aimed at mineral resources listed companies in China. Indexes which reflecting the profitability, growth ability, solvency, cash flow capacity and operating capacity of the enterprise operating situation are determined according to the classification settings of CSMAR “Financial Index Analysis Database of Listed Companies in China”; and then the early-warning models of factor analysis method and Logistic regression method are constructed by screening of 19 ST listed companies and 57 non ST companies from 2010 to 2014 as the sample data, and the early-warning effects of the these two models are compared. The outcome of the empirical research shows that Logistic regression analysis method is more effective than factor analysis method in the forecasting business crisis or financial difficulties, and then this paper put forward five policy suggestions for preventing business risks.

Introduction

Sustainable development of enterprises not only to consider environmental issues but also should pay more attention to risk monitoring. How to implement the scientific and effective risk monitoring? As far as the basic data in China’s theory and practice, the indexes which can reflecting the enterprise operating situation are mainly financial indexes. So it is an good ideal choice to use the financial index data of enterprises as the basic data to analyze the enterprise operating situation. The financial data of enterprises are basically available, which provides the conditions for quantitative analysis. Using the financial data of enterprises, with the help of analytical tools or mathematical modeling innovation analysis methods, become the mainstream of risk prediction research. In order to facilitate the research, this paper choose mineral resources listed companies in China as the research object. Based on the financial data of selected sample companies, using the factor analysis method and Logistic regression method to construct the early-warning model of mineral resources enterprises. And the early-warning effects of the two models are compared and analyzed, so as to determine a relatively ideal prediction method, so that enterprises can adjust the business strategies to avoid risks.

Literature Review

Domestic and foreign scholars have done a lot of research on the early-warning of enterprise risk. Cuiping Li et al.¹ (2016) determine the index weights by subjective and objective combination weighting method, and then establish the GM(1,1) early-warning model for iron and steel enterprises. The feasibility and validity of the proposed early-warning model was tested and verified through its application in a functioning industrial plant. Bao et al.⁵ (2016) use balanced panel data of manufacturing industry from 2001 to 2010 to build the panel data logit model for financial distress early-warning after choosing the financial indexes and non-financial indexes. Zhang et al.⁸
(2015) use information fusion technique to build a finance early-warning model based on data mining methods such as SVM and Logistic model, which can integrate the respective strengths from different data mining methods to improve the prediction accuracy rate, it fuses the different data mining results to gain the prediction results for reliable decision.

From the above literature, it can be found that different methods are chosen according to the different distribution of basic data. The effectiveness of early warning model is limited by the distribution of financial data and the range of sample selection. In addition, the sample selection of different industries will also affect the warning effect of early-warning model. Taking these factors into account, this paper takes the mineral resources listed companies in China as samples, according to the classification settings of CSMAR “Financial Index Analysis Database of Listed Companies in China”, indexes which reflecting the profitability, growth ability, solvency and cash flow capacity and operating capacity of the enterprise operating situation are determined. The early-warning models of factor analysis method and Logistic regression method are constructed by screening of 19 ST listed companies and 57 non ST companies from 2010 to 2014 as the sample data, and the early-warning effects of the these two models are compared.

Empirical Analysis

The selection of sample

In view of the representative and convenience of the research, we take the mineral resources listed companies in China as the research objects.

According to the Chinese Securities Regulatory Commission issued the “listing corporation industry classification guidelines”, select the ferrous metal mining industry (B08), nonferrous metal mining industry (B09), non metal mining industry (B10), nonmetal mineral products industry (C30), ferrous metal smelting and rolling processing industry (C31), non-ferrous metal smelting and rolling processing industry (C32) and the metal products industry (C33) as the selection range of minerals enterprise, from where pick out 19 ST listed companies in 2010-2014(excluding the companies whose financial data are incomplete in these four years) as the research samples, and in accordance with the ratio of 1:3 select 57 non ST companies who have the same accounting year, the similar size of assets and the same industry segmentation as the paired samples for analysis. Select the financial indexes one year, two years and three years before the ST year of sample companies from the China listed corporation financial index analysis database of CSMAR solution.

The selection of index

In order to fully and accurately reveal the status of enterprise management, this paper draws on the setting and classification of the relevant indexes of “Financial Index Analysis Database of Listed Companies in China”, CSMAR solution. Choose 29 financial indexes as alternative variables including profitability, growth ability, solvency, cash flow capacity and operating ability. According to the data of three years before ST year of 19 ST companies, analyze 29 financial indexes with T test of paired samples, a total of 10 indexes are authentic under the significance level of 5%, which are net profit rate of total assets ($X_1$), assets return ratio ($X_2$), liqidity ratio ($X_3$), equity to debt ratio ($X_4$), cash ratio ($X_5$), inventory turnover rate ($X_6$), turnover of current assets ($X_7$), asset turnover ($X_8$), turnover of fixed assets ($X_9$) and equity multiplier ($X_{10}$). Test results show that there are significant differences between ST companies and non ST companies in terms of debt paying ability, profit ability, cash flow and operating capacity. This paper uses the above 10 financial indexes to build the early-warning model of enterprise management.

Analysis of empirical results of factor analysis method

KMO and spherical Bartlett test

The output shows that KMO>0.5, the significant probability of Bartlett test statistic is 0.000, which indicates that the sample is suitable for factor analysis.
The construction of model

This paper uses SPSS software to carry on a principal component analysis based on the standardized data of estimate sample group in the one year before the ST year. The characteristic value and contribution rate of each principal component are calculated. This paper selects the first four principal components which containing the information content (cumulative contribution rate) of the original index 86.728%. Four principal component contribution rate are 30.541%, 22.673%, 0.289% and 13.224% respectively after the maximum variance rotation.

According to the factor score matrix table, the linear expression of principal components (F) on financial indexes can be got as follows:

\[
F_1 = -0.002X_1 - 0.023X_2 + 0.329X_3 + 0.333X_4 + 0.333X_5 + 0.124X_6 - 0.019X_7 - 0.020X_8 + 0.030X_9 - 0.009X_{10} \\
F_2 = 0.016X_1 + 0.021X_2 + 0.018X_3 + 0.004X_4 - 0.030X_5 - 0.030X_6 + 0.294X_7 + 0.425X_8 + 0.435X_9 - 0.143X_{10} \\
F_3 = 0.490X_1 + 0.493X_2 - 0.007X_3 - 0.015X_4 - 0.016X_5 - 0.055X_6 - 0.051X_7 - 0.004X_8 + 0.086X_9 + 0.049X_{10} \\
F_4 = 0X_1 - 0.004X_2 + 0.30X_3 + 0.061X_4 + 0.081X_5 + 0.653X_6 + 0.210X_7 - 0.043X_8 - 0.273X_9 + 0.553X_{10}
\]  

(1)  

(2)  

(3)  

(4)  

According to the contribution rate of each principal component factor, the management early-warning model of listed enterprises of mineral resources can be got as follow:

\[
Z_1 = 0.30541F_1 + 0.22673F_2 + 0.019F_3 + 0.020F_4 \\
Z_2 = 0.30952F_1 + 0.22299F_2 + 0.21787F_3 + 0.13224F_4
\]  

(5)  

Among them, F_1 represents the solvency; F_2 represents the operating capacity; F_3 represents the profitability; F_4 represents the short-term solvency level.

Construct the early-warning models of companies in the two years and three years before the ST year according to the above method:

\[
Z_1 = 0.30541F_1 + 0.22673F_2 + 0.019F_3 + 0.020F_4 \\
Z_2 = 0.30952F_1 + 0.22299F_2 + 0.21787F_3
\]  

(6)  

(7)  

Discriminant analysis of the model

According to the early-warning model equations, SPSS software automatically calculate the Z value of the listed companies in the different accounting year, this paper makes a discriminant classification of Z value of the listed company in each year. The figure of “0” means the ST listed company, the figure of “1” means the non ST listed company. The discriminant result are shown in table 1.

Table 1. The discriminant results of factor analysis.

<table>
<thead>
<tr>
<th>Year of sample</th>
<th>The year before ST</th>
<th>Two years before ST</th>
<th>Three years before ST</th>
<th>Actual results</th>
<th>Predicted results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>count</td>
<td>%</td>
<td>count</td>
<td>%</td>
<td></td>
</tr>
<tr>
<td>Initial</td>
<td>0</td>
<td>11</td>
<td>19</td>
<td>15</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>16</td>
<td>41</td>
<td>57</td>
<td>23</td>
</tr>
<tr>
<td>Cross</td>
<td>0</td>
<td>11</td>
<td>19</td>
<td>15</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>16</td>
<td>41</td>
<td>57</td>
<td>23</td>
</tr>
</tbody>
</table>

Table 1 says that in the validation of the one year before the ST year, the discrimination correct rate of ST companies is 42.1% and that of the non ST companies is 71.9%. In the validation of the two years before the ST year, the discrimination correct rate of ST companies is 78.9% and that of
non ST companies is 59.6%. In the validation of the three years before the ST year, the
discrimination correct rate of ST companies is 63.2% and that of the non ST companies is 45.6%.
As can be seen from the above discriminant results, although the correct rates of discrimination
are mostly less than 70%, the correct rate of discrimination on the financial early-warning of two
years before the ST year is highest in the correct rates of discrimination on the financial
early-warning of one year, two years and three years before the ST year. The financial
early-warning result of two years before the ST year can be served as a reference of business
early-warning for the enterprises.

Analysis of empirical results of Logistic regression method

Assuming that the variable Y indicates whether the listed company appeared ST, Y=0 means the
ST listed company, Y=1 means the non ST listed company.
The m independent variables that affect the dependent variable Y are \(X_1, X_2, \ldots, X_m\). The
conditional probability of occurrence of an event is \(P(Y=0|X_i) = p_i\), Logistic model discriminant
can be written as equation (1):
\[
P_i = \frac{1}{1 + e^{-(\alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_m X_m)}}
\]

(8)

Construction of early-warning model by Logistic regression analysis method based on
variables determined by principal component method

Due to the Logistic model is more sensitive to multicollinearity of independent variables,
therefore on the basis of the above model, considering the effect of multicollinearity between
variables, the Logistic regression models are constructed by respectively using the factor analysis
results of financial indexes of the one year, two years and three years before the ST year. The
Logistic regression results of the one year before ST year are as shown in table 2:

<table>
<thead>
<tr>
<th>Variables in the equation</th>
<th>B</th>
<th>S.E.</th>
<th>Wals</th>
<th>df</th>
<th>Sig.</th>
<th>Exp (B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>step 1*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FAC1_1</td>
<td>.212</td>
<td>.370</td>
<td>.327</td>
<td>1</td>
<td>.567</td>
<td>1.236</td>
</tr>
<tr>
<td>FAC2_1</td>
<td>-.235</td>
<td>.264</td>
<td>.793</td>
<td>1</td>
<td>.373</td>
<td>.790</td>
</tr>
<tr>
<td>FAC3_1</td>
<td>.131</td>
<td>.318</td>
<td>.169</td>
<td>1</td>
<td>.681</td>
<td>1.140</td>
</tr>
<tr>
<td>FAC4_1</td>
<td>-.598</td>
<td>.267</td>
<td>5.007</td>
<td>1</td>
<td>.025</td>
<td>.550</td>
</tr>
<tr>
<td>constant</td>
<td>1.190</td>
<td>.287</td>
<td>17.194</td>
<td>1</td>
<td>.000</td>
<td>3.286</td>
</tr>
</tbody>
</table>

a. Variables entered in step 1: FAC1_1, FAC2_1, FAC3_1, FAC4_1.

In the four factor scores, only the significance level of FAC4_1 is less than 0.05, indicating that
FAC4_1 has a significant effect on the discriminant equation, while the effects of other three factors
on the discriminant equation are not significant. Therefore, the discriminant equation for this model
is:
\[
P_i = \frac{e^{0.212 \times FAC1_1 - 0.235 \times FAC2_1 + 0.131 \times FAC3_1 - 0.598 \times FAC4_1 + 1.190}}{1 + e^{0.212 \times FAC1_1 - 0.235 \times FAC2_1 + 0.131 \times FAC3_1 - 0.598 \times FAC4_1 + 1.190}}
\]

(9)

Same way to get the discriminant models in the two years and three years before the ST year, \(P_2\)
and \(P_3\), respectively are:
\[
P_2 = \frac{e^{2.039 \times FAC1_1 - 0.183 \times FAC2_1 + 3.306 \times FAC3_1 + 1.975}}{1 + e^{2.039 \times FAC1_1 - 0.183 \times FAC2_1 + 3.306 \times FAC3_1 + 1.975}}
\]

(10)
\[ P_3^* = \frac{e^{1.038 \text{FAC}_1 + 0.877 \text{FAC}_2 + 0.387 \text{FAC}_3 + 1.384}}{1 + e^{1.038 \text{FAC}_1 + 0.877 \text{FAC}_2 + 0.387 \text{FAC}_3 + 1.384}} \] (11)

**Analysis of discriminant results**

The figure of “0” means the ST listed company, the figure of “1” means the non ST listed company. The predicted results, actual results and the number of their discrimination results are summarized in table 3.

Table 3. Logistic regression discriminant result (Based on the principal component factor variables).

<table>
<thead>
<tr>
<th>Year of sample</th>
<th>The year before ST</th>
<th>Two years before ST</th>
<th>Three years before ST</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct number</td>
<td>Actual results</td>
<td>Correct rate (%)</td>
<td>Correct rate (%)</td>
</tr>
<tr>
<td>0</td>
<td>4</td>
<td>15</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>57</td>
<td>100.0</td>
</tr>
<tr>
<td>Total percentage</td>
<td>79.4</td>
<td>89.5</td>
<td>76.3</td>
</tr>
</tbody>
</table>

It can be seen from the above table that in the one year before the ST year, the discrimination correct rate of ST companies is 21.1% and that of non ST companies is 100%. The total discrimination correct rate is 80.3%. In the two years before the ST year, the discrimination correct rate of ST companies reached 68.4% and that of non ST companies is 96.5%. The total correct rate is 89.5%. In the three years before the ST year, the discrimination correct rate of ST companies is 5.3% and that of non ST companies is 100%. The total discrimination correct rate is 76.3%.

According to the above discriminant results, it can be seen that the discrimination correct rate of the financial early-warning model of two years before the ST year is highest in the discrimination correct rates of the financial early-warning of one year, two years and three years before the ST year. The early-warning result of two years before the ST year can be served as a reference of business early-warning for the enterprises.

**Comparative analysis of the application of early-warning model of factor analysis method and Logistic regression method**

By analyzing the discriminant results of the above models, we can see:
1. The discrimination correct rate of Logistic regression model is higher than that of factor analysis model. So the model constructed by Logistic regression method can fit the characteristics of the whole sample better.
2. The discrimination correct rate of two years before ST year is the highest for every method in two models of both factor analysis method and Logistic regression method. This result suggests financial indexes of two years before ST year have the best early-warning effect of the company. The financial indexes of two years before ST year are the best measurement to determine if the company is in business crisis or financial difficulties.

**Conclusions and Suggestions**

**Conclusions**

The conclusions on this empirical study are that,
1. The financial early-warning models of two years before ST year on both factor analysis method and Logistic regression analysis method all can play a role in forecasting business crisis or financial difficulties.
2. The early-warning effect of financial early-warning model of Logistic regression method is better than that of factor analysis method.
3. Compared with the financial indexes of the other year, the early-warning model constructed by the financial indexes of two years before the ST year has the best early-warning effect.
4. It is a gradual process that a enterprise becomes a ST company, so inevitably reflected by the financial indexes, therefore, the early-warning model by the financial indexes is more reliable than other data.
5. The use of early-warning model based on Logistic regression analysis method is more simple, more flexible and much easier to apply.

**Policy suggestions**

1. The scientific design of financial early-warning model should be brought into the strategic management of enterprises.
   In order to achieve sustainable development, enterprises must pay attention to train the risk consciousness of the whole staff and establish a sound scientific risk prevention mechanism, and the design and application of early-warning model is an indispensable part. Effective prediction of operating risk is a reliable guarantee of enterprises to improve the ability to resist risks.
2. The financial indexes database of enterprises over the years is established and renewed in order to ensure the authenticity, accuracy and comprehensiveness of the basic data.
   The design of early-warning model by using big data requires a lot of data information of financial indexes which reflecting the operating status of enterprises. These basic data lays the foundation for the effective design and application of early-warning model.
3. Establish operational risk monitoring system, improve the internal control system and carry out daily management in company's rules and regulations, organizational structure, staff and workers and business and so on.
   There are inevitable risks in the exchange and operation of the enterprises in the environment of market economy. Managers should strengthen the sense of responsibility, establish and improve the risk monitoring system, meanwhile establish targeted and quantifiable risk prediction and monitoring system and carry out all-round supervision at any time to ensure that the risk is manageable.
4. Establish a complete comprehensive risk management system with reference to international financial early-warning models.
   There is no fixed method of risk management, the enterprise must establish a flexible response mechanism for active control and put the risk management system completely within the internal control system.
5. Improve domestic risk control laws and regulations, improve enterprise risk management and control standards which are in line with international risk management methods.

**Contributions of this Paper**

Research findings of the early-warning models based on financial indexes have been emerging constantly. These methods are on different research perspectives, different index options, different basic models and the comparison of early-warning effects of different models. Such as Z-Score model (Zhao et al.\[11\], 2015), method of variable precision weighted average roughness decision tree (Bao et al.\[12\], 2015), credit risk model and financing constraint model (Zou et al.\[14\], 2016). And this paper chooses mineral resources listed companies in China as research sample, and uses the factor analysis method and the Logistic regression analysis method to construct the enterprise's early-warning models. The empirical research conclusions can be directly applied to the early-warning of the mineral resources enterprises in China. The use of early-warning model of Logistic regression analysis method can help enterprises to prevent operating risks, while providing a warning to investors to avoid investment risks. The suggestions of this paper provide a good policy reference and method guidance for enterprises to carry out scientific early-warning. The research results provide theoretical foundation and empirical case for enterprises to carry out scientific early-warning management.
Acknowledgement

This research was supported by the foundation of System Science and Enterprise Development Research Center of Sichuan Province (project name: The impact of environmental regulation and ecological innovation on enterprise competitiveness; grant No. Xq16C11), Science & Technology Department of Sichuan Province (Soft science project of Sichuan science and technology: Mechanism research on the deeply integrated of military and civilian promote the development of strategic emerging industries in Sichuan province; grant No. 2017ZR0121), Research Office of Sichuan Provincial People’s Government (project name: A investigation report of the study on the path of military and civilian integration based on strategic emerging industries in Sichuan province; grant No. Cdut2016-218) and Education Department of Sichuan (The key project of humanities and social sciences: The exploration of the innovation ability construction of Sichuan titanium industry; grant No. 14SA0035).

References


