Research on Enterprise Financial risk Prediction Method Based on Regression Analysis

Shirui Zhang*

Department of accounting, Tianjin University of Finance and Economics, Tianjin, 300202, China

Abstract: The prediction and control of enterprise financial risk is an important research topic in the financial field. Over the years, the financial risks of enterprises have seriously affected the healthy development of enterprises, credit institutions, securities investors and even the whole country. However, from the current practice and theoretical research of enterprises, how to effectively evaluate and control the financial risk of Chinese enterprises is still very scarce. Therefore, this paper conducts detailed analysis on the prediction and control of financial risk based on regression analysis to better identify the financial risks of enterprises.

1 Introduction

At present, the operation and management mode of many large group companies is changing. The previous asset management are gradually transformed into the current capital management, and has entered the enterprise management stage which focus on finance [1]. Finance plays a vital role in the development of enterprises. The life and death of an enterprise today and its subsequent growth are closely related to its financial status. With the further deepening of the theoretical system of socialist market economy in practice, Chinese enterprises are also facing more and more complex financial crises in the process of continuous innovation and development. The existence of these crises has greatly impacted the operation of enterprises, and has had a considerable impact on the social stability and economic development of the whole country [2]. The high risks associated with economic activities urge enterprises to identify before the financial crisis they face, and to understand the nature of the crisis and predict the possible losses caused by risk accidents. On the basis of this identification and understanding, enterprises need to implement the most effective crisis prevention measures to minimize losses, avoid the deterioration of their financial situation and maintain their normal business activities.

Therefore, only by deeply analyzing and studying the financial crisis can the enterprise maintain a sound, stable and rapid growth. Foreign experts' research on financial risk is mainly based on Fitzpatrick's single variable early warning model and Altman's multivariable warning model. The accuracy of multivariate warning model is very high. In recent years, scholars in several countries have also carried out similar research on the basis of this model. However, the research on financial risk in China is still in its infancy, and most of the research on enterprise financial risk is still on financial indicators and managers' performance [3]. On the basis of being familiar with the causes of financial risks, it is of great practical significance to create a relatively perfect risk control system [4].

2 METHOD and result analysis.

2.1 Significance test of financial indicators

Non-parametric tests are used to determine whether data come from the same population when the population does not know whether it obeys the normal distribution and the distribution is unknown. There are numerous nonparametric testing methods. The Mann-Whitney U test with two independent samples is used in this paper. Mann-Whitney U test is made on the financial indicators of 32 ST companies and 32 normal companies in the t-3 years. Finally, the test results are as follows:

*Corresponding author. Email: 2628910182@qq.com

Table 1. Mann-Whitney U test of 27 financial indicators

	Mann-Whitney U	Wilcoxon W	Z	Asymp. Sig.
Earnings per share x1	51	579	-6.1899888	0.00000000
Earnings per share growth rate x2	263	791	-3.3433996	0.00082759
Debt ratio per share x3	446	974	-0.8861922	0.37551399
Current ratio x4	253	781	-3.4776727	0.00050579
Quick ratio x5	247	775	-3.5581958	0.00037341
Cash-to-debt ratio x6	389	917	-1.6515399	0.09862837
Shareholder equity ratio x7	300	828	-2.8465566	0.00441949
Current debt ratio x8	380	908	-1.7748239	0.07592696
Debt-to-equity ratio x9	427	955	-1.1413081	0.25374174
Inventory turnover rate x10	253	781	-3.4776329	0.00050586
Accounts receivable turnover x11	345	873	-2.2423347	0.02493975
Main cost ratio x12	358	886	-2.0677817	0.03866056
Total assets turnover x13	150	678	-4.8606297	0.00000117
Operating net interest rate x14	126	654	-5.1828814	0.00000022
Operating gross profit margin x15	352	880	-2.1483446	0.03168639
Net asset interest rate x16	75	603	-5.8676662	0.00000000
Return on net assets x17	108	636	-5.4251912	0.00000006
Growth rate of net assets x18	194	722	-4.2698838	0.00001956
Pre-tax profit growth rate x19	284	812	-3.0614261	0.00220285
Main revenue growth rate x20	259	787	-3.3971088	0.00068102
Net profit growth rate x21	248	776	-3.5448092	0.00039290
Total assets growth rate x22	251	779	-3.5045273	0.00045742
Shareholder's equity growth rate x23	197	725	-4.2296019	0.00002341
Main profit growth rate x24	289	817	-2.9942896	0.00275085
Asset-liability ratio x25	308	836	-2.5846554	0.00974764
Long-term debt-asset ratio x26	459	894	-0.0722624	0.94239306
Main income cash content x27	332	860	-2.2546994	0.02415220

As can be seen from Table 1, if α =0.01 is considered as the significance level, the probability of companionship Asymp. Sig. of variables x3, x6, x8, x9, x11, x12, x15, x26, x27 are greater than α , which means that their data are not significantly different between ST and non-ST companies and cannot be used to determine their difference, which means that they are to be excluded from the model is to be excluded from the model [5]. In this way, there are 18 variables in the model.

2.2 Univariate Logistic Regression Analysis

To improve the goodness of fit of the model, we use Logistic regression analysis to check the binary relationship between each independent variable (the financial indicators selected for the first time above) and the dependent variable. We test the goodness of fit of the model for the significant financial indicators selected for the first time in t-3. The final test results are as follows:

Table 2. Wald test of goodness of fit of financial indicators

financial target	Wald	freedom	P value
Earnings per share x1	23.41276	1	0.000001
Earnings per share growth rate x2	8.524978	1	0.003503
Current ratio x4	3.3371	1	0.067734
Quick ratio x5	2.509493	1	0.113162
Shareholder equity ratio x7	6.595435	1	0.010224
Inventory turnover rate x10	0.436597	1	0.508770
Total assets turnover x13	24.15945	1	0.000001
Operating net interest rate x14	11.16412	1	0.000834
Net asset interest rate x16	20.33995	1	0.000006
Return on net assets x17	4.675405	1	0.030597
Growth rate of net assets x18	7.60093	1	0.005834
Pre-tax profit growth rate x19	8.27266	1	0.004025
Main revenue growth rate x20	1.967001	1	0.160767
Net profit growth rate x21	8.76832	1	0.003065
Total assets growth rate x22	2.57025	1	0.108890
Shareholder's equity growth rate x23	5.474756	1	0.019293
Main profit growth rate x24	2.202166	1	0.137817
Asset-liability ratio x25	6.283355	1	0.012188

According to the discriminant criteria in Table 2, we choose the following indicators: x1, x2, x7, x13, x14, x16, x18, x19, x21. Further factor analysis will be conducted based on these variables.

2.3 Factor Analysis

The above nonparametric test and univariate regression analysis resulted in the selection of nine financial indicators with significant differences that are conducive to improving the model's prediction accuracy. Although it has been simplified, using these financial indicators to build models is still complicated and time-consuming, and there may be a strong correlation between them. [6] As a result, the factor analysis method is required to further simplify these financial indicators and use a few factors to describe the relationship between many indicators or factors. Based on data from the previous three years, this paper employs the factor analysis method to re-streamline the financial indicators chosen the second time. The following steps are roughly included in factor analysis: (1) Determine whether some of the original variables to be analyzed are suitable for factor analysis; (2) Rotate the constructed factor variables to make them more interpretable; and (3) Calculate the factor variable scores. Following confirmation, the following outcomes are obtained:

Table 3. KMO and Bartlett's inspection

Kaiser-Meyer-Olkin Measure of Sampling Adequacy. (KMO)		0.781
Bartlett's Test of Sphericity	Approx. Chi-Square	892.507
	Sig.	0.000

The Table 3 shows the results of KMO test and Bartlett sphericity test. The value of KMO is 0.781. According to the standard given by statistician Kaiser, the value of KMO is greater than 0.6, which is suitable for factor analysis. In addition, the associated probability given by Bartlett sphericity test is 0.000, which is less than the significance level of 0.05. Therefore, the zero hypothesis of Bartlett sphericity test is rejected and considered suitable for factor analysis.

Table 4. Contribution rate of factor variance

	Initial Eigenvalues		
Componen	Total (eigenvalue)	% of Variance	Cumulative% (cumulative)
1	5.016032546	55.73369496	55.73369496
2	1.914052942	21.26725491	77.00094987
3	0.851372222	9.459691361	86.46064123
4	0.644178308	7.157536757	93.61817799
5	0.361186316	4.013181292	97.63135928
6	0.127779235	1.419769282	99.05112856
7	0.049939303	0.554881143	99.60600971
8	0.034793452	0.386593906	99.99260361
9	0.000665675	0.007396389	100
Extraction Method: Principal Component Analysis. (principal component analysis)			

The overall description of the original variables by the initial solution of factor analysis is shown in Table 4. The second column contains the variance contribution (eigenvalue) of factor variables, which serves as an index for determining the importance of factors. The first factor's eigenvalue is 5.02, indicating that it describes 5.02 of the total variance of the original variables 9, while the variance described by the later factors gradually decreases. The third column shows the variance contribution rate of each factor, and the fourth shows the cumulative variance contribution rate of factor variables, which shows the proportion of total variance described by the first M factors to total variance of the original variables [7]. The cumulative contribution rate of the first three factors can reach 86.46 percent, indicating that the first three factors can be used for factor analysis of principal components. According to the score factor regression matrix calculated by the regression algorithm output by SPSS, the following factor score function can be obtained:

$$\begin{cases} F1 = -0.047x_1 + 0.27x_2 + 0.011x_7 - 0.153x_{13} - 0.052x_{14} - 0.028x_{16} + 0.268x_{18} + 0.274x_{19} + 0.268x_{21} \\ F2 = 0.391x_1 - 0.078x_2 - 0.242x_7 + 0.498x_{13} + 0.168x_{14} + 0.35x_{16} - 0.084x_{18} - 0.101x_{19} - 0.072x_{21} \\ F3 = -0.09x_1 + 0.004x_2 + 0.852x_7 - 0.272x_{13} + 0.358x_{14} - 0.027x_{16} - 0.002x_{18} + 0.021x_{19} - 0.001x_{21} \end{cases} \quad (1)$$

According to this score function, the information of nine variables can be converted into the information of three factors, that is, the variables simplified into three principal components. The following three factors are used for Logistic regression analysis.

2.4 Logistic Regression Analysis

Using the scoring functions of the above three factors, the new factor variables F1, F2 and F3 can be obtained by inputting the values of the original nine variables into the scoring function. The following results are obtained:

Table 5. Parameter Table of Regression Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1(a)	F1	-5.54703	5.18409	1.144921	1	0.004615	0.003899
	F2	-8.83341	2.921117	9.144474	1	0.002495	0.000146
	F3	-6.37548	2.70134	5.57016	1	0.010269	0.001703
	Constant	1.546161	0.952518	2.634895	1	0.010539	4.69342

a. Variable(s) entered on step 1: F1, F2, F3.

In Table 5,B, S.E., Wald, df, Sig. ,Exp(B) respectively represent variable coefficient, standard deviation, Wald score, degree of freedom, adjoint probability and coefficient logarithm. From Table 5, it can be seen that the probability of companionship of all three variables and constant quantities is very small, and if judged by the default significance level $\alpha = 0.05$, all will be less than the significance level, i.e., they can meet the requirements. Therefore, it can be seen that the regression model obtained from the three factorial variables fits well. In

addition, the logistic prediction probability model can be obtained from Table 5, namely

$$\hat{p} = \frac{exp(1.5462-5.5470F1-8.8334F2-6.3755F3)}{1+exp(1.5462-5.5470F1-8.8334F2-6.3755F3)} \quad (2)$$

According to the probability model, if the calculated probability p<0.5, it can be predicted that the enterprise is risk-free; if the probability p>0.5, it can be predicted that the enterprise is risk-free, or in other words, it will be in danger of special treatment. In this way, this probability model can be used for prediction and inspection.

2.5 Testing and Prediction of Models

Using the above model formula, 64 modeling samples are brought into the calculation, and after statistics, the following results are obtained:

Table 6. Comparison between predicted value based on our method and actual value of model

reality / predict	0	1	percentage
0	31	1	96.9%
1	4	28	87.5%
add up to			92.2%

Among them, sensitivity, also known as true positive rate, is the probability that the predicted result is also 1 among individuals actually classified as 1; Specificity, also known as true negative rate, is the probability that the predicted result is also 0 among individuals actually classified as 0; Missed diagnosis rate, also known as false negative rate, is the probability that the predicted result is zero among individuals actually classified as 1; Misdiagnosis rate, also known as false positive rate, is the probability that the predicted result is 1 among individuals actually classified as 0 [8]; The overall prediction accuracy rate is the rate of correct prediction results. Table 6 shows that the sensitivity =87.5%, specificity =96.9%, missed diagnosis rate =12.5%, misdiagnosis rate =3.1%, and the overall prediction accuracy rate is 92.2%, indicating that the overall prediction effect is very good.

In order to further prove the effectiveness of the model, we use the Z-Core model to re analyze the financial risks of 64 enterprises, and compare the results and the calculation, and after statistics, the following results are obtained:

Table 7. Comparison between predicted value based on Z-core model and actual value

reality / predict	0	1	percentage
0	30	2	93.8%
1	2	30	93.8%
add up to			93.8%

Through comparison in Table 7, the overall prediction effect of the two groups of models is similar. However, Z-core model is effective for financial early warning of Chinese enterprises, but its standard value is not suitable for China. We need to make some adjustments when applying this model. In contrast, our model is based on the Chinese market and is more suitable for the situation in China. In addition, the model we built is more brief and convenient to use. In conclusion, our model is a competitive method.

In order to better illustrate the representativeness and accuracy of the model, we selected another 48 enterprises for testing, and the results are as follows: Sensitivity =87.5%, specificity =91.67%, missed diagnosis rate =12.5%, misdiagnosis rate =8.33%, and overall prediction accuracy rate is 89.58%, which shows that this method has certain applicability and popularization.

Multiple regression analysis takes six selected indexes as independent variables and financial coefficient as observed variables, and uses stepwise regression method to study the factors that affect the financial risk of the company. The regression equation we attain is: Y=0.997x-1.

3 CONCLUSION

Nowadays, the financial risk of enterprises has caused many enterprises to go bankrupt or close to bankruptcy. Facing the problems brought by financial risks, we are required to avoid financial risks as much as possible with a positive attitude [9]. In this paper, we establish a competitive financial risk prediction model based on regression analysis and verify its reliability and generalization. The management of financial crisis and establishment of financial early warning mechanism are essential for the sustainable and healthy development of enterprises. This requires enterprises to have a clear understanding of financial risks in the process of operation. Financial risks must be fully considered when making relevant decisions [10]. Various financial information of the company should be frequently compared and analyzed. Finally, potential risk factors in the actual operation of the company should be clarified.

REFERENCES

1. Guoyinan Enterprise financial risk prevention and management optimization Countermeasures [j] Modernization of shopping malls, 2019,05:105-106
2. Li min Research on financial risk evaluation of GEM listed companies [d] Shanxi: School of economics, Taiyuan University of technology, 2018
3. Wu Yubei Research on financial risk control [d] Shandong: School of economics and management, Shandong University, 2016
4. Caoweizhong Analysis and prevention of financial risks of small and medium-sized enterprises [j] Modern economy (second half of modern property), 2019, (02): 91-103
5. SUN Hao Research on financial early warning based on logistic model [d] Hunan: School of economics and management, Nanhua University, 2017
6. Niejianping On financial risk prevention of Listed Companies in China [j] Commodity and quality, 2018, (S6): 88
7. Xuxinqiang Enterprise financial risk prevention and control from the perspective of cash flow [m] Accounting communication, 2019, (23): 143-144
8. Fanguangjuan Research on financial risk control of listed companies [d] Shanxi: School of economics, Shanxi University of Finance and economics, 2017
9. Zhubeiyun Analysis on financial risk management of Listed Companies in China [j] Financial circles (Academic Edition), 2019,09:191-192
10. Zhangqizhong Research on financial risk early warning management of small and medium-sized enterprises [d] Wuhan: School of economics, Wuhan University of technology, 2020