ISSN:2790-1661 DOI: 10.56028/aemr.4.1.52.2023

Research on financial early warning of listed companies based on factor analysis and logistic regression

Xianfu Zeng [1, a], Jayson Xu [2, b]

[1] College of International Economics and Trade, Ningbo University of Finance and Economics, Ningbo 315175, Zhejiang, China;

[2] Ningbo Foreign Language School, Ningbo 315121, Zhejiang, China.

[a] zengxianfu@nbufe.edu.cn, [b] jaysonxuruichen2024@outlook.com

Abstract. The establishment of an accurate, effective, sensitive and stable financial crisis early warning model for listed companies has a role that cannot be underestimated for the stability and healthy development of China's financial order. In this paper, financial statements of listed companies in Shanghai and Shenzhen Stock markets from 2018 to 2021 were selected as sample data. On the basis of the four traditional financial indicators, namely profitability, solvency, operating ability and development ability, two new financial indicators, namely profit quality and market valuation, were added. The factor analysis method and binary Logistic regression method are combined to construct the financial early warning model of listed companies. The results show that the model has high fitting accuracy and discriminant ability, which proves that the financial indicators disclosed by listed companies can provide empirical basis for various stakeholders to provide an empirical basis for early warning of corporate financial risks.

Keywords: financial crisis; factor analysis; logistic regression model; financial warning.

1. Introduction

Facing the impact of economic globalization and increasing market competition pressure, how to remain invincible is the premise of the existence and development of each enterprise. Enterprises establish a set of effective and sensitive financial early warning models, track and monitor the company's financial operation process when the financial situation is operating normally, comprehensively analyze the company's financial data indicators, issue early warning signals, and inform operators of the potential dangers faced by the company, so that they can take corresponding measures to eliminate or weaken the degree of destruction of multiple companies in the industrial chain and avoid greater losses, in order to maintain an undefeated position in the fierce market competition [1].

As the main body of self-management, self-financing and self-development, the listed companies are faced with the ever-changing market environment and will be tested by financial crisis at any time. Therefore, for listed companies, an accurate, effective, sensitive and stable financial crisis early-warning model can make them insight into the financial crisis before it occurs, quickly take appropriate measures to effectively prevent, to help listed companies to prevent financial crisis, to help enterprise managers to better manage enterprises, is also conducive to improve the regulatory capacity of government departments, reduce the risk of our country's securities market, it plays an important role in the stability and healthy development of our financial order.

Based on the deep analysis of the current financial problems and operating conditions of listed companies in our country, this paper selects the financial statement data of listed companies in Shanghai and Shenzhen stock markets from 2018 to 2021, from the six angles of listed companies' financial status, such as profitability, debt paying ability, management ability, development ability, profit quality and valuation analysis, by using factor analysis method and Logistic regression model, this paper reveals the signaling function of financial index in the development of enterprises, and proves that under the function of scientific, effective, sensitive and accurate financial early-warning model, enterprises can better avoid risks to avoid bankruptcy, healthy management.

ISSN:2790-1661 DOI: 10.56028/aemr.4.1.52.2023

2. Literature Review

A financial crisis is an economic phenomenon in which a business is unable to pay its debts or expenses as they fall due, ranging from the failure of money management techniques to bankruptcy and everything in between. It exists in all life cycles of the enterprise and has the uncertainty of outbreak, so an effective financial early warning system is very important [2].

The earliest financial early warning models originated in the West. Fitzpatrick used the unitary discriminant analysis method to predict the bankruptcy of the company, and the results showed that the two indexes with the strongest discriminative ability of financial crisis were shareholders' equity/liabilities and net profit [3]. But for a company, simply using a variable to measure the health of its financial situation is too one-sided and unconvincing. Altman used 33 bankrupt companies and 33 non-bankrupt companies as a research sample, and established a Z-Score model by using multiple discriminant methods [4]. The premise of the use of the multivariate discriminant method is that the covariance matrices of each group of samples are equal and the multivariate normal distribution between the discriminant variables is required, which is difficult to meet in reality, so its use value is very small [5]. Ohlson made significant improvements to the financial early warning system using a logistic regression method, which is loosely required for sample data, and found that the four statistically significant factors in estimating the likelihood of bankruptcy were firm size, capital structure, operating performance, and short-term liquidity [6]. Because logistic regression models are sensitive to multicollinearity, if explanatory variables with a high degree of correlation are present, the prediction results of the model will be affected. As an improvement, Koyuncugil and Ozgulbas combined with data mining technology to study and apply the enterprise financial early warning model [7].

In recent years, relying on China's national conditions and the characteristics of domestic listed companies, Chinese scholars have established a variety of financial early warning models. Huang Dezhong and Zhu Chaoqun selected the data of 96 normal companies that were first matched by 48 listed companies of ST from 2010 to 2013, respectively constructed a financial early warning model containing only commonly used financial indicators and an asset quality model containing asset quality indicators, compared the results of these two early warning models, and found that the introduction of asset quality indicators can improve the accuracy of financial risk early warning models [8]. Some scholars believe that surplus management variables have a great influence on the prediction results of the financial early warning system, such as Gu Xiaoan, Wang Bingqi, Li Wenqing used the 2012-2016 A-share listed companies as a sample to establish a Logistic regression model, and introduced earnings management financial indicators, which proved that the financial early warning model after the introduction of earnings management indicators was better than before the introduction [9].

Most scholars' research on financial early warning models is based on the static analysis of cross-sectional data at a certain point in time, but the occurrence of financial crisis is not a static process, but a dynamic result of gradual deterioration, so cross-sectional data cannot particularly reflect the dynamic trend. Based on the traditional static indicators, Fan Linyu introduces vertical financial information indicators that reflect the development trend of enterprises, and constructs a more efficient dynamic financial crisis early warning model, and the results show that the improved model introducing longitudinal financial information has better goodness of fit and forecast accuracy [10].

3. Research Design

3.1 Sample Selection and Data Sources

This paper selects a total of 50 listed companies in Shanghai and Shenzhen, of which the training sample includes 11 listed companies that lost money for two consecutive years in 2018 and 2019 and 39 listed companies with normal financial conditions in 2020. In the selection of training

ISSN:2790-1661 DOI: 10.56028/aemr.4.1.52.2023
samples of ST company, if the data of three years or more before the financial crisis of listed companies are used, the model will have a poor effect and the prediction result will be inaccurate. According to the ST definition of the listed company, the financial situation of the listed company is abnormal for two consecutive years, and the average of its financial data of the previous two years by ST is taken as the financial data of ST company, and finally the financial data of the listed company's 2021 annual report is used as the test object of the model, that is, the test set, to describe the accuracy of the model warning. The data used in this article is mainly derived from the CSMAR database.

3.2 Main Variables

This list extracts 20 financial indicators from six aspects, including profitability, solvency, operating ability, development ability, profit quality and valuation analysis, as the variables of the research sample: total asset profit margin, EBIT and total assets, operating net interest rate, return on net assets attributable to the parent enterprise TTM, current ratio, quick ratio, total corporate working capital formed by active operation/enterprise current liabilities, asset-liability ratio, inventory turnover rate, current asset turnover rate, net asset turnover speed, Net asset growth rate, net profit growth rate, operating income growth rate, net flow growth rate formed by active enterprise operation, net cash content of total assets, cash content of operating income, operating indicators, net cash flow per share of operating activities, price-to-book ratio.

4. Empirical Results and Analysis

4.1 Descriptive Statistics

Table 1 shows the descriptive statistical results for each variable. It can be seen that the mean values of the variables X14, X21, X22, X23, X24 and X41 are close to the median, indicating that the data of these indicators are centrally distributed and discrete. The standard deviations of X31, X44, X51 and X53 are 12.0228, 16.8418, 11.5578 and 32.2159, respectively, which indicates that some relevant indicators involving operating capacity, development trend and profit quality of listed companies in the research sample are quite different.

Table 1. Descriptive statistics of variables

Variable abbreviation	Number of cases	maximum	minimum	average value	median	standard deviation
x11	50	0.1072	-0.4999	0.0000	0.0182	0.0911
X12	50	0.1147	-0.6318	0.0133	0.0335	0.1113
X13	50	0.3077	-27.1465	-0.6028	0.0316	3.8520
X14	50	0.4201	-1.8918	-0.0597	0.0450	0.3875
X21	50	5.3173	0.2457	1.7197	1.4445	1.0229
X22	50	5.1479	0.2372	1.2597	1.0121	0.9653
X23	50	0.7150	-0.9700	0.0597	0.0734	0.2394
X24	50	0.9481	0.0683	0.5715	0.5865	0.2251
X31	50	56.5507	0.0179	7.8511	3.3254	12.0228
X32	50	3.8789	0.0162	0.9941	0.6738	0.8513
X33	50	3.0630	0.0122	0.6091	0.3584	0.6425
X41	50	0.5302	-0.4375	0.0458	0.0434	0.1690
X42	50	12.2920	-15.0800	-0.9096	-0.1215	3.7334
X43	50	1.5916	-0.5342	0.0042	0.0092	0.3166
X44	50	96.7650	-29.4666	0.6811	-0.1497	16.8418
X51	50	28.3701	-213.9769	-3.2174	0.9009	32.2159
X52	50	0.7475	-4.4439	0.0042	0.0576	0.6760
X53	50	52.9888	-40.8437	0.0190	0.8479	11.5578

ISSN:2790-1661 DOI: 10.56028/aemr.4.1.52.2023

X54	50	8.4411	-5.7007	0.7364	0.2536	2.1126
X61	50	51.8186	0.1172	3.6831	1.8604	7.4207

4.2 Factor Analysis

4.2.1 KMO and Bartley tests

Factor analysis must be preceded by testing the original variables to determine whether they are suitable for factor analysis, of which KMO and Bartlett tests are currently the most common test methods, calculated as follows.

$$KMO = \frac{\sum\sum_{i\neq j} r_{ij}^2}{\sum\sum_{i\neq j} r_{ij}^2 + \sum\sum_{i\neq j} r_{ij\bullet 1,2\ldots k}^2} \quad (1)$$

From the results of KMO and Bartlett test in Table 2, the KMO value is 0.525, and the results of Bartlett spherical test show that the P value is less than the significance level, indicating that there is a correlation between the variables, and factor analysis can be performed through the test.

Table 2. KMO and Bartley test

KMO value		0.525
Bartlett sphericity test	Approximate chi-square	776.973
	df	190.000
	p	0.000***

Note: ****represent significance levels of 1%.

4.2.2 Eigenvalues and variance contribution rates

The factor analysis of the selected financial indicator variables shows the characteristic values and contribution rates shown in Table 3. There are six large eigenvalues of 4.240, 3.112, 2.567, 2.284, 1.392 and 1.209, and the cumulative contribution rate of variance of the first six factors reaches 74.014%. Therefore, the original 20 financial indicators can be replaced by 6 common factors, which are denoted as F1, F2, F3, F4, F5 and F6.

Table 3. Explanation table of total variance

Total variance explained							
ingredients	Feature root			ingredients	Feature root		
	Feature root	Variance percentage	Variance percentage		Feature root	Variance percentage	Variance percentage
1	4.240	21.198%	21.198%	11	0.576	2.878%	92.527%
2	3.112	15.559%	36.757%	12	0.484	2.422%	94.949%
3	2.567	12.834%	49.591%	13	0.315	1.573%	96.522%
4	2.284	11.42%	61.011%	14	0.286	1.428%	97.95%
5	1.392	6.959%	67.97%	15	0.125	0.624%	98.574%
6	1.209	6.044%	74.014%	16	0.111	0.553%	99.127%
7	0.979	4.894%	78.908%	17	0.083	0.415%	99.542%
8	0.805	4.026%	82.934%	18	0.057	0.287%	99.829%
9	0.720	3.602%	86.536%	19	0.031	0.155%	99.984%
10	0.623	3.113%	89.649%	20	0.003	0.016%	100.0%

Because some factors are related to multiple financial indicator variables, which will increase the difficulty of interpreting financial indicator variables, the factor load matrix transformation is performed with the maximum variance orthogonal rotation, and the results are shown in Table 4. It can be seen that F1 has a large load on total asset profit margin (X11), EBIT and total assets (X12), operating net profit rate (X13), return on net assets attributable to the parent company TTM (X14),

ISSN:2790-1661 DOI: 10.56028/aemr.4.1.52.2023
and net profit growth rate (X42), so F1 is called the enterprise profitability factor; The load of F2 in the asset-liability ratio (X24) is very large, the load of current asset turnover ratio (X32) and total asset turnover ratio (X33) is also at a large value, so F2 is called the solvency factor of the enterprise; F3 has a large load on operating net profit margin (X13), cash content of operating income (X52), and net cash flow per share of operating activities (X54), which can be called the profit quality factor of the enterprise; F4 has a large load on inventory turnover rate (X31), current asset turnover rate (X32), and total asset turnover rate (X33), which can be called the operating ability factor of the enterprise; The load of F5 on the price-to-book ratio (X61) is very large, which can be called the valuation power factor of the enterprise; F6 has a large load on the net profit growth rate (X42) and can be a factor for the development ability of the enterprise.

Table 4. Table of factor load coefficients after rotation

component	Factor load matrix						Commonality (common factor variance)
	F1	F2	F3	F4	F5	F6	
1	0.916	0.096	0.143	0.041	0.060	0.037	0.876
2	0.881	0.044	0.153	-0.007	0.076	0.073	0.813
3	0.147	0.176	0.885	0.083	-0.011	0.059	0.846
4	0.683	0.052	0.269	0.142	-0.053	-0.125	0.581
5	-0.017	0.943	0.049	-0.035	-0.020	-0.112	0.906
6	-0.045	0.947	0.012	-0.070	-0.026	0.070	0.910
7	0.517	-0.443	0.202	-0.091	-0.159	-0.091	0.546
8	-0.164	-0.730	-0.126	0.215	0.335	0.156	0.758
9	0.065	-0.040	0.114	0.926	0.034	0.168	0.907
10	-0.116	0.101	0.131	0.353	-0.046	0.590	0.516
11	0.077	-0.098	0.072	0.898	-0.043	0.178	0.861
12	0.686	-0.027	-0.438	0.194	-0.106	-0.206	0.754
13	0.375	0.538	0.039	0.144	0.118	0.231	0.521
14	0.802	0.077	0.005	-0.070	-0.407	0.138	0.839
15	-0.073	0.083	0.095	-0.043	-0.110	-0.812	0.695
16	-0.094	-0.033	0.165	0.035	0.835	-0.147	0.757
17	0.111	0.054	0.916	0.028	0.057	-0.071	0.864
18	-0.060	-0.050	-0.015	-0.018	0.747	0.279	0.642
19	0.308	-0.216	0.618	0.097	0.282	-0.084	0.619
20	0.065	-0.025	-0.473	0.539	0.155	-0.227	0.594

4.3 Logistic Regression Model

The essence of the logistic model is a linear probability model to solve the 0-1 problem, whether the listed company is facing a financial crisis, in fact, is also a 0-1 problem, the result is nothing more than the existence of a financial crisis and the absence of a financial crisis, so in the choice of model, this paper adopts the logistic regression model, and its formula is as follows.

$$p = \frac{1}{1+\exp(-(\beta_0+\sum_{i=1}^{k}\beta_i x_i))} \quad (2)$$

4.3.1 Logistic model metermination

Assume that the dependent variable of the model is Y, which follows binomial distribution (namely st and non-ST), and the independent variables are F1, F2, F3, F4, F5 and F6. The specific coefficients of the regression equation can be obtained by analyzing the sample data through SPSS software, as shown in Table 5. From the analysis of the results of the F test, we can see that the significance P value is 0.000989, which rejects the null hypothesis, so it can be considered that the model constructed in this paper basically meets the requirements. For variable collinearity, VIF is

ISSN:2790-1661 DOI: 10.56028/aemr.4.1.52.2023
all less than 10, so the model has no multicollinearity problem, and the conclusion that the model is well constructed.

Table 5. Variables in the equation

item	Non-normalized coefficients		Normalization coefficients	T	p	VIF	R^2	R^2 *	F
	B	standard error	Beta						
Constant	0.2800	0.0584	-	4.799	0.000***	-			
F1	-0.2167	0.0590	-0.436	-3.676	0.001***	1			F=4.659
F2	0.0860	0.0589	0.173	1.459	0.152	1	0.394	0.309	
F3	0.0060	0.0589	0.012	0.102	0.919	1			P=0.000989
F4	0.0975	0.0589	0.196	1.655	0.105	1			
F5	-0.1126	0.0589	-0.227	-1.911	0.063*	1			
F6	-0.1434	0.0589	-0.289	-2.432	0.019**	1			

Note: ****, **, and * represent significance levels of 1%, 5%, and 10%, respectively.

The corresponding logistic regression model is

$$p(Y)=\frac{1}{1+\exp\left(-\left(0.280-0.217\mathrm{F}_1+0.086F_2+0.006F_3+0.098F_4-0.113F_5-0.143F_6\right)\right)} \quad (3)$$

4.3.2 Sample data logistic regression results

Make predictions on the sample data of the training set, substitute the data into the above logistic regression model, and calculate their probability values p. When the probability value p is less than 0.5, it is considered that the company's financial situation is in an abnormal state and is likely to face ST financial crisis; when p is greater than 0.5, the company's financial situation is considered to be in a normal state. The accuracy of the model for normal companies is 81.82%, the accuracy of judgment for companies with financial crisis is 97.44%, and the overall accuracy rate has reached 94%, and the specific classification situation and results are shown in Table 6.

Table 6. Classification table

Measured		forecast		
		ST value		Correct percentage
		0.0	1.0	
ST value	0.0	9	2	81.82%
	1.0	1	38	97.44%
Overall percentage				94%

4.3.3 Statistics on the accuracy rate of early warning

The test set data is substituted into the model, , and the prediction results are shown in Table 7, and the overall prediction accuracy rate is 90%, indicating that the financial early warning model constructed in this paper can make a more accurate judgment on the financial status of listed companies.

Table 7. Classification table

Measured		forecast		
		ST value		Correct percentage
		0.0	1.0	
ST value	0.0	8	3	72.7%
	1.0	2	37	94.87%
Overall percentage				90%

ISSN:2790-1661 DOI: 10.56028/aemr.4.1.52.2023

5. Conclusion

Using the financial data of listed companies from 2018 to 2021, the financial early warning model of listed companies was constructed by factor analysis and Logistic regression method, and the p-value reflecting the probability of financial risk of enterprises was obtained. The results show that profitability-related indicators have a great impact on the health of the company's financial status and are the financial fundamentals of the company. The early warning model obtained in this paper has high discriminating ability and can effectively identify whether an enterprise has a financial crisis. By observing the contribution rate of factors, it is found that the accuracy of traditional financial indicators is improved after adding profit quality and market valuation, so it is believed that the financial status of enterprises is greatly affected by the market, and operators need to improve their sensitivity to market changes. The research results of this paper can provide a good financial reference for the majority of investors and business operators, and also provide some reference for enterprises to avoid risks.

References

[1] Wang Qi. Research on financial early warning system of RFS (Shanghai) Co., Ltd.[J]. China Economy and Trade,2011(22):179.

[2] Long Jingnan. An empirical study on the financial early warning system of biopharmaceutical listed companies in China based on logistic regression model[J]. Oriental Corporate Culture,2010(4):93-94.

[3] Fitzpatrick P.J.. A comparison of ratios of successful industrial enterprises with those of failed firms[J]. Certified Public Accountant,1932(2):589-605.

[4] Altman E.I.. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy[J]. Journal of Finance, 1968,23(4):589-609.

[5] Chen Xinxin, Guo Hongtao. The combined application of factor analysis and logistic regression in financial early warning of agricultural listed companies[J]. Mathematical Statistics and Management, 2022(1):11-24.

[6] Ohlson J.. Financial ratios and the probabilistic prediction of bankruptcy[J]. Journal of Accounting Research, 1980,18(1):109-131.

[7] Koyuncugil A.S., Ozgulbas N.. Financial early warning system model and data mining application for risk detection[J]. Expert Systems with Applications,2012, 39(6):6238-6253.

[8] Huang Dezhong, Zhu Chaoqun. Financial risk early warning model for enterprise asset quality index[J].Finance and Accounting Monthly, 2016(8):48-52.

[9] Gu Xiaoan, Wang Bingqi, Li Wenqing. Research on Improvement of Early Warning Accuracy of Logistic Financial Early Warning Model-Analysis of Introducing Earnings Management Variables[J].Journal of Nanjing Audit University,2018,15(4):45-52.

[10] Fan Linyan. Research on early warning of financial crisis by introducing longitudinal financial information[D].Xiamen:Jimei University,2016.