Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

Research on Financial Warning for Chinese Listed Companies by Using Panel Data Model

Xijuan Shao
Institute of Emerging Industrialization Development,
School of Business Administration
South China University of Technology
Guangzhou, P.R. China.510640
Email: xjshao@scut.edu.cn

Yaohui Chen
Institute of Emerging Industrialization Development,
School of Business Administration
South China University of Technology
Guangzhou, P.R. China.510640
Email: chencyh3312@gmail.com

Haibo Wang
A.R. Sanchez School of Business,
Texas A&M International University
Laredo, Texas. USA. 78041
Email: hwang@tamiu.edu

***Abstract*—This paper studies how to establish models for predicting financial distress in China's listed companies. We firstly select 26 companies with financial distress and 54 matching companies' panel data as samples, then use panel data model to conduct an empirical study. The research indicates that: (1) The predictability precision is 91.25%、92.5%、91.25% and 87.5% for T-1、T-2、T-3 and T-4, respectively, superior to the previous research. (2) The panel data model warning analysis is of stable predictability, which is suitable in constructing mid-term and long-term prediction models.**

***Keywords*—Financial crisis warning; Panel data mode; Random effects model**

I. INTRODUCTION

Nowadays, companies are suffering from financial crisis which is caused by the sub-prime mortgage loan crisis. Conflicts within the macroeconomic development have shown in the performance of the company, especially the financial risk faced by the companies is increasing. Therefore, the financial crisis warning system is one of the key aspects of the company risk management. Both academics and the practitioners are paying much more attentions to this issue.

Since the mid-20th century, researches about financial crisis prediction had been carried out for nearly 40 years. The pioneering researches included Beaver (1966) and Altman (1968) worked in business failure classification by the use of the univariate and multivariate discriminant analysis (MDA). The former used univariate discriminant analysis with 5 financial ratios as variables, 79 bankrupt firms and 79 matching non-bankrupt firms as samples to predict the financial crisis and found that cash flow and debt to assets can more actually discriminate the companies' financial status. The essence of the latter technique is an issue of classification under several assumptions. It assigns a score (the Z score) to each of the observed firms. The Z score is a linear combination of several independent variables and a cutoff score is estimated to divide the firms into healthy and unhealthy ones. The accuracy rates of the first three years of the Z score model are 95%, 72%, 48%. Obviously, the Z score model is useless to predict financial crisis after the second year. Since then, extensive researches had been conducted to predict financial crisis. Researches mainly focus on the following two aspects:

(1) Models for financial crisis prediction. In the area of financial crisis prediction research, the most popular methods are multivariate linear discriminant analysis and logistic regression methods. Since the nineties twenty-first century, the neural network technology had been gradually applied to the financial crisis studies, so far, many had made considerable achievements. Kumar and Bhattacharya (2006) used BP neural network model to construct the financial warning model and compared with the discriminant analysis, results showed that the predicting precision of BP neural network model is better than discriminant analysis, furthermore neural network model can be applied to missing data, and therefore can be applied in a wider scope. Yang Shu-e and Wang Le-ping (2007) constructed a BP neural networks model to predict the financial status of Chinese listed companies, by taking panel data composed of the financial data of T-2 and T-3 as the sample. The research indicated that with panel data, the BP neural networks warning analysis was of stable and continuous predictability, which was suitable in constructing practical mid-term and long-term prediction models to make the model more applicable and the predictability precision was 88.46 % and 75.64 % for T-3 and T-4, respectively, superior to that of counterparts and the previous research.

(2) The selection of Variables. Since Beaver (1966) began to study the issue of the financial crisis prediction, most of the extensive researches also used financial ratios as the variables of discriminant analysis. Yang Shu-e and Huang Li (2005) added the information of cash flow as the discriminant variables based on the traditional financial indicators so that the prediction model was more effective. Wu Ying-yu, Cai Qiuping and Wu Fan (2008) considered that the company's financial status was not affect by financial factors or non-financial factors alone, but rather the result of both, so they combined with financial and non-financial indicators to improve the companies' financial warning model.

However, from the traditional model of discriminant analysis and logistic regression model to the neural network model, the difference between them is only the arise of precision and the advance of the methods, in the data collection section, studies mostly used cross-section data as samples, thus, they had not taken into account the business continuity of the companies' financial status. Therefore, the results of such models prevalence of low accuracy and decline rapidly. It is not suitable in predicting mid-term and long-term financial crisis. However, the panel data contains cross-sectional and time series information, it not only considers the difference between the financial status of different companies, but also takes into account the time continuity of the companies' financial status,

978-1-4244-2794-9/09/$25.00 ©2009 IEEE

thus, panel data model should be more scientific than others warning models.

By studying the relevant literatures about panel data model, we find that the earliest studies about panel data collection and model was the research about income dynamics done by the University of Michigan and the survey about countries' labor market. Wang Wei (2008) analyzed the determinants of household saving with the dynamic panel data on Chinese urban and rural household saving rate at provincial level of the period 1995-2004. Tian Zhe-yong (2008) used a panel data, applying the fixed coefficient model to study the relation empirically between the investments and the regional economic growth in China. The results showed that the fixed asset investments was the most important cause for the economic growth in China, the private economy was the most force for the regional economic growth stably due to the high output elasticity. However, it's found that the panel data model is mainly used in the researches about macro-economic, and is almost not applied in the area of financial management researches. Only Lu Rui-di (2007) did the empirical research by using panel data model with 81 Chinese listed companies as samples. Her research found that the panel data model warning analysis is of high accurate and stable predictability, the predictability precision is 100%、93.83%、87.65% and 85.19% for the first forth years. However, in the variables selection, it is subjective for her to select ROA, retained earnings to total assets, debt-to-assets, working capital to total assets, assets turnover as variables by the historical frequency of the variables had been selected in the former researches.

Therefore, through using the empirical model frame of Lu Rui di (2007). However, the methodology of choosing the variables is modified in this paper. Namely, variables are chosen from the amounts of financial indicators to represent the five key aspects of corporate capabilities. The rest of the paper is organized as follows. In section 2, we introduce the Sample selection and variables construction. In section 3, we select and estimate the panel data model. In section 4, we present and analyze the empirical results. In section 5, we come up with a conclusion.

II. SAMPLE SELECTION AND VARIABLES CONSTRUCTION

A. Sample selection

The sample includes financial crisis and matching non-financial crisis companies.

(1) Definition of financial crisis companies. Researchers have made definition of financial crisis companies with different standards from vary angles. Walters (1957) considered the exhaustion of the cash as the determination. Deakin (1972) definited bankrupt, default or liquidation as financial crisis. Blum (1974) made the definition of corporate debt are not fully paid when debt are due. Golstein (1988) used a dynamic process to define the financial crisis companies in three groups. The first group is management failure, referring to the companies failed to achieve their required return on investment or significantly lower than similar competitors, which means poor performance. The second group is financial loser, referring to the losers suffered in serious and long-term losses so that it is not sufficient to settle the liabilities. The third group is legal losers, referring to the legal declaration of bankruptcy for the reason that it can not reverse the poor performance for long time. However, in Chinese, because of the undeveloped of capital market, as a result, we could not follow foreign researchers' routes to conduct our research, instead, we turned to the stock market and chose ST and non-ST companies as our samples.

According to the definition of financial crisis and the sample selection methods, firstly we use financial panel data of Chinese listed companies from 1996 to 2007, considering in the period of 2006-2007 Chinese economic and capital market are developing in high-speed, the number of ST companies are significantly less than previous years, so we just use 1996-2005 financial panel data as the research object. Furthermore, we select 26 ST companies that have had negative cumulative earnings over two consecutive years or net asset value (NAV) per share below its par value as the financial crisis companies' sample.

(2) Definition of matching non-financial crisis companies. We select 54 non-ST companies which have the similar assets at the same time and also in the same industries as the comparable firms to test the correct of the model constructed in this paper.

Sample size of each group is shown in Table I. All of the data comes from Resset and Gildata financial database and uses the data at the end of each year.

Table I Research samples of each group

	ST	**Non-ST**	**Total**
Shanghai A	12	26	38
Shenzhen A	14	28	42
Total	26	54	80

B. Variables construction

(1) Selection of dependent variables. According to II.A, we use net profit from the income statement as the dependent variables, referring that companies will be special treatment as long as they have negative cumulative earnings over two consecutive years or net asset value (NAV) per share below their par value.

(2) Selection of independent variables. To improve the subjective selection method of independent variables in the paper of Lu Rui di (2007). We use the objective quantitative methods similar to the relative financial warning model researches to determine the independent variables of our model. Firstly, 18 financial ratios are selected and calculated to cover the aspects of profitability, solvency, asset management efficiency, sustainable growth and cash flow information see Appendix 1. Furthermore, to exclude the correlation and multicollinearity between the variables, we extract the above financial ratios by correlation analysis, excluding the financial ratios with greater correlation than 0.7. Finally, we use the remaining 10 financial ratios as the independent variables of our model. See Table II below.

Table II Financial ratio set

	Ratios		**Ratios**
Liquidity	X_1=current ratio	**Solvency**	X_4=debt to assets
	X_3= net working capital to total assets		X_5=interest coverage
Profitability	X_6= net profit margin	**Growth**	X_{13}=growth rate of operational profit

	X_7=ROA		X_{14}=retained earnings to total assets
	X_8=ROE	**Cash flow**	X_{18}=operational cash flow to total debt

III. SELECTION AND ESTIMATION OF PANEL DATA MODEL

A. Panel data model classification

The basic model given in (3.1) is variable coefficient model.

$$y_{it} = \alpha_i + \beta_i x_{it} + \varepsilon_{it}$$

i=1，2…80 t=1996，1997…2005 (3.1)

Where i is the different individual corresponding to the panel data, t is the different time corresponding to the panel data. $x_{it} = \left(x_{it}^1, \ldots x_{it}^k\right)^T$ represents a 1*k vector where k is the number of independent variables. The cross-section coefficients and slope coefficients are vary with the cross-section individuals.

Assuming the slope coefficients of model(3.1) are constant, it can obtain the variable intercept model which is given in (3.2).

$$y_{it} = \alpha_i + \beta x_{it} + \varepsilon_{it}$$

i=1，2…80 t=1996，1997…2005 (3.2)

Furthermore, assuming the cross-section coefficients of model(3.2) are also constant, it can obtain the mixed regression model which is given in (3.3).

$$y_{it} = \alpha + \beta x_{it} + \varepsilon_{it}$$

i=1，2…80 t=1996，1997…2005 (3.3)

There are two main approaches to validate the kind of the panel data model, such as F-test and Hausman test.

B. Model selection

There are many approaches to test the types of panel data model. We use Hausman test to examine the models constructed by the panel data T-1, T-2, T-3, T-4 respectively. Results show that the probability P values of the above four models are less than the critical value at 5 % level of significance, which means it should not use fixed-effect model rather than should use random effects model. Hausman test results present in TableIII.

Table III Hausman test results

Samples	P value	Conclusion
T-1 (1996-2004)	0.0933	Random effects model
T-2 (1996-2003)	0.1105	Random effects model
T-3 (1996-2002)	0.2225	Random effects model
T-4 (1996-2001)	0.2839	Random effects model

Meanwhile, according to the experience, we expect to analysis the general through the empirical research result, referring to use the specific constant term as the random distribution of the cross individual differences. In other words, the 80 listed companies' samples are randomly selected from the total listed companies in Shanghai and Shenzhen stock market. So random effects model should be selected which is consistent with the Hausman test results.

C. Model estimation

By using the random effects model, panel data of 80 Chinese listed companies from 1996-2004、1996-2003、1996-2002、1996-2001 is used as the research object respectively. We estimate the T-1, T-2, T-3, T-4 panel data model (results see Table IV) respectively which represent the financial warning models four years prior to their crisis

Table IV Estimation results of models

Variables	T-1	T-2	T-3	T-4
C	-2181273	-9734284	764421.8	14672623
X_1=current ratio	-1980084	4708695	2494959	-5845131
X_3=net working capital to total assets	-31435742	-56592496*	-57845908*	-37064308
X_5=interest coverage	5134.778	-32344.7	-39875.38	130990.6
X_4=debt to assets	46566952*	54211840*	35853767*	26682707
X_6=net profit margin	4477611	2430370	2042395	24538609
X_7=ROA	1.02E+09***	9.09E+08***	9.28E+08***	8.77E+08***
X_8=ROE	817412.6	3713997	1907028	2255996
X_{14}=retained earnings to total assets	41325014*	42599064**	45348567**	42509010
X_{13}=growth rate of operational profit	169208.3	-61512.4	-11658.28	345800.6
X_{18}=operational cash flow to total debt	38468352**	15550106	-4406878	-10680236
R^2	0.558034	0.552437	0.59413	0.649153
F value	51.38847	41.84348	39.67023	37.00492
P value	0.00	0.00	0.00	0.00

Values significantly different from zero at 10%, 5% and 1% are marked *, ** and *** respectively

According to Table IV, we can construct four panel data warning models in different observation points respectively. See formulas below.

T-1 panel data warning model:

$$y_{ij} = \gamma - 2181273 - 1980084x_{1j} - 31435742x_{3j} + 5134.778x_{5j} + 46566952x_{4j} + 4477611x_{6j} + 1020000000x_{7j} + 817412.6x_{8j} + 41325014x_{14j} + 169208.3x_{13j} + 38468352x_{18j}$$

γ presents the random influence coefficient, see Appendix II.

T-2 panel data warning model:

$$y_{ij} = \gamma - 9734284 - 4708695x_{1j} - 56592496x_{3j} - 32344.7x_{5j} + 54211840x_{4j} + 2430370x_{6j} + 909000000x_{7j} + 3713997x_{8j} + 42599064x_{14j} - 61512.4x_{13j} + 15550106x_{18j}$$

γ presents the random influence coefficient, see Appendix III.

T-3 panel data warning model:

$$y_{ij} = \gamma + 764421.8 + 2494959x_{1j} - 57845908x_{3j} - 39875.38x_{5j} + 35853767x_{4j} + 2042395x_{6j} + 928000000x_{7j} + 1907028x_{8j} + 45348567x_{14j} - 11658.28x_{13j} - 4406878x_{18j}$$

γ presents the random influence coefficient, see Appendix IV.

T-4 panel data warning model:

$$y_{ij} = \gamma + 14672623 - 5845131x_{1j} - 37064308x_{3j} + 130990.6x_{5j} + 26682707x_{4j} + 24538609x_{6j} + 877000000x_{7j} + 2255996x_{8j} + 42509010x_{14j} + 345800.6x_{13j} - 10680236x_{18j}$$

γ presents the random influence coefficient, see Appendix V.

As can be seen from Table IV, among the four models, debt to assets, net profit margin, ROA, ROE, retained earnings to total assets change in the same direction with the net profit, that means it is useful to improve the companies' financial status by increasing the above 5 ratios.

On the contrary, working capital to total assets changes in the opposite direction with net profit, that means although the increase in working capital will benefit to enhance the solvency, because of its poor profitability, it is not conducive to improve the net profit.

Through comparing the coefficients, it is easy to find that the coefficients of working capital to total assets, debt to assets, ROA, retained earnings to total assets , cash flow to total debt are significantly larger than the coefficients of another indicator which means that the above 5 ratios play decisive roles in companies' financial status. Through further analysis, it can be found that working capital to total assets represents solvency, ROA represents profitability, retained earnings to total assets represents the growth rate of the company, cash flow to total debt represents cash flow information which means that the determinants of financial status represents one of the abilities of companies. It can be reflected that companies' financial status are not only relate to the profitability of the business, but also have a great relationship with liquidity, solvency, growth and cash flow. It indicates that in the operation and investment process, on the one hand, it should pursuit the profit generated from the operating activities, on the other hand, the influence of the operating and investment activities must be taken into consideration. Only trading-off the benefits and losses, it can attain sustain growth in net profit.

IV. EMPIRICAL RESULTS AND ANALYSIS

Using the methods mentioned above, we calculate the net profit of listed companies from 2005, 2004-2005, 2003-2005, 2002-2005 respectively and then compare with the actual net profit of the corresponding companies. According to the results, we obtained the predictability precision of the above four models. As shown in Table V.

Table V predictability precision of models

			Warning value		
		Groups	**ST**	**Non-ST**	**Total**
T-1	**Numbers**	**ST**	19	7	26
		Non-ST	0	54	54
	Percent	**ST**	73.08%	26.92%	100%
		Non-ST	0	100%	100%
Precision			**91.25%**		
			Warning value		
		Groups	**ST**	**Non-ST**	**Total**
T-2	**Numbers**	**ST**	20	6	26
		Non-ST	0	54	54
	Percent	**ST**	76.92%	23.08%	100%
		Non-ST	0	100%	100%
Precision			**92.5%**		
			Warning value		
		Groups	**ST**	**Non-ST**	**Total**
T-3	**Numbers**	**ST**	20	6	26
		Non-ST	1	53	54
	Percent	**ST**	76.92%	23.08%	100%
		Non-ST	1.85%	98.15%	100%
Precision			**91.25%**		
			Warning value		
		Groups	**ST**	**Non-ST**	**Total**
T-4	**Numbers**	**ST**	19	7	26
		Non-ST	3	51	54
	Percent	**ST**	73.08%	26.92%	100%
		Non-ST	5.56%	94.44%	100%
Precision			**87.5%**		

From Table V, the panel data warning analysis is of stable and continuous predictability, the predictability precisions are 91.25%, 2.5%, 91.25%, 87.5% respectively, superior to the logistic regression warning model constructed by Wu Shi-long (2001) whose predictability precisions are 93.53%-73.17%, and the BP neural networks warning model constructed by Yang Shu-e and Wang Le-ping (2007) whose predictability precisions are 89.74%-75.64%.

Furthermore, panel data warning models have higher precision in predicting non-financial crisis companies than multivariate linear discriminant analysis, logistic regression methods, and neural networks methods, the predictability precision four years prior to their crisis are 100%, 100%, 98.15%, 94.44% respectively.

In addition, we note that the predictability precision of T-1 warning model is less than T-2 warning model, this may be for the reason that T-1 warning model just only predict the net profit previous one year, lacking of considering the influence of time continuity factors. As a result, it is easy to suffer in the earning manipulation. Thus, there is a certain deviation in judging the ST companies base on the standard that companies will be special treatment as long as they have negative cumula-

tive earnings over two consecutive years. However, predictability precision of T-1 warning model is just a little less than the T-2 warning model, moreover, the predictability precisions of T-2, T-3, T-4 warning model are declining smoothly which shows that the information contained by the panel data warning models constructed in this paper have already covered the characteristics of financial crisis and the models are able to identify the potential financial crisis of Chinese listed companies with satisfactory accuracy even four years prior to their crisis.

Through the above analysis, we consider that panel data model is suitable in constructing practical both mid-term and long-term prediction models. It has higher predictability precision even four years prior to their distress than others models. What is more, the performance of the model will present stable estimation status along with the advancement of the observed periods of which the predictability precision is declining slowly with the observation time advancing. So, it is useful to make the model in warning application.

V. CONCLUSIONS

This paper studies how to establish models for predicting financial crisis in Chinese listed companies. We firstly select 26 companies with financial crisis and 54 matching companies' panel data as samples, 10 financial indicators will be selected as the variables from 18 financial indicators which represent the capabilities of five key aspects via correlation analysis. And the net profit will be deemed as the explained variable, constructing panel data models to forecast Chinese listed companies' financial status and bringing out the following conclusions:

(1) Panel data model is suitable in constructing practical mid-term and long-term prediction models. Empirical research results indicate that the predictability precisions are all above 90% for T-1, T-2 and T-3. Even for T-4, the predictability precision can also achieve 87.5%, which is higher than the traditional statistics models and artificial intelligence models.

(2) The performance of the panel data model does not decrease dramatically along with the advancement of the observed periods, instead, it present the trend of falling downward. Hence, this is considered to be beneficial to the estimation of the corporate financial crisis in the middle and long term perspective.

The financial crisis warning model is constructed base on the panel data of Chinese listed companies. Therefore, the panel data warning model is suitable to forecast the financial crisis of Chinese listed companies. However, how effective of this model used to predict financial crisis of foreign listed companies need further study.

REFERENCES

[1] Altman,E.Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. Journal of Finance, 1968, pp.589-609.

[2] Beaver,W. Financial Ratios as Predictors of Failure. Journal of Accounting Research, 1966, pp.71-111.

[3] Blum,M. Failure Company Discriminant Analysis. Journal of Accounting Research, 1974, pp.12,1-25.

[4] Deakin, E. B. A Discriminant Analysis of Predictors of Business Failure, Journal of Accounting Research, 1972, pp.167-179.

[5] Golstein, A. S. Corporate Comeback. John Wiley & Son.24. 1988.

[6] Kuldeep Kumar and Sukanto Bhattacharya. Artificial neural network vs linear discriminant analysis in credit ratings forecast A comparative study of prediction performances.Review of Accounting and Finance, 2006, 5, pp.216-227.

[7] Walters, J. E.The Determination of Technical Solvency. Journal of Business,1957, vol.30, pp.30-45.

[8] Zmijewski, M. E. Methodological Issues Related to the Estimation of Financial Distress Prediction Models.Journal of Accounting Research, 1984, vol.22, pp.59-86.

[9] Yang Shu-e, Wang Li, Financial Crisis Warning Model based on BP Neural Network. Systems Engineering-Theory&Practice,2005, vol.1, pp.12-19

[10] Yang Shu-e, Wang Le-ping, Research on Financial Warning for Listed Companies by Using BP Neural Networks and Panel Data. Systems Engineering-Theory&Practice, 2007, vol.2, pp.61-67.

[11] Wu Ying-yu, Cai Qiu-ping Wu Fan. Pre-warning study of corporations'financial crisis based on ANN technique. Journal of Southeast University, 2008, vol.1, pp.22-26.

[12] Wang Wei. Determinants of Chinese Household Saving Rate:An Analysis Based on Dynamic Panel Data at Provincial Level of the Period 1995-2005. Journal of Finance and Economics, 2008, vol.2, pp.53-64.

[13] Tian Zhe-yong, Jiang Ke-shen, Xie Zhong-qiu. An Comparable Research of Fixed Assets Investments to China's Economic Growth—Based on the Eastern, Central and Western Panel Data. Forecasting, 2008, vol.1, pp.29-40.

[14] Lu Rui-di. Financial Crisis Warning Model based on Panel data: Evidence from Chinese listed companies. Master degree paper of Zhejiang University,2007.

[15] Wu Shi-nong, Lu Xian-yi. Study of Models for Predicting Financial Distress in China's Listed Companies. Economic Research Journal, 2001, vol.6, pp.46-55.

All Appendix tables are available at http://www.myikmc.com/appendix.doc