International Journal of
Organizational Innovation

THE ROLE OF AUDIT QUALITY IN FINANCIAL DISTRESS: EVIDENCE FROM CHINA

Shu-Ling Chang
Department of International Business,
Southern Taiwan University
of Science and Technology, Taiwan, R.O.C.

Long-Jainn Hwang*
Department of Marketing and Distribution Management,
WuFeng University, Taiwan, R.O.C.
*(corresponding author) ljhwang@wfu.edu.tw

Abstract

For stakeholders, to predict the survival probability of a company in a financial crisis is very important. The objective if this article is to investigate whether the firm's financial distress is predictable using artificial intelligence techniques research methods. We analyze whether audit quality is the key factor to affect the occurrence of company's financial distress in China. Using binary choice model and life test method, the evidence indicates that audit quality of the firm is negatively correlated with the probability of firm's financial distress. The finding support that firm with higher audit quality would be more likely to reduce the probability of financial distress.

Keywords: audit quality, financial distress prediction, binary choice model

Introduction

This study investigates whether the firm's financial distress is predictable using neural network of financial engineering research methods. Since the 2008 global financial crisis, investors suffer huge losses in the stock market. How to confirm or determine which companies are unable to survive during the financial crisis is a very worthwhile issue. Claessens and Kose (2013) reviews the prior literature on financial crises and analyze the main factors explaining financial crises. They conclude that financial crisis can be driven by a variety of factors, but financial crises often are preceded by asset and credit booms that then turn into busts. They conclude that it is necessary to put together new data series and to design new methodologies to get a better understanding of crises episodes (p.40).

Most of the prior research use cross-country data to analysis financial crises. Cerutti, Claessens and McGuire (2012) argue that supervisors and other agencies need more and better data to construct even rudimentary measures of risks in the international financial system. Therefore, we collect data from China, a fast developing country, in order to provide more evidence for single country to contribute this line of research.

For a deeper understanding of crises and the policy issues surrounding these events, we design new methods to classify crises in a more robust manner. Using binary choice model, we examine whether periods of financial disruptions are necessarily changing into crises. The empirical results show that audit quality is the key factor to affect the occurrence of financial distress in China, and is negatively related to firm's financial distress. This finding support that demonstrate the firm with higher audit quality are inclined to reduce the probability of financial distress.

The remainder of this paper is organized as follows. Section 2 reviews related literature and develops hypothesis. Section 3 describe empirical design. Section 4 analyzes the empirical results. Section 5 concludes the paper.

Literature Review and Hypotheses

To forecast or predict business financial crises is always one of the major research problems in the accounting and finance fields (Tsai, 2011). The company declared bankruptcy, debt defaults, bank overdrafts and unpaid dividends on preferred stock that is known as the financial crisis (Beaver, 1966). Liu, Uchida, and Yang (2012) claim that Chinese state-owned enterprises performed better during the crisis if the firm performed poorly before the global financial crisis. Claessens and Kose (2013) reviews the prior literature on financial crises focusing on the main factors explaining financial crises, the major types of financial crises, and implications of crises.

Ohlson (1980) presents empirical results of predicting company failure using logistic regression. He developed a measure of Ohlson O-Score, the combination result of a 9 business ratios. Pastena and Ruland (1986) show that size and ownership concentration are important factors to explain firm's financial distress using UK dataset. Almeida and Philippon (2008) proposes a new method for valuing expected financial distress costs and argue that net asset value play a very important role in estimating financial distress.

According to agency theory, managers may not work hard without effective monitoring mechanisms. Instead, they may thoughtlessly expand the size of the firm to increase their own interests or reputation. This will lead to the consequence of damaging the firm's performance and exposes the firm to financial difficulties. Audits, as an effective mechanism, can respond to issues in a timely manner to reduce risk. Lu and Ma (2016) indicate that better audit quality can reduce the likelihood of financial distress.

Kluger and Shields (1989) provide empirical evidence about auditor change behavior to the quality of comparative bankruptcy prediction.

Sundgren (2009) shows that liquidating bankruptcy is less common among Big 4 audited firms. Based on the above analysis, the hypothesis is developed.

Hypothesis: Ceteris Paribus, the audit quality of the firm is negatively related to firm's financial distress.

Data and Research Methodology

Sample Selection

China currently has two exchanges, the Shanghai Stock Exchange (SSE) and the Shenzhen Stock Exchange (SZSE). The types of shares issued by listed companies in the Shanghai and Shenzhen stock markets are divided into A shares and B shares. A-shares represent publicly listed Chinese companies that trade on Chinese stock exchanges such as the Shenzhen and Shanghai Stock Exchanges. These stocks trade in Chinese Yuan (CNY). B-shares are domestically listed foreign investment shares. They list on the Shenzhen and Shanghai exchanges, and trade in foreign currencies. Another H-share, traded on Hong Kong's exchanges, are regulated by Chinese law and are freely tradable by anyone. These shares trade using the Hong Kong dollar (HKD).

Our sample period is from 2002 to 2018. All data of financial variables are taken from Taiwan Economic Journal (TEJ) China Equity Database and two exchanges (SSE and SZSE). We collect quarterly data of 54 predicting variables as independent variables including financial, corporate governance and other variables to include basic information. Our sample excludes those firms in the financial related industry. Table 1 includes all these 54 explanatory variables.

Variable Definitions

The original explanatory variables are collected from TEJ China Equity Database and their definitions are as defined and can be found in the TEJ. These variables can be classified into some functional parts, including profitability, stability analysis of financial structure, short-term solvency analysis, long-term financial stability, management capability, and most importantly corporate governance. We specially consider corporate governance associated variables because they are related to audit quality.

To measure the dependent variable of firm's occurrence of financial crisis, we use whether the company is delisted in the stock market as the dummy variable. To test hypothesis, we measure audit quality as defined in the Taiwan Economic Journal (TEJ) China Equity Database Corporate Governance Module.

Research Methods

Many empirical research methods have been developed to predict financial crisis and bankruptcy. These empirical research methods include data mining (Tae, Namsik, and L. Gunhee, 1999), fuzzy neural network (Li, 2005; Lee, and Booth, 2005), genetic programming (Lensberg, Eilifsen, and McKee, 2006).

Table 1. List of Variables

No.	Variables	No.	Variables	No.	Variables
X1	Number of ordinary shares in issue (thousands)	X19	Quick ratio	X37	Impairment losses on goodwill
X2	Return on assets (net)	X20	Total liabilities / total equity	X38	Impairment of intangible assets
X3	Return on equity (after tax)	X21	The gearing ratio	X39	Suitable for long-term funding ratio
X4	Total net assets ratio	X22	Corporate restructuring costs	X40	Dependence on borrowing
X5	Net asset value (NAV)	X23	Gain or loss on debt restructuring	X41	Inventories and accounts receivable on net worth
X6	Price-Book ratio	X24	Derivative financial assets	X42	Total assets turnover (times)
X7	The net growth rate	X25	Impairment loss of oil and gas asset	X43	Accounts receivable turnover ratio
X8	Operating profit margin	X26	Impairment losses of productive biological assets	X44	Fixed asset turnover
X9	Net profit margin	X27	Impairment losses on construction projects	X45	Net turnover
X10	Outside the industry revenue and expenditure rate	X28	Construction materials impairment losses	X46	Accounts payable turnover
X11	Operating expense ratio	X29	Impairment of fixed assets	X47	Rate per person with
X12	Cash flow ratio	X30	Investment real estate impairment losses	X48	Dividend Yield
X13	Revenue growth rate	X31	Long-term equity investment impairment losses	X49	Earnings
X14	Operating income growth rate	X32	Impairment losses on investments held-to-maturity	X50	P/S ratio
X15	Net income growth rate	X33	Impairment losses on financial assets available for sale	X51	Tradable float rate
X16	Total assets growth rate	X34	For inventory obsolescence	X52	After holding the transfer rate
X17	Growth rate of fixed assets	X35	Bad debt losses	X53	Total shareholding ratio
X18	Current ratio	X36	Impairment losses on other assets	X54	Outstanding shares of ownership rate

This study combines the Artificial Neural Networks (ANN) and Logistic Binary Regression models, to construct the financial crisis prediction mechanism using China stock market data. In the first step, we select significant predictor variables from the artificial intelligence technique to abstractvariables. Then the next step is to test whether the selected variables do have significantly effect on the financial crisis.

First, we retrieve valuable information from a large database by Data Mining in order to explore and analyze data. Data Mining is the core knowledge of the database (Han and Kamber, 2000), which the ultimate goal is to propose valuable information as a basis for a favorable decision (Simoudis, 1996). We combine ANN and Logistic regression model to abstract valuable variables from 54 original variables to predict. Next, to investigate whether there is effect of company's audit quality on financial crisis, we use Chi-square test based on the selected variables in the first step. We separate sample firm into two different types, firm with audit quality equals to one and firm without audit quality equals to zero. Independent samples T-test method is to obtain the expected value.

Third, we further compare the sample listed company with delisted company to find the difference of key indicator variables between them. Using the Binary Choice Model method for listed companies and delisted companies in China stock market may provide more insight of the company's financial crisis verification.

Usually, the measure the probability of occurrence of the dependent variable is:

$$P(y_i = 1|X_i, \beta) = 1 - F(-\beta_0 - \beta_1 x_1 - ..... - \beta_k x_k) \quad (1)$$

However, to overcome the limitations of the linear probability model in this study, consider the model as shown in Binary Choice Model:

$$P(y_i = 1|X_i, \beta) = P(y_i = 1|x_0, x_1, x_2, ..., x_k) \quad (2)$$

Empirical Findings

Experimental Results

The empirical results are based on data mining modeling using combination of ANN and Logistic regression model to abstract key variables from 54 original variables. In order to build the model, we randomly select 60% of data as the training data, and use 20% of the data as the testing data, and the remaining 20% of the data as the verification data.

From Table 2 to Table 4, we show the results of abstracting variables of affecting the occurrence firm's financial crisis in A, B and H shares, separately. These tables basically show that the artificial neural networks (ANN) method seem to focus on the company's profit variables.

Table 2 show that the correct prediction rate is 0.8257 and type 1 error is 0.1237.

About 70% of total 51,431 sample are low audit quality. Elected from 54 original variables by the ANN, the return on assets (net), return on equity (after tax), operating margin, cash flow ratio, operating income growth rate, net income growth rate and quick ratio are the key financial crisis variables.

Table 3 shows that there are 20 key explanatory variables in the B-share listed companies, among them four factors are significant losses and tradable float rate and number of ordinary share in issue in both the corporate governance variables. The forecast accuracy is 0.853. In Table 4, the forecast accuracy is 0.853.

Table 2. ANN Group Statistics: A shares Listed Companies

Type I error: 0.1237					Type II error: 0.0507
Forecast accuracy: 0.8257					Test samples: 10286
Variables	Audit(1/0)	Sample	Mean	Std. Deviation	Std. Error Mean
x2	1	15225	0.74	1.38	0.01
	0	36206	0.82	1.40	0.01
x3	1	15225	54.25	241.67	1.96
	0	36206	74.36	289.19	1.52
x8	1	15225	-23.93	1051.31	8.52
	0	36206	-20.69	922.37	4.85
x12	1	15225	18.00	484.79	3.93
	0	36206	3.61	85.91	0.45
x14	1	15225	-1.37	1859.44	15.07
	0	36206	14.88	2111.53	11.10
x15	1	15225	-49.04	1548.05	12.55
	0	36206	63.00	4013.09	21.09
x19	1	15225	112.58	186.92	1.51
	0	36206	104.90	211.90	1.11
x22	1	15225	157.11	11163.52	90.47
	0	36206	3.45	462.09	2.43
x23	1	15225	1148.86	31924.45	258.73
	0	36206	474.94	13812.49	72.59
x24	1	15225	792.78	25519.31	206.82
	0	36206	324.55	14931.65	78.47
x27	1	15225	254.52	8177.49	66.27
	0	36206	48.11	4507.86	23.69
x30	1	15225	5336.02	122249.98	990.76
	0	36206	1084.47	86380.28	453.97
x35	1	15225	4133.08	57986.39	469.95
	0	36206	805.58	18848.32	99.06
x36	1	15225	84.14	3311.79	26.84
	0	36206	12.92	1011.24	5.32
x37	1	15225	225.35	11811.89	95.73
	0	36206	35.50	1822.47	9.58
x38	1	15225	135.98	3170.02	25.69
	0	36206	31.96	1716.89	9.02

Table 3. ANN Group Statistics: B shares Listed Companies

Type I error: 0.1596 | Type II error: 0.0437
Forecast accuracy: 0.7967 | Test samples: 12000

Variables	Audit(1/0)	Sample	Mean	Std. Deviation	Std. Error Mean
x1	1	14928	124.68	92.50	0.76
	0	45072	127.97	114.82	0.54
x3	1	14928	395.82	2945.97	24.11
	0	45072	529.92	5342.98	25.17
x6	1	14928	2.55	2.33	0.02
	0	45072	2.55	2.37	0.01
x7	1	14928	-35.95	2311.45	18.92
	0	45072	6046.78	312484.67	1471.89
x9	1	14928	-12.21	342.33	2.80
	0	45072	-23.75	881.23	4.15
x21	1	14928	96.37	460.46	3.77
	0	45072	99.27	465.03	2.19
x23	1	14928	4322.27	71382.63	584.24
	0	45072	921.93	30932.95	145.70
x24	1	14928	2390.57	45119.14	369.28
	0	45072	307.78	5850.25	27.56
x27	1	14928	1023.96	16252.53	133.02
	0	45072	11.25	325.11	1.53
x28	1	14928	14.66	336.52	2.75
	0	45072	0.00	0.00	0.00
x29	1	14928	7425.50	48510.47	397.04
	0	45072	516.51	9613.54	45.28
x41	1	14928	58.42	643.55	5.27
	0	45072	99.44	762.91	3.59
x42	1	14928	0.71	0.95	0.01
	0	45072	0.33	1.62	0.01
x43	1	14928	321.06	7371.09	60.33
	0	45072	141.23	3692.62	17.39
x44	1	14928	6.12	16.43	0.13
	0	45072	3.22	9.51	0.05
x45	1	14928	2.04	4.40	0.04
	0	45072	3.17	74.00	0.35
x46	1	14928	8.64	14.81	0.12
	0	45072	8.45	16.98	0.08
x48	1	14928	1.45	2.38	0.02
	0	45072	1.32	2.05	0.01
x49	1	14928	51.48	284.28	2.33
	0	45072	48.18	216.68	1.02
x51	1	14928	2.26	7.90	0.07
	0	45072	2.63	8.86	0.04

Table 4. ANN Group Statistics: H shares Listed Companies

Type I error: 0.1167					Type II error: 0.0298
Forecast accuracy: 0.8535					Test samples: 7550
Variables	Audit(1/0)	Sample	Mean	Std. Deviation	Std. Error Mean
x1	1	8205	83.24	59.36	0.66
	0	29545	64.28	61.00	0.36
x2	1	8205	142190.03	940620.96	10384.26
	0	29545	12114.69	239766.09	1394.91
x3	1	8205	8010.68	126429.62	1395.76
	0	29545	2174.66	62524.05	363.75
x6	1	8205	4.32	2.94	0.03
	0	29545	3.40	3.15	0.02
x7	1	8205	20.87	67.37	0.74
	0	29545	13.07	47.98	0.28
x8	1	8205	10.40	14.90	0.16
	0	29545	5.75	92.80	0.54
x9	1	8205	9.34	14.20	0.16
	0	29545	5.36	77.00	0.45
x40	1	8205	83.94	145.10	1.60
	0	29545	112.67	1362.79	7.93
x42	1	8205	0.76	0.64	0.01
	0	29545	0.21	0.45	0.00
x43	1	8205	19.04	38.90	0.43
	0	29545	7.33	17.75	0.10
x45	1	8205	1.86	2.75	0.03
	0	29545	0.76	1.92	0.01
x46	1	8205	9.01	10.26	0.11
	0	29545	6.60	8.46	0.05
x48	1	8205	1.47	1.69	0.02
	0	29545	1.73	2.59	0.02
x49	1	8205	27.40	60.43	0.67
	0	29545	35.37	110.98	0.65
x50	1	8205	2.80	3.78	0.04
	0	29545	3.86	19.66	0.11

Next, we focus on building financial crisis detecting model of delisted companies in A, B and H shares, separately. In Table 5 to Table 7, we use artificial intelligence method again to filter out what variables that significantly play important role in the group of delisted companies. For example, as it can be seen in Table 6, only the variables x2, x3, x41, x42, x48 are significantly different and will be put in to the predicting model in the next step.

Table 5. ANN Group Statistics: A shares Delisted Companies

Variables	Audit(1/0)	Sample	Mean	Std. Deviation	Std. Error Mean
Type I error: 0.1592					Type II error: 0.0625
Forecast accuracy: 0.7783					Test samples: 2639
x1	1	3920	86.87	189.10	3.02
	0	9279	88.10	148.70	1.54
x2	1	3920	-1.35	26.19	0.42
	0	9279	1.46	8.12	0.08
x3	1	3920	2.25	33.67	0.54
	0	9279	2.46	24.84	0.26
x5	1	3920	11.04	106.14	1.70
	0	9279	18.03	162.60	1.69
x41	1	3920	62.85	106.28	1.70
	0	9279	62.84	234.99	2.44
x42	1	3920	0.68	0.75	0.01
	0	9279	0.34	0.60	0.01
x45	1	3920	1.92	3.03	0.05
	0	9279	1.17	1.57	0.02
x46	1	3920	8.26	10.70	0.17
	0	9279	19.48	266.29	2.76
x48	1	3920	0.92	2.27	0.04
	0	9279	0.88	1.87	0.02

Table 6. ANN Group Statistics: B shares Delisted Companies

Type I error: 0.1705					Type II error: 0.0000
Forecast accuracy: 0.8295					Test samples: 440
Variables	Audit(1/0)	Sample	Mean	Std. Deviation	Std. Error Mean
x1	1	600	76.1	45.88	1.87
	0	1604	73.64	44.44	1.11
x2	1	600	-15.78	70.13	2.86
	0	1604	1.86	7.14	0.18
x3	1	600	10.3	9.01	0.37
	0	1604	-30.81	321.39	8.02
x5	1	600	5.71	17.79	0.73
	0	1604	8.15	51.27	1.28
x41	1	600	38.76	40.73	1.66
	0	1604	47.48	47.77	1.19
x42	1	600	0.78	0.65	0.03
	0	1604	0.22	0.41	0.01
x45	1	600	1.37	1.1	0.05
	0	1604	1.51	11.05	0.28
x46	1	600	7.64	5.51	0.23
	0	1604	7.85	6.46	0.16
x48	1	600	2.53	2.93	0.12
	0	1604	2.18	2.26	0.06

Table 7. ANN Group Statistics: H shares Delisted Companies

Type I error: 0.0573					Type II error: 0.2328
Forecast accuracy: 0.7099					Test samples: 262
Variables	Audit(1/0)	Sample	Mean	Std. Deviation	Std. Error Mean
x1	1	368	94.44	43.36	2.26
	0	961	88.84	45.29	1.46
x2	1	368	19369.61	50879.34	2652.27
	0	961	374.39	2085.92	67.29
x3	1	368	203.83	63.71	3.32
	0	961	205.27	69.74	2.25
x6	1	368	5.71	2.20	0.12
	0	961	6.01	2.53	0.08
x7	1	368	19.58	22.16	1.16
	0	961	17.58	21.26	0.69
x8	1	368	10.29	5.65	0.29
	0	961	9.87	6.64	0.21
x9	1	368	9.26	7.14	0.37
	0	961	9.25	6.90	0.22
x41	1	368	66.36	28.37	1.48
	0	961	76.31	40.12	1.29
x45	1	368	2.21	1.50	0.08
	0	961	1.28	1.05	0.03
x46	1	368	7.47	1.91	0.10
	0	961	7.13	2.79	0.09
x47	1	368	2213.71	6273.94	327.05
	0	961	2676.39	7539.21	243.20
x48	1	368	2.37	2.88	0.15
	0	961	2.26	2.14	0.07
x49	1	368	11.14	6.84	0.36
	0	961	11.55	7.09	0.23
x50	1	368	1.20	0.86	0.05
	0	961	0.92	0.72	0.02

Analysis of Variance

In order to build financial crisis model, single financial indicator constructing is not enough. Based on of the results of ANN technique in Table 2 to Table 7, we further test the robustness of prediction variables using analysis of variance. Table 8 is relatively average independent samples T-test inspection for listed company using Table 2 to Table 4, and Table 9 is for delisted company using Table 5 to Table 7. Both Table 8 and Table 9 show that these key variables affect significantly based on the pretty small p-value.

Table 8. Analysis of Variance: Listed Companies

Market	Variables	Levene's Test for Equality of Variances F	Sig.	t-test for Equality of Means t	df	Sig.(2-tailed)	Mean Difference	Std. Error Difference
A	x3	51.686	0	-7.541	51429	0	-20.101	2.666
	x12	35.083	0	5.451		0	14.396	2.641
	x23	44.354	0	3.341		0.001	673.922	201.692
	x24	26.063	0	2.592		0.01	468.226	180.639
	x27	52.686	0	3.659		0	206.410	56.406
	x30	65.061	0	4.474		0	4251.548	950.189
	x35	355.417	0	9.761		0	3327.494	340.882
	x36	83.848	0	3.702		0	71.222	19.238
	x37	34.372	0	2.975		0.003	189.848	63.809
	x38	92.961	0	4.792		0	104.026	21.706
B	x1	12.138	0	-3.171	59998	0.002	-3.285	1.036
	x3	29.110	0	-2.923		0.003	-134.104	45.879
	x7	21.727	0	-2.378		0.017	-6082.729	2557.610
	x23	219.437	0	8.079		0	3400.341	420.888
	x24	346.042	0	9.561		0	2082.793	217.848
	x27	688.107	0	13.221		0	1012.712	76.599
	x28	343.157	0	9.246		0	14.655	1.585
	x29	2733.438	0	28.589		0	6908.986	241.662
	x41	4.871	0.027	-5.910		0	-41.025	6.941
	x42	40.874	0	26.789		0	0.376	0.014
	x43	53.176	0	3.907		0	179.830	46.031
	x44	882.797	0	26.341		0	2.891	0.110
	x48	186.590	0	6.235		0	0.126	0.020
	x51	67.366	0	-4.595		0	-0.375	0.081
H	x1	59.391	0	25.054	37748	0	18.961	0.757
	x2	1413.979	0	21.398		0	130075.348	6078.729
	x3	119.008	0	5.786		0	5836.020	1008.683
	x6	73.050	0	23.747		0	0.921	0.039
	x42	2210.376	0	88.912		0	0.550	0.006
	x43	1588.111	0	39.121		0	11.710	0.299
	x45	321.480	0	41.260		0	1.096	0.027
	x46	56.318	0	21.776		0	2.413	0.111
	x48	41.007	0	-8.521		0	-0.258	0.030
	x49	81.032	0	-6.255		0	-7.973	1.275
	x50	33.803	0	-4.836		0	-1.055	0.218

Table 9. Analysis of Variance: Delisted Companies

Market	Variables	Levene's Test for Equality of Variances		t-test for Equality of Means				
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
A	x2	1683.218	0	-9.32	13197	0	-2.807	0.301
	x5	22.791	0	-2.48		0.013	-6.997	2.821
	x42	207.465	0	26.997		0	0.333	0.012
	x45	532.172	0	18.589		0	0.748	0.040
	x46	21.262	0	-2.637		0.008	-11.218	4.255
B	x2	589.922	0	-9.937	2202	0	-17.632	1.774
	x3	25.705	0	3.132		0.002	41.102	13.125
	x41	32.987	0	-3.964		0	-8.719	2.200
	x42	229.436	0	24.122		0	0.566	0.023
	x48	50.888	0	3.001		0.003	0.353	0.118
H	x2	426.163	0	11.555	1327	0	18995.223	1643.874
	x6	12.39	0	-1.974		0.049	-0.295	0.150
	x41	161.788	0	-4.357		0	-9.947	2.283
	x45	101.739	0	12.782		0	0.934	0.073
	x46	7.696	0.006	2.134		0.033	0.337	0.158
	x50	13.291	0	6.053		0	0.281	0.046

Binary Choice Model

In order to explore whether audit quality can improve business performance, we define Grade equals to one if the audit quality is good and also the firm's operating performance has been improved, and zero otherwise.

The following equation (3) to equation (5) indicate that which financial variables are the key variables to improve the firm's operating performance in the A-shares, B-shares, and H-shares, separately. The t-values are all significant meaning that these explanatory variables will reduce the probability of the occurrence of the financial crisis.

$$\text{Grade(A)}_i = 2.031 + 1.150\ Audit_i + \beta_2 X3_i + \beta_3 X12_i + \beta_4 X22_i + \beta_5 X23_i + \beta_6 X24_i + \beta_7 X27_i + 0.002\ X30_i + \beta_9 X33_i + \varepsilon_i \quad (3)$$

$$\text{Grade(B)}_i = 0.213 + 0.427\ Audit_i + \beta_2 X3_i + \beta_3 X7_i + \beta_4 X23_i + \beta_5 X27_i + \beta_6 X29_i + \beta_7 X41_i - 0.021\ X42_i + \beta_9 X43_i + \beta_{10} X44_i + \beta_{11} X48_i + \beta_{12} X51_i + \varepsilon_i \quad (4)$$

$$\text{Grade(H)}_i = -6.583 + 8.21\ Audit_i + \beta_2 X1_i + 0.11\ X2_i + \beta_4 X3_i + 58.88\ X6_i + \beta_6 X42_i + 1.11 X43_i + 0.13\ X45_i + \beta_9 X46_i + \beta_{10} X48_i + \beta_{11} X49_i + \beta_{12} X50_i + \varepsilon_i \quad (5)$$

Grade(A), Grade(B), and Grade(H) are all binary logic model. We will only take A-share listed company as an example to show the binary logit regression results in the Table 10. By a 9 iterations to achieve convergence, the result again shows that the audit quality is significantly correlated with the occurrence of the financial crisis.

Table 10. Logistic model

Dependent Variable: GRADE-A
Method: ML - Binary Logit (Quadratic hill climbing)
Sample: 51,431
Convergence achieved after 9 iterations

Variable	Coefficient	Std. Error	z-Statistic	Prob.
AUDIT	1.14985	0.045114	25.48765	0
X3	0.000931	0.000918	1.014653	0.3103
X12	2.95E-05	4.14E-05	0.713614	0.4755
X22	-4.67E-06	1.34E-05	-0.347718	0.7281
X23	5.28E-07	1.25E-06	0.422857	0.6724
X24	8.60E-06	5.45E-06	1.577479	0.1147
X27	5.43E-05	4.40E-05	1.234536	0.217
X30	0.001909	0.003552	0.5374	0.591
X33	1.39E-05	1.91E-05	0.728201	0.4665
C	2.031304	0.016618	122.232	0

McFadden R-squared	0.027411	Mean dependent var	0.907099
S.D. dependent var	0.290297	S.E. of regression	0.288196
Akaike info criterion	0.601836	Sum squared resid	4270.878
Schwarz criterion	0.603557	Log likelihood	-15466.5
Hannan-Quinn criter.	0.602374	Deviance	30933.04
Restr. deviance	31804.85	Restr. log likelihood	-15902.4
LR statistic	871.8031	Avg. log likelihood	-0.30072
Prob(LR statistic)	0		
Obs with Dep=0	4778	Total obs	51431
Obs with Dep=1	46653		

Table 11. Expectation-Prediction Evaluation for Binary Specification

Panel A	Estimated Equation					
	Dep=0			Dep=1		
	A	B	H	A	B	H
E(# of Dep=0)	507.05	573.71	6929.07	4270.95	2629.37	3.93
E(# of Dep=1)	4270.95	2716.29	3.93	42382.05	54080.63	30813.07
Total	4778	3290	6933	46653	56710	30817
Correct	507.05	573.71	6929.07	42382.05	54080.63	30813.07
% Correct	10.61	17.44	99.94	90.85	95.36	99.99
% Incorrect	89.39	82.56	0.06	9.15	4.64	0.01
Total Gain*	1.32	11.95	81.58	0.14	0.85	18.35
Percent Gain**	1.46	12.65	99.93	1.46	15.44	99.93

Panel B	Constant Probability					
	Dep=0			Dep=1		
	A	B	H	A	B	H
E(# of Dep=0)	443.88	180.4	1273.28	4334.12	3109.6	5659.72
E(# of Dep=1)	4334.12	3109.6	5659.72	42318.88	53600.4	25157.28
Total	4778	3290	6933	46653	56710	30817
Correct	443.88	180.4	1273.28	42318.88	53600.4	25157.28
% Correct	9.29	5.48	18.37	90.71	94.52	81.63
% Incorrect	90.71	94.52	81.63	9.29	5.48	18.37

*Change in "% Correct" from default (constant probability) specification

**Percent of incorrect (default) prediction corrected by equation

Table 12. Goodness-of-Fit Test

Goodness-of-Fit Evaluation for Binary Specification
Andrews and Hosmer-Lemeshow Tests
Grouping based upon predicted risk (randomize ties)

	Quantile of Risk		Dep=0		Dep=1		Total	H-L
	Low	High	Actual	Expect	Actual	Expect	Obs	Value
1	0	0.8915	1721	1030.53	4279	4969.47		558.56
2	0.8915	0.9103	542	596.557	5458	5403.44		5.54
3	0.9103	0.9282	298	482.935	5702	5517.06		77.02
4	0.9282	0.9457	215	375.843	5785	5624.16		73.43
5	0.9457	0.9603	108	278.757	5892	5721.24	6000	109.70
6	0.9603	0.9726	109	199.534	5891	5800.47		42.49
7	0.9726	0.9831	71	131.805	5929	5868.19		28.68
8	0.9831	0.9913	107	76.4147	5893	5923.59		12.40
9	0.9913	0.9983	101	28.5739	5899	5971.43		184.46
10	0.9983	1	18	2.13625	5982	5997.86		117.85
		Total	3290	3203.1	56710	56796.9	60000	1210.12

H-L Statistic	1210.122	Prob. Chi-Sq(8)	0
Andrews Statistic	1137.075	Prob. Chi-Sq(10)	0

In order to take a step further, we calculate Grade* which is the estimated value of Grade Logic expected value in the equation (3). As shown in equation (6) and (7), the expected value will be calculated to generate expectation-prediction table.

(6)

$$\text{Grade*} = \ln\left(\frac{P(Grade = 1|X,\beta)}{1 - P(Grade = 1|X,\beta)}\right) = 2.031 + 1.15\ \text{AUDIT}$$

(7)

$$e^{\text{Grade *}} = \left(\frac{P(Grade = 1|X,\beta)}{1 - P(Grade = 1|X,\beta)}\right)$$

The calculation of expected value is as in equation (8), if the predicted probability is greater than the cutoff value = 0.5, then it will be classified as one, otherwise it will be classified as zero.

(8)

$$e^{\text{Grade *}} = \left(\frac{P(Grade = 1|X,\beta)}{1 - P(Grade = 1|X,\beta)}\right)$$

Table 11 show the results of expectation-prediction evaluation for binary specification. We use estimated equation in the Panel A and constant probability in the Panel B. As it can be seen, the percentage of correct prediction rate are all about 90% in the column of dependent-equal-to-one.

Finally, Table 12 shows the test results of Goodness-of-Fit. As shown in this table, the sample is divided into 10 groups of observations. The quantile of risk prediction probability values are displayed up to 90%. This table again confirm our findings.

Conclusion

The ability to accurately predict financial distress for enterprise is a very important issue in financial decision-making. This study investigates whether financial crisis prediction is available. We use big data techniques including binary choice model and life test method to predict financial distress. We first collect 54 financial indicators and corporate governance variables, then through an independent t-test, and the Binary Choice Model analysis techniques such as artificial intelligence methods to reduce variable.

The empirical results confirm that the selected variables perform well in predicting the financial crisis. This means that these variables have a significant effect on predict financial distress. Using artificial intelligence techniques to investigate the probability of survival analysis, we find audit quality will significantly reduce the occurrence of financial crises. The evidence can provide a more effective decision-making reference for business and financial data management, supporting to help enterprises industry to adjust business strategy and defuse financial risks.

References

Almeida, H., & T. Philippon, 2008, Estimating risk-adjusted costs of financial distress, Applied Corporate Finance, 20(4), 105-109.

Beaver, W. H., 1966, Financial ratios as predictors of failure, Journal of Accounting Research, 4(3), 71 -111.

Cerutti, E., S. Claessens & P. McGuire, 2012, Systemic risks in global banking: What available data can tell us and what more data are needed? NBER Working Paper No. 18531.

Claessens S. & M.A. Kose, 2013, Financial crises: explanations, types, and implications, International Monetary Fund, WP/13/28.

Han, J., & M. Kamber, 2000, Data mining: concepts and techniques, Morgan Kaufmann.

Kluger, B.D., & D. Shields, 1989, Auditor changes, information quality and bankruptcy prediction, Managerial And Decision Economics, 10(4), 275-282.

Lee, K. D. & P.A. Booth, 2005, A comparison of supervised and unsupervised neural networks in predicting bankruptcy of Korean firms, Expert Systems with Applications, 29, 1-16.

Lensberg, T., A. Eilifsen, & T.E. McKee, 2006, Bankruptcy theory development and classification via genetic programming, European Journal of Operational Research, 169, 677-697.

Li, B. X., 2005, Research on forecasting firms financial distress based on nonlinear combining of fuzzy neural network, Journal of Industrial Engineering and Engineering Management, 19(1), 19-23.

Liu, C., K. Uchida, & Y. Yang, 2012, Corporate governance and firm value during the global financial crisis: Evidence from China, International Review of Financial Analysis, 21, 70-80.

Lu, Y., & D. Ma, 2016, Audit quality and financial distress: Evidence from China, WSEAS Transactions on Business and Economics, 13, 330-340

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

Pastena, V. & W. Ruland, 1986, The merger/bankruptcy alterative, The Accounting Review, 61(2), 288-301.

Simoudis, E., 1996, Reality check for data mining, IEEE Intelligent Systems, 11, 26-23.

Sundgren, S., 2009, Perceived audit quality, modified audit opinions and the likelihood of liquidating bankruptcy among financially weak firms, International Journal of Auditing, 13(3), 203-221

Tae, K. S., C. Namsik, & L. Gunhee, 1999, Dynamics of modeling in data mining: interpretive approach to bankruptcy prediction, Journal of Management Information Systems, 16(1), 63-85.

Tsai, C.F., 2011, Combining cluster analysis with classifier ensembles to predict financial distress, Information Fusion, 16, 46-58.