Credit Risk Assessment Model of Small and Medium-Sized Enterprise Based on Logistic Regression

Hui Sun, Mingyuan Guo
College of Management and Economics, Tianjin University, Tianjin, China
(leu2@163.com)

Abstract - **Small and medium-sized enterprises play a very important role in China's economic and social development. Their development is inseparable from the financial support, and the development of credit business is also an important aspect of bank's operation. At the time of solving the problem of financing difficulties of small and medium-sized enterprises, how to manage credit risk will become an important subject to study. In this paper, we choose 50 money-borrowing enterprises' financial data in a commercial bank as samples to construct logistic risk assessment model. We make regression analysis, Hosmer-Lemeshow test and Hosmer-Lemeshow randomness test. The empirical results show that the logistic risk assessment model has good fitness of the data.**

Keywords - **Commercial banks, small and medium-sized enterprises, credit risk, logistic model**

I. INTRODUCTION

In more than 30 years of reform and opening-up, China's economy maintains rapid growth. It has made great achievements and also spawned a large number of small and medium-sized enterprises. In today's market economy, small and medium-sized enterprises have become an important part of national economic development. They play a very important role in the economic and social development. As the main indirect financing way of small and medium-sized enterprises (hereinafter referred to as "SME"), commercial banks increase support for SME financing. However, due to unstable operation, imperfect financial system and shallow credit consciousness, compared with large enterprises, SME will produce a greater credit risk. It has also become the current focus of banking and financial experts.

In recent years, many scholars in China have studied on risk management of commercial banks. Zhang Jie (2012) believes that loan risk is basically a kind of credit risk [1]. Tong Ming-jun (2012) argue that credit risk should also include management risk and market risk [2]. Wang Ying-chi, Xue Jing (2013) argue that in addition to operating risk, market risk, credit risk, small and medium-sized enterprise risk should also include policy risk [3]. Duan Shi-jie (2012) constructs a detailed analysis of the influencing factors of the commercial Banks' credit risk of small and medium-sized enterprise, and put forward three measures to improve commercial bank's credit risk management [4]. Xu Qing (2010) points out that commercial banks should build a risk management control system specifically for small and medium-sized enterprise [5]. Sun Ji, Liu Shi-yun(2012) put forward three suggestions which include a clear direction of the development of small and medium-sized enterprise credit policy, reforming credit system and setting up service system for small and medium-sized enterprise [6]. Liu Yu-xin, Wei Can-qiu (2003) set up the analysis method to forecast commercial bank loan default probability [7]. In addition, Yang Peng-bo, Zhang Cheng-hu, Zhang Xiang (2009) use logistic regression model to establish a credit default probability forecasting model for listed company [8].

II. LOGISTIC REGRESSION MODEL AND ANALYSIS

Logistic regression model mainly makes probability forecast for the event influenced by various factors. It is an extension of the general linear regression model, which is divided into binary logistic regression and multiple logistic regression. Martin (1977) first uses logistic regression method to study operation risk of the bank [9]. Wu Shi-nong (2001) uses 70 special treatment and normal companies as the research object. The research results show that logistic model has the lowest misjudgment rate compared with other forecasting model. [10]. In this paper, the default risk of the money-borrowing enterprises only has two situations, which includes the default and non-default. Therefore, the binary logistic regression model is suitable for analysis.

The model is suitable for forecasting commercial bank's credit risk in China. The reasons are as follows: Firstly, binary logistic regression model has been verified and has a good forecast ability; Secondly, this model can fit the data which don't follow normal distribution. We don't need to make any hypothesis for the prior probability of credit defaults and the distribution of the sample data.

Logistic regression model is as follows:

$$P = \frac{1}{1+e^{-(\beta_0+\beta_1F_1+\beta_2F_2+\beta_3F_3+\beta_4F_4+\beta_5F_5)}} = \frac{1}{1+e^{-(\beta_0+\sum_{i=1}\beta_iF_i)}} \quad (1)$$

Where, P is the default probability. F_i is independent variables. β_0 is the regression intercept. β_i is the regression coefficient.

III. EMPIRICAL ANALYSIS

A. Data Source

We select 50 samples from a commercial bank in 2012 including 30 non-default companies and 20 default enterprises. Based on the evaluation system of large commercial banks, we choose 12 financial indicators as explanatory variables of money-borrowing companies to measure the degree of prosperity, the liquidity, the profitability and the growth of the enterprise.

Asset-liability ratio, current ratio and cash flow liability ratio can reflect short-term debt paying ability. Net profit margin, return on equity and sales gross profit margin can reflect corporate profitability. Total asset turnover, inventory turnover and accounts receivable turnover ratio can reflect the money-borrowing companies' operation ability. Sales growth rate, net profit growth rate and capital increment rate can reflect the development ability.

Using logistic regression, we need to consider the collinearity problem of various indicators. Because the linear relationship of independent variable in the regression function will cause the multi-collinearity problem and reduce the explanatory power of the model.

B. Factor Analysis

We do factor analysis in this section. Principal component analysis (PCA) in factor analysis is used to extract principal components and factors whose characteristic value greater than 1. According to our results, it is appropriate to extract five factors.

TABLE I
FINANCIAL INDICATORS

Basic indicators	Symbol	Indicators instructions
Asset-liability ratio	X_1	(total debt/total assets)×100%
Current ratio	X_2	current assets/current liabilities
Cash flow liability ratio	X_3	(net cash flows from operating activities/ interest-bearing debt due within one year)×100%
Net profit margin	X_4	(net profit/main business profit)×100%
Return on equity	X_5	(net income/average Stockholders' equity)×100%
Sales gross profit margin	X_6	(sales net income-the product cost)/sales net income×100%
Total asset turnover	X_7	sales net income/average total assets
Inventory turnover	X_8	main business cost/average inventory
Accounts receivable turnover	X_9	sales net income/average accounts receivable
Sales growth rate	X_{10}	(this year's sales growth/net sales revenue last year)×100%
Net profit growth rate	X_{11}	(the final net profit-initial net income)/initial net profit×100%
Capital increment rate	X_{12}	(terminal owner's equity - initial owner's equity)/initial owners' equity×100%

TABLE Ⅱ
FACTOR LOADING DIAGRAM

	Factor				
	1	2	3	4	5
Asset-liability ratio	-0.915	-0.030	-0.066	0.005	-0.143
Current ratio	0.904	-0.045	0.004	-0.002	-0.068
Sales gross profit margin	0.523	-0.056	0.256	-0.467	0.466
Net profit growth rate	-0.045	0.941	0.008	0.025	0.076
Capital increment rate	-0.008	0.917	-0.103	-0.145	0.042
Cash flow liability ratio	0.140	0.600	0.537	0.047	0.248
Inventory turnover	-0.033	-0.025	-0.884	0.079	0.207
Sales growth rate	-0.496	-0.299	0.545	0.212	0.233
Sales net profit margin	0.440	-0.041	0.534	-0.367	0.265
Accounts receivable turnover	0.174	-0.111	0.153	0.804	-0.085
Total asset turnover	-0.265	-0.001	-0.240	0.767	0.065
Return on equity	0.025	0.185	-0.074	-0.023	0.908

TABLE Ⅲ
FACTOR SCORE COEFFICIENT MATRIX

	Factor				
	1	2	3	4	5
Asset-liability ratio	-0.397	0.000	0.025	-0.137	-0.044
Current ratio	0.412	-0.012	-0.044	0.112	-0.119
Cash flow liability ratio	0.022	0.263	0.296	0.134	0.090
Sales net profit margin	0.099	-0.056	0.238	-0.132	0.118
Return on equity	-0.054	-0.018	-0.120	0.071	0.747
Sales gross profit margin	0.124	-0.093	0.048	-0.190	0.297
Total asset turnover	-0.008	0.019	-0.077	0.475	0.163
Inventory turnover	0.031	-0.044	-0.526	0.006	0.261
Accounts receivable turnover	0.184	-0.008	0.141	0.569	-0.013
Sales growth rate	-0.243	-0.151	0.336	0.124	0.239
Net profit growth rate	-0.021	0.433	0.015	0.056	-0.037
Capital increment rate	-0.019	0.418	-0.061	-0.066	-0.071

In factor F_1, variables with larger loading include asset-liability ratio and liquidity ratio which reflect the short time debt paying ability of money-borrowing companies. We defined them as short time debt paying ability factor. In factor F_2, variables with larger loading include net profit growth rate, capital increment rate, operating cash flow ratio and sales gross profit rate which reflect future growth expectation of enterprises. We defined them as future growth ability factor. In factor F_3, variables with larger loading include inventory turnover, sales revenue growth rate and sales net profit margin. In factor F_4, variables with larger loading include accounts receivable turnover and total asset turnover, which reflect assets operation ability of enterprises. We defined them as operating capacity factor. In factor F_5, variables with larger loading include return on equity, which reflects asset returns of the money-borrowing companies. We defined them as profit ability factor. Regression method is used to obtain factor coefficient matrix and score function of five factors. It is shown as follows:

$$F_1 = -0.397x_1 + 0.412x_2 + 0.022x_3 + 0.099x_4 - 0.054x_5 + 0.124x_6 - 0.008x_7 + 0.031x_8 + 0.184x_9 - 0.243x_{10} - 0.021x_{11} - 0.019x_{12} \quad (2)$$

$$F_2 = -0.012\ x_2 + 0.263\ x_3 - 0.056\ x_4 - 0.018\ x_5 - 0.093\ x_6 + 0.019\ x_7 - 0.044\ x_8 - 0.008\ x_9 - 0.151\ x_{10} + 0.433\ x_{11} + 0.418\ x_{12} \quad (3)$$

$$F_3 = 0.025x_1 - 0.044x_2 + 0.296x_3 + 0.238x_4 - 0.120x_5 + 0.048x_6 - 0.077x_7 - 0.526x_8 + 0.141x_9 + 0.336x_{10} + 0.015x_{11} - 0.061x_{12} \quad (4)$$

$$F_4 = 0.137x_1 + 0.112x_2 + 0.134x_3 - 0.132x_4 + 0.071x_5 - 0.190x_6 + 0.475x_7 + 0.006x_8 + 0.569x_9 + 0.124x_{10} + 0.056x_{11} - 0.066x_{12} \quad (5)$$

$$F_5 = -0.044\,x_1 - 0.119\,x_2 + 0.134\,x_3 - 0.132\,x_4 + 0.071\,x_5 - 0.190\,x_6 + 0.475\,x_7 + 0.006\,x_8 + 0.569\,x_9 + 0.124\,x_{10} + 0.056\,x_{11} - 0.066\,x_{12} \quad (6)$$

Where, x_1 represents asset-liability ratio, x_2 represents liquidity ratio, x_3 represents operating cash flow ratio, x_4 represents net profit margin, x_5 represents return on equity, x_6 represents sales gross profit margin, x_7 represents total asset turnover, x_8 represents inventory turnover, x_9 represents accounts receivable turnover, x_{10} represents sales growth rate, x_{11} represents net profit growth rate and x_{12} represents capital increment rate.

C. Logistic Regression Analysis

Regression analysis includes total 50 sample data. The score of the five factors is calculated based on the financial data of 50 money-borrowing companies of H Bank's Tangshan Branch. Test of goodness of fit of the results is performed. Table Ⅳ shows the parameter estimation result of logistic regression model.

TABLE Ⅳ
PARAMETER ESTIMATION RESULT OF LOGISTIC REGRESSION MODEL

	β	standard error	significance level
F_1	0.173	0.066	0.009
F_2	0.035	0.014	0.012
F_3	-0.001	0.047	0.044
F_4	0.076	0.092	0.039
F_5	0.411	0.153	0.007
constant	-14.269	5.336	0.007

β represents for regression coefficient. The smaller the value is, the more significant the corresponding variable is. In Table IV, significance level of all variables is less than 0.05, which means all variables are significant and logistic regression model is meaningful.

According to the regression results in Table IV, logistic regression model is obtained:

$$P = \frac{1}{1 + e^{-(\beta_0 + \beta_1 F_1 + \beta_2 F_2 + \beta_3 F_3 + \beta_4 F_4 + \beta_5 F_5)}} = \frac{1}{1 + e^{(14.269 - 0.173 F_1 - 0.035 F_2 + 0.001 F_3 - 0.076 F_4 - 0.411 F_5)}} \quad (7)$$

Hosmer-Lemeshow test is a commonly used statistical method to test the goodness of fit of logistic regression model. The test results are shown in Table V. Generally, Hosmer-Lemeshow test and Hosmer-Lemeshow randomness test are performed simultaneously. The smaller the gap between observed value and expected value in randomness test, the higher goodness of fit is. The results obtained through SPSS software are shown in Table V and Table VI.

It can be seen from Table V that chi-square statistic is 1.968. Significance =0.982 > 0.05, which shows the logistic regression model has high goodness of fit.

TABLE Ⅴ
HOSMER-LEMESHOW TEST

Number	Chi-square	Degrees of freedom	Significance level
1	1.968	8	0.982

It can be seen from Table VI that observed value is approximately equal to expected value, which shows the model fitting is ideal.

TABLE Ⅵ
HOSMER-LEMESHOW RANDOMNESS TEST

	P= 0		P= 1	
	Observed value	Expected value	Observed value	Expected value
1	5	4.993	0	0.007
2	5	4.956	0	0.044
3	5	4.671	0	0.329
4	3	3.438	2	1.562
5	1	1.492	4	3.508
6	1	0.381	4	4.619
7	0	0.059	5	4.941
8	0	0.009	5	4.991
9	0	0.000	5	5.000
10	0	0.000	5	5.000

IV. CONCLUSION

In this paper, we first select financial indicators which affect credit risk of small and medium-sized enterprises in Tangshan, China. Secondly, based on financial data and the performance of money-borrowing companies in 2012, we use logistic model to do the analysis. The main conclusions of this paper are as follows:

1. Factor analysis is used to extract principal component based on sample data. Among financial indexes of money-borrowing companies, debt paying ability index, future growth ability index, cost index, operation ability index and profitability index have great impact on credit risk.

2. Logistic regression model is performed with Hosmer-Lemeshow test and Hosmer-Lemeshow randomness test. It is found that logistic regression model has good fitness of the data.

To sum up, with increasing numbers of commercial banks, competition will become increasingly fierce. So, credit risk management is important for commercial banks to achieve great development. It requires commercial banks to establish a unified database and information management system and use modern risk management techniques for effective management.

REFERENCES

[1] Zhang Jie, "Brief discussion on strengthening the construction of state-owned commercial bank credit ability," *The world of labor security (theory version)*, vol. 24, no. 6, pp. 70-72, Jun. 2012.

[2] Tong Ming-jun, "How to strengthen the management of small and medium-sized enterprise credit risk for banks," *Management and Administration*, vol. 30, no. 2, pp. 72-73, Feb. 2012.

[3] Wang Ying-chi, Xue Jing, "A Brief Analysis of Issues Concerning Commercial Bank Management of SME Credit Risks," *Business Economy*, vol. 32, no. 4, pp. 105, Apr. 2013.

[4] Duan Shi-jie, "The influence factors of small and medium-sized enterprise credit risk and management countermeasures in commercial bank," *Chinese & Foreign Entrepreneurs*, vol. 30, no. 10, pp. 33-35, Oct. 2012.

[5] Xu Qing, "Small and medium-sized enterprise credit risk control research in China's commercial Banks," *Journal of Chongqing University of Science and Technology (Social Sciences Edition)*, vol. 12, no. 2, pp. 1-3, Sep. 2010.

[6] Sun JI, Liu Shi-yun, "Small and medium-sized enterprises credit strategy analyses in commercial Banks," *Business*, vol. 11, no. 21, pp.120, Jun. 2012.

[7] Liu Yu-xin, Wei Can-qiu, "Measuring the Credit Risk of the Credit Derivative with EDF Model," J*ournal of Sichuan University (Natural Science Edition)*, vol. 40, no. 4, pp. 1-5, Aug. 2003.

[8] Yang Peng-bo, Zhang Cheng-hu and Zhang Xiang, "A Research into Listed Companies' Credit Default Probability Forecast Model Based on Logistic Regression Analysis," *Economic Survey*, vol. 26, no. 2, pp.144-148, Jun. 2009.

[9] Martin D., "Early Warning of Bank Failure," *Journal of Banking and Finance*, vol. 1, no. 3, pp. 249-276, Nov. 1977.

[10] Wu Shi-nong, Lu Xian-yi, "A Study of Models for Predicting Financial Distress in China's Listed Companies," *Economic Research Journal*, vol. 47, no.6, pp. 46-55, Jun. 2001.

[11] Yao Jian-ping, "Risk and Profit Model of Credits to the Medium and Small-Sized Businesses," *Journal of Chongqing Jianzhu University*, vol. 23, no. 4, pp. 1-5, Aug. 2001.