ATLANTIS PRESS
Advances in Economics, Business and Management Research, volume 96
2019 3rd International Conference on Education, Management Science and Economics (ICEMSE 2019)

Evaluation Analysis of Credit Risk for Listed Real Estate Companies Based on Logistic Model

Dongpeng Huang
Foreign language Department
Beihang University
Beijing, 100191, China

***Abstract*—With the aim to measure the credit risk of real estate companies, this study took the default probability of companies as a criterion and developed the evaluation model of credit risk. The sample included forty real estate companies listed in China's stock exchanges whose data has been accessible from 2014 to 2018. we used applied Kolmogorov-Smirnov test to judge the distribution of selected fifteen indexes, then, applied independent-samples T test and Mann-Whitney U test to select indexes with significance, after that, principal component analysis was used to condense fifteen indicators into three principal component factors, finally, a Logistic regression model was built with those three principal component factors. Results show that the correctness of the prediction model is more than 85%; indexes representing profitability, operating capacity and growth capacity of the company have a significant impact on the credit risk. The Logistic model can make a more accurate judgment on the credit risk of real estate companies.**

Keywords— principal component analysis; Logistic model; credit risk; listed real estate company

I. INTRODUCTION

With the development of economic globalization, domestic real estate companies face various risks in the market competition. How to assess and control risks has become a problem that must be faced in the development of real estate companies in China.

A lot of researchers at home and abroad kept applying Logistic model to credit risk measurement. Altman[1] (1968) used multivariate discriminant analysis for the first time to predict financial distress. Sjur [2] (2001) established a logistic regression model to predict the probability of a company going bankrupt or defaulting. Shinong Wu [3] (2001) and Fuxia Guo [4] (2012) all took the Logistic model to construct relevant financial early-warning and evaluation system. J.H. Li, et al. [5](2011) used the Logistic model to evaluate and test the credit rating of listed companies, the results showed that the model has a good predictive ability. Junzi Wang, et al.[6](2017) compared the early-warning application of different types of credit risk measurement models in real estate credit risk, and showed that the Logistic model was relatively accurate. Sheng Hu, et al. [7] (2018) used principal component analysis and Logistic model to measure the credit risk of Chinese real estate listed companies. On the basis of the research of scholars at home and abroad, a Logistic model was established in this study to discuss the credit risk of real estate enterprises in China.

II. SAMPLE AND FINANCIAL INDEX SELECTION

A. Sample Selection

In this study, forty listed real estate companies from China's stock exchanges were selected as independent samples with financial data from 2014 to 2018. The selected companies were divided into two groups. Group 1 was composed of twenty companies with continuous profits in the past five years without credit risk. Group 2 was composed of twenty companies with credit risk that have suffered losses for two consecutive years in the past five years.

B. Financial Index Selection

In this study, fifteen financial indexes were selected to measure the company's credit risk from five aspects of financial condition, those are profitability index included rate of net income on total assets X_1, net profit margin on sales X_2, return on equityX_3; Operating capacity index included accounts receivable turnover ratio X_4, total assets turnover ratio X_5 and inventory turnover ratio X_6, current asset turnover ratio X_7; Solvency index included current ratio X_8, quick ratio X_9, debt asset ratio X_{10} and equity ratio X_{11}; The growth capacity index included growth rate of main operating revenue X_{12} and net asset growth rate X_{13}; Cash flow capacity index included operating cash flow to sales ratio X_{14} and return on operating cash flow of assets X_{15}.

III. LOGISTIC MODEL ESTIMATE

A. Significant Test of Index

Since the population distribution of fifteen indexes is unknown, the normality of the samples was first tested by the Kolmogorov-Smirnov test, and then, Independent sample T test was used to test the significance of discrimination for indexes subject to normal distribution, and Mann-Whitney U method was used to test the significance of discrimination for indexes not subject to normal distribution.

1) Testing of the Normality: In this study, the K-S test was used for the normality test. Let the null hypothesis H0 in K-S test was as Equation (1)

$$H_0\text{: the index Xi is normally distributed} \quad (1)$$

In case the probability is more than 0.05, H_0 will be accepted at 95% level of confident, otherwise, it will be rejected. The results of testing normality have been shown in Table I.

TABLE I. ONE-SAMPLE KOLMOGOROV-SMIRNOV TEST

	N	Normal Parameters[a,b]		Most Extreme Differences			Test Statistic	Asymp. Sig. (2-tailed)
		Mean	*Std. Deviation*	*Absolute*	*Positive*	*Negative*		
X_1	40	0.159	4.242	0.148	0.121	-0.148	0.148	.028[c]
X_2	40	-30.969	85.647	0.295	0.225	-0.295	0.295	.000[c]
X_3	40	-17.886	96.115	0.401	0.349	-0.401	0.401	.000[c]
X_4	40	425.956	1726.633	0.446	0.446	-0.403	0.446	.000[c]
X_5	40	0.254	0.154	0.133	0.133	-0.105	0.133	.073[c]
X_6	40	321.332	1875.648	0.508	0.508	-0.432	0.508	.000[c]
X_7	40	0.426	0.491	0.296	0.296	-0.231	0.296	.000[c]
X_8	40	2.042	1.114	0.221	0.221	-0.138	0.221	.000[c]
X_9	40	1.035	1.085	0.296	0.296	-0.235	0.296	.000[c]
X_{10}	40	63.273	19.282	0.135	0.069	-0.135	0.135	.064[c]
X_{11}	40	386.611	672.694	0.357	0.357	-0.290	0.357	.000[c]
X_{12}	40	1044.011	5926.224	0.473	0.473	-0.429	0.473	.000[c]
X_{13}	40	55.052	152.197	0.389	0.389	-0.338	0.389	.000[c]
X_{14}	40	-0.157	1.363	0.308	0.250	-0.308	0.308	.000[c]
X_{15}	40	0.015	0.059	0.201	0.201	-0.143	0.201	.000[c]

a. Test distribution is Normal.
b. Calculated from data.
c. Lilliefors Significance Correction.

The bilateral significant P values of K-S test for two indexes X_5 and X_{10} equals 0.073 and 0.064, respectively, both of which are more than 0.05, so the two indexes are in a normal distribution. However, the probability of the other thirteen indexes are all less than 0.05, they do not conform to the normal distribution.

2) T-Test of Two Independent Samples: The T-test process of independent samples is used to check the mean values of various indicators are significantly at different dependent variable levels. If the probability of the tested indexes is less than 0.05, it indicates that the corresponding index is significant, and vice versa. The results of T-test had been shown in Table II and Table III.

TABLE II. GROUP STATISTICS

	Group	N	Mean	Std. Deviation	Std. Error Mean
X_5	1	20	.2220700	.07424699	.01660213
	2	20	.2850950	.20213008	.04519766
X_{10}	1	20	67.3298600	14.31253167	3.20037937
	2	20	59.2165250	22.88381290	5.11697612

TABLE III. T-TEST OF INDEPENDENT SAMPLES

		Levene's Test for Equality of Variances		t-test for Equality of Means		
		F	*Sig.*	*T*	*df*	*Sig.(2-tailed)*
X_5	Equal variances assumed	12.283	0.001	-1.309	38	0.198
	Equal variances not assumed			-1.309	24.036	0.203
X_{10}	Equal variances assumed	5.951	0.019	1.344	38	0.187
	Equal variances not assumed			1.344	31.892	0.188

It has been shown that the probability of T- test for two indexes X_5 and X_{10} equals 0.203 and 0.183 respectively, both of which are more than 0.05. Therefore, there is no significant difference between the two groups of sample companies in terms of index X_5 and X_{10}. These two indexes will be eliminated in the following analysis. the second affiliation.

3) Mann-Whitney U test: Non-parametric test statistical method is mainly used for the case where the overall distribution does not consider the distribution of the object being studied and if the distribution is known, and there is no assumption about the overall distribution. Therefore, the Mann-Whitney U test method was used to test the significance of the remaining thirteen indexes that do not obey the normal distribution. The results of Mann-Whitney U test had been shown in Table IV.

TABLE IV. RESULTS OF MANN-WHITNEY U TEST

	Mann-Whitney U	Wilcoxon W	Z	Asymp. Sig. (2-tailed)	Exact Sig. [2*(1-tailed Sig.)]
X_1	35.000	245.000	-4.463	0.000	.000
X_2	24.000	234.000	-4.761	0.000	.000
X_3	26.000	236.000	-4.707	0.000	.000
X_4	142.000	352.000	-1.569	0.117	.121
X_6	103.000	313.000	-2.624	0.009	.008
X_7	156.000	366.000	-1.190	0.234	.242
X_8	174.000	384.000	-0.703	0.482	.495
X_9	128.000	338.000	-1.948	0.051	.052
X_{11}	186.000	396.000	-0.379	0.705	.718
X_{12}	55.000	265.000	-3.922	0.000	.000
X_{13}	150.000	360.000	-1.353	0.176	.183
X_{14}	190.000	400.000	-0.271	0.787	.799
X_{15}	180.000	390.000	-0.541	0.589	.602

As Table IV shows, at the significance level of 0.05, five indexes including X_1, X_2, X_3, X_6 and X_{12} have passed Mann-Whitney U test, which indicate that there is a statistically significant relationship between the five indexes in two groups of sample companies. However, the probability of Mann-Whitney U test for other eight indexes including X_4, X_7, X_8, X_9, X_{11}, X_{13}, X_{14} and X_{15} are more than 0.05, it failed the test and will be eliminated in the next selection of the independent variable to ensure the accuracy of the forecast

B. Analysis of the Factors

There is the possibility of mutual conversion between financial indicators, and such multicollinearity between financial indicators may reduce the stability of the model and even affect the final results. In this study, the principal component was extracted by factor analysis, and the information of discrete index was condensed into the index needed for prediction.

1) Testing of Kaiser-Meyer-Olkin and Bartlett's sphericity: Kaiser-Meyer-Olkin (hereinafter referred to as KMO test) and the Bartlett's sphericity test were used to test the correlation between variables.The test results had been shown in Table V.

TABLE V. KMO AND BARTLETT'S TEST

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		.716
Bartlett's Test of Sphericity	Approx. Chi-Square	53.823
	df	10
	Sig.	.000

As Table V shows, the KMO statistics equals 0.559, which indicate that variables can be used for factor analysis. The chi-square value equals 53.823, and the probability of KMO and Bartlett's test equals 0.000, which is less than 0.05, indicating that the variables are independent. In summary, the data can be further analyzed by factor analysis and principal component analysis.

2) Communalities: The commonalities is the degree of common factors that can be extracted from the original information contained in each variable. This study analyzed the common factor variance ratio of variables, the results had been shown in Table VI.

TABLE VI. COMMUNALITIES

	Initial	**Extraction**
X_1	1.000	.804
X_2	1.000	.808
X_3	1.000	.774
X_6	1.000	.996
X_{12}	1.000	.996

As Table VI shows, the value of the commonalities of all variables is almost 80%, so the three factors extracted can be used to explain the variables.

3) Rotation of the Factors: In this study, the above factors were rotated by the maximum variance method for ensuring the interpretability of the factors, and then the information after rotation was redistributed, that is to divide five variables into three Eigenvalues. The eigenvalue represents the average amount of interpretation of the original variable information after the factor is introduce.The results was shown in Table VII.

As Table VII shows, the eigenvalue of three factor is close to 1, so the first three common factors are extracted. After rotation, the cumulative variance contribution rate of the three common factors changes, and the cumulative variance contribution rate of the first three factors reaches 87.583%.Therefore, the impact of the five indexes can be explained by these three common factors.

4) Expression of the Factor: In this study, the variance maximization method in the orthogonal rotary valve was used to convert the factor. The rotated factor had a stronger explanatory power. The results have been shown in Table VIII.

TABLE VII. ROTATED PRINCIPAL COMPONENT MATRIX [A]

	Component		
	1	***2***	***3***
X_1	.891	-.099	.008
X_2	.892	.060	-.095
X_3	.878	-.024	.055
X_6	-.030	.998	.000
X_{12}	-.013	.000	.998

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization
[a.] Rotation converged in 4 iterations.

As Table VIII shows, the higher load on factor F_1 is the index X_1 (rate of net income on total assets, X_2 (net profit margin on sales) and X_3(return on equity) ;

The higher load on factor F_2 is the index X_6 (inventory turnover ratio);

The higher load on factor F_3 is the index X_{12} (growth rate of main operating revenue);

Combined with these high-value indicators on factors F_1, F_2 and F_3, a scoring coefficient matrix of principal components is established to calculate the values of principal components. The results have been shown in Table IX.

TABLE VIII. COMPONENT SCORE COEFFICIENT MATRIX

	Component		
	1	***2***	***3***
X_1	.375	-.066	.023
X_2	.380	.091	-.078
X_3	.373	.008	.071
X_6	.024	.990	.008
X_{12}	.012	.008	.990

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

The factor expressions for each major component can be obtained from Table VIII as follows:

$$F_1 = 0.375X_1 + 0.380X_2 + 3.73X_3 + 0.24X_6 + 0.12X_{12} \quad (2)$$

$$F_2 = -0.66X_1 + 0..91X_2 + 0.08X_3 + 0.990X_6 + 0.08X_{12} \quad (3)$$

$$F_3 = 0.23X_1 - 0.78X_2 + 0.71X_3 + 0.08X_6 + 0.990X_{12} \quad (4)$$

As the e equation show, the factor F_1 includes the index X_1, X_2 and X_3, which mainly represent the profitability of a company, F_2 includes the index X_6, which mainly represents the company's operation ability, and F_3 includes the index X_{12}, which mainly represents the growth ability of a company.

TABLE IX. TOTAL VARIANCE EXPLAINED

Component	Initial Eigenvalues			Rotation Sums of Squared Loadings		
	Total	*% of Variance*	*Cumulative %*	*Total*	*% of Variance*	*Cumulative %*
1	2.368	47.362	47.362	2.361	47.226	47.226
2	1.015	20.307	67.669	1.010	20.190	67.417
3	0.996	19.915	87.583	1.008	20.167	87.583
4	0.338	6.761	94.344			
5	0.283	5.656	100.000			

C. Logistic Model Estimate

The prediction of credit risk is a dichotomous qualitative analysis. In this study, the three principal component factors extracted above were taken as covariates, and the probability of corporate default was used as dependent variables to establish a binary Logistic regression model. The model took the default probability equal to 0.5 as the critical value. When the probability of the model is more than 0.5, the company will fall into a credit crisis; otherwise, when the probability of the model is less than 0.5, the company has no default risk. The results of parameter estimation of the three factors have been shown in Table X, Table XI and Table XII.

TABLE X. VARIABLES IN THE EQUATION

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	*F1*	-5.477	1.833	8.932	1	.003	.004
	Const	1.371	.763	3.227	1	.072	3.940

a. Variable(s) entered on step 1: F_1

TABLE XI. VARIABLES NOT IN THE EQUATION

			Score	df	Sig.
Step 1[b]	*Variables*	F_2	.309	1	.578
		F_3	.397	1	.529
	Overall Statistics		.763	2	.683

b. Variable(s) entered on step 1: F_2, F_3

TABLE XII. OMNIBUS TESTS OF MODEL COEFFICIENTS

		Chi-square	df	Sig.
Step 1	*Step*	26.054	1	.000
	Block	26.054	1	.000
	Model	26.054	1	.000

As above Table shows, the significance probability of the model test is less than 0.05, probability of the estimated principal component factor F_1equals 0.003, which is less than 0.05, so factor F_1 is statistically significant, while factor F_2 and F_3 are both meaningless and been eliminated from the regression model. It indicates that only factor F_1 has an impact on company's credit risk. Moreover, the regression coefficient is less than 0, indicating that the weaker the profitability, the higher the probability of credit risk. And then the Logistic regression prediction model is as follows:

$$P = \frac{1}{1 + \mathrm{Exp}[-(1.371 - 5.477F_1)]} \tag{5}$$

The chi-square value of the model test equals 26.054; the significance probability of the model test is less than 0.05, it indicate that the model is fit well.

According to above model, the company's credit risk is predicted. The sample data were brought into the model to predict the credit risk of the real estate industry in China. The results had been shown in Table XIII.

TABLE XIII. CLASSIFICATION TABLE [A]

	Observed		Predicted		
			Default		Percentage Correct
			1	2	
Step 1	*Default*	1	18	2	90.0
		2	4	16	80.0
	Overall Percentage				85.0

a. The cut value is 0.500

As Table XIII shows, there are eighteen companies without credit risk in group 1, and the number of misjudgments is two, so the prediction accuracy of the model reaches 90%; while in the group 2, there are sixteen companies with credit risk, and the number of misjudgments is four, the prediction accuracy of the model reaches 80%. On the whole, the prediction accuracy of the model is 85%, which indicates that the model is suitable.

IV. CONCLUSION

In this study, the principal component analysis method was used to screen the main influencing factor indexes, and the credit risk measurement model of China's listed real estate company was constructed based on the Logistics model. The conclusions are as follows:

1)Through the parameter estimation and goodness of fit test of the model, the Logistic regression model established in this study had achieved 85% prediction accuracy for the latest data samples, indicating that the model can be suitable for the measurement of credit risk and early warning of Chinese real estate companies.

2) Based on the results of model regression, the main factors that determine the credit risk of real estate enterprises in China were the profitability factor F_1 of the enterprise, the operational capability factor F_2 of the enterprise and the growth capability factor F_3 of the enterprise. The main relevant financial indicators were rate of net income on total assetsX_1, net profit margin on salesX_2, return on equityX_3, inventory turnover rateX_6, and growth rate of main operating revenueX_{12}. It shows that only when real estate company have a higher rate of return on assets, faster asset turnover rate and a higher main business growth rate can credit risks be reduced. A company caught in a credit crisis is not only a problem of shortfalls in capital or temporary losses, but also a deeper problem of business management. If these problems are not effectively solved, the credit risk of real estate enterprises will be intensified.

REFERENCES

[1] E.I. Altman, "Financial ratios, discriminant analysis and prediction of corporate bankruptcy," Journal of Finance, vol.23.1968, pp.589- 609.

[2] W. Sjur and W. Nicovander, "Default probabilities in a corporate bank portfolio: a logistic model approach," European Journal of Operational Research, 2001, pp.338-349.

[3] Shinong Wu and Xianyi Lu, "A study of models for predicting financial distress in China's listed companies," Economic Research Journal, vol.6, 2001, pp.46-55. (In Chinese)

[4] Fuxia Guo, Lan Lu and Qisheng Gao, "Evaluation analysis of credit risk for listed company based on principal component analysis and Logistic model," Journal of Qingdao university, vol.25, 2012, pp.106-110. (In Chinese)

[5] Jinhai Li, Linna Fu and Jian-hua Li, "Logistic mode-based credit rating assessment of listed companies," Journal of Hebei University of Technology, vol.40, 2011, pp.113-118. (In Chinese)

[6] Junzi Wang and Lantao Liu, "Research on real estate credit risk of commercial banks in China based on Logistic model," Review of Economy and Management, vol.2, 2017, pp. 86-95. (In Chinese)

[7] Sheng Hu, Huanhuan Lei and Huaqiang Hu, "Research on Credit Risk Measurement of China's Real Estate Enterprises Based on Logistic Model," China Soft Science, vol.12, 2018, pp.157-164. (In Chinese)