

基于 Logistic 模型的房地产行业信用风险研究

肖 冰,李春红

(重庆大学 经济与工商管理学院,重庆 400030)

摘 要:本文运用 Logistic 模型对房地产行业的信贷违约风险进行评估和预测,并对模型的拟合优度进行了检验,论证了该模型的实用性和准确性,得出房地产行业具有特殊的信贷风险特征的结论,并指出除财务指标外,宏观经济指标也是影响我国房地产公司信用风险的重要因素。

关键词:房地产行业;信贷风险;Logistic 回归分析;主成分分析

中图分类号:F832.45 **文献标识码:**A **文章编号:**1002-980X(2010)03-0065-04

信用风险是目前我国金融市场面临的主要风险之一,房地产信用风险更是占据了信用风险的很大一部分,如何提高房地产行业的信用风险管理水平是我国银行业面临的重大挑战,因此,对房地产行业的信用风险进行分析和预测变得非常迫切。从我国现状来看,目前商业银行推行内部信用风险模型的条件尚不成熟,因此,从量化分析的角度对其展开研究十分必要。本文利用 Logistic 模型对目前我国上市房地产公司的信用风险进行分析和预测,采用的样本数据为 2003 年以来的 97 家房地产上市公司的主要季度财务指标,并加入房地产企业特有的定性指标以及跟房地产业相关的宏观经济指标进行综合分析,从而得出 Logistic 模型回归的结果,并对房地产企业的信用风险进行预测。

1 文献综述

国内外关于房地产信贷风险的研究比较多见。Rupert Nabarro 和 Tony Key 对房地产行业的信贷周期和信用风险进行了研究,证明了房地产信贷周期中的关键要素能够被度量和检测,并且论证了房地产需求、租金和资本价格、贷款规模等指标的剧烈波动可能是房地产信贷危机的警示信号^[1];香港城市大学的王坤、王泽森对银行业房地产信贷风险的成因及对策进行了研究,结果表明灾难短视症、羊群效应、道德风险以及信息不对称等因素是房地产信贷风险的主要成因^[2];张婷、嵇玲竹于 2006 年对房地产价格与商业银行房地产信贷风险的关系进行了定性与定量分析,发现我国商业银行房地产信贷风险正随着房价的不断攀升和房地产泡沫的膨胀而逐

渐增大,房价因素已成为影响商业银行房地产信贷风险的重要因素^[3]。

此外,国内外用数学模型对信贷风险进行的研究也比较多见,能够应用于信贷风险研究的模型包括 KMV 模型、EDF 模型、人工神经网络模型、MDA 模型、Logistic 模型等。由于 Logistic 模型采用最大似然估计法进行参数估计,且不要求样本数据呈正态分布,因此本文主要采用 Logistic 模型。Logistic 模型最早是由 Martin 用来预测公司的破产及违约概率^[4],他在 1970—1977 年间从大约 5700 家美联储成员银行中界定出 58 家困境银行,并从 25 个财务指标中选取总资产利润率等 8 个财务比率来预测公司的破产及违约概率,建立了 Logistic 回归模型。该模型可以根据银行和投资者的风险偏好设定风险警戒线,以此为对象进行风险定位和决策。他还将 Z-Score 模型、ZETA 模型和 Logistic 模型的预测能力进行了比较,发现 Logistic 回归模型明显优于前两种模型。Ohlson 也利用 Logistic 模型做了关于财务比率和企业破产之间关系的研究^[5],从上千家企业财务报表中选取了财务数据,建立了有条件的逻辑模型,研究结果表明,公司规模、资本结构、公司业绩和变现能力这 4 个因素显著影响公司破产概率;乔卓、薛锋等应用 Logit 模型对我国上市公司的财务困境进行实证研究,结果表明 Logit 模型具有良好的预测精度,可以作为证券投资者和财务分析人员使用的一种有效的财务困境预测工具^[6]。随着信用风险研究的深入,应用于信用风险的度量模型也越来越多。为此,张宗益、朱小宗等于 2005 年测算了各模型的预测结果,发现 Logistic 模

收稿日期:2009-12-18

作者简介:李春红(1969—),女,云南人,重庆大学经济与工商管理学院副教授,博士,研究方向:金融市场研究;肖冰(1985—),女,湖南人,重庆大学经济与工商管理学院硕士研究生,研究方向:商业银行经营管理。

型的预测效果要明显优于 Z 模型、Z 模型和 ZETA 模型,犯第一类错误的概率显著下降,神经网络的预测效果相对于 MDA 模型和 Logit 模型并没有明显的改善^[7]。宋素荣等从上市公司的财务品质、公司战略和管理者行为等方面选取预警指标,并运用 Logistic 模型对测试样本进行检验,证实该模型的预测具有较高的准确性^[8];庞素琳的研究表明,Logistic 回归信用评价模型对总体的判别准确率达到了 99.06%,同样说明了 Logistic 具有较高的判别准确性^[9];南开大学经济学院的李萌证明了 Logit 模型具有非常可信的识别、预测及推广能力,是商业银行信用风险评估的有效工具^[10]。此外,国内将 Logistic 模型应用于上市公司财务风险预警和信用担保风险的研究也比较多见。

2 实证分析

2.1 Logistic 模型简介

Logistic 模型是一种非线性分类的统计方法,实际上是普通多元线性回归模型的推广。Logistic 模型采用最大似然估计法进行参数估计,不要求样本数据呈正态分布。该模型的最大优点在于它解决了因变量不连续回归的问题,特别是因变量为分类变量时非常适合使用该模型进行研究。Logistic 模型采用的是逻辑概率分布函数,它的具体形式为:

$$P(y_i = 1 | x_i) = P_i = \frac{1}{1 + e^{-(\alpha + \beta x_i)}} \quad (1)$$

其中, P_i 为第 i 个案例发生的概率,它是由一个由解释变量 x_i 构成的非线性函数,在式(1)两边取自然对数得到如下的线性函数:

$$\ln\left(\frac{P_i}{1 - P_i}\right) = \alpha + \beta x_i \quad (2)$$

式(2)将 Logistic 函数做了自然对数转换,这称作 Logit 形式,也称作 y 的 Logit,即 $\text{Logit}(y)$ 。Logistic 回归应用于因变量为二分类评定的回归分析,

应用 Logistic 回归模型可以识别自变量对因变量的影响方向,当 $P_i \geq 0.5$ 时,说明发生违约的概率较大,容易发生违约。模型系数采用 Wald 统计量检验,Wald 检验值越大,表明该自变量的作用越显著。

2.2 主成分分析

根据来自中国经济研究服务中心(China Center for Economic Research,CCER)的数据,本文选取了从 2003 年第 1 季度到 2008 年第 4 季度的 96 家房地产行业上市公司的主要财务数据,共得到 2178 个样本。根据银行对房地产客户采用的信用等级评价指标,再加入部分宏观经济指标得到以下 19 个分析指标(如表 1 所示)。

表 1 原始分析指标及其分类

指标分类	分析指标
偿债能力指标	资产负债率(X1)
	每股经营活动净现金流量(X2)
	流动比率(X3)
	负债现金流量比率(X4)
	利息保障倍数(X5)
盈利能力指标	总资产报酬率(X6)
	净资产收益率(X7)
	销售利润率(X8)
	每股收益(X9)
经营能力指标	存货周转率(X10)
	销售收入增长率(X11)
	净利润增长率(X12)
	净资产增长率(X13)
宏观经济指标	对美元汇率(X14)
	中长期贷款利率(X15)
	房屋销售价格指数(X16)
	GDP 变动率(X17)
	国房景气指数(X18)
	全国土地交易价格指数(X19)

对表 1 的分析指标进行主成份分析,并且采用方差最大化正交旋转法(Varimax)改变信息量在不同主成份上的分布,便于为主成份找到合理的经济解释。得到的结果如表 2 所示

表 2 主成分特征值与贡献率

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.353	25.106	25.106	3.353	25.106	25.106
2	2.537	20.458	45.564	2.537	20.458	45.564
3	2.012	14.676	60.240	2.012	14.676	60.240
4	1.508	10.522	70.772	1.508	10.522	70.772
5	1.133	6.480	77.252	1.133	6.480	77.252
6	1.082	4.611	81.863	1.082	4.611	81.863
7	1.042	4.023	85.886	1.042	4.023	85.886

根据主成份分析结果,得到如下 7 个重要因子。这 7 个因子的可解释度为 85.886%,基本符合要求。根据主成分分析结果得到的 7 个因子分别是:

$$Y_1 = -0.931 \times X_{14} - 0.908 \times X_{15};$$
$$Y_2 = 0.883 \times X_6 + 0.875 \times X_{10};$$
$$Y_3 = 0.876 \times X_{16} + 0.863 \times X_{19};$$

$$Y_4 = 0.828 \times X_{17};$$
$$Y_5 = 0.819 \times X_7 + 0.813 \times X_8 + 0.766 X_9;$$
$$Y_6 = 0.736 \times X_2 + 0.787 \times X_3;$$
$$Y_7 = 0.862 \times X_{13}.$$

以上公式说明: Y_1 为包含对美元汇率和中长期贷款利率在内的因子,主要代表了宏观经济指标的利率和汇率因素; Y_2 为包含总资产报酬率和存货周转率在内的因子,主要代表了企业的盈利能力和经营能力; Y_3 为包含房地产销售价格指数在内的因子,主要代表了宏观经济指标的房价和地价因素; Y_4 为包含 GDP 变动率在内的因子,主要代表了企业所处的宏观经济条件; Y_5 为包含销售利润率、净资产收益率和每股收益在内的因子,主要代表了企

业的盈利能力; Y_6 为包含每股经营活动净现金流量和流动比率在内的因子,主要代表了企业的偿债能力; Y_7 是包含净资产增长率在内的因子,主要代表了企业的经营能力。

2.3 回归分析

根据主成分分析的结果,将样本数据中的 ST 企业定为具有信贷违约风险的企业,得到 246 个违约样本。以 2003 年第一季度至 2006 年第四季度的 182 个 ST 数据及对照的 1238 个非 ST 数据为训练样本,以 2007 年第一季度至 2008 年第四季度的 64 个 ST 数据及对照的 694 个非 ST 数据为预测样本,采用 SPSS16.0 进行 Logistic 回归,得到的 Logistic 模型参数如表 3 所示。

表 3 Logistic 模型参数表

因子	B	S. E.	Wald	df	Sig.	Exp(B)
Y1	- 1.007	0.193	24.586	1.000	0.000	0.245
Y2	0.000	0.000	0.015	1.000	0.872	1.000
Y3	- 0.776	0.292	15.612	1.000	0.011	0.291
Y4	- 0.242	0.353	4.201	1.000	0.035	0.411
Y5	- 0.827	0.226	20.285	1.000	0.001	0.213
Y6	- 0.379	0.332	8.956	1.000	0.021	0.391
Y7	- 0.551	0.241	12.997	1.000	0.015	0.307
Constant	- 1.799	0.146	169.843	1.000	0.000	0.159

根据表 3,得到 Logistic 回归函数为

$$P = \frac{1}{1 + e^{-1.799 - 1.007Y_1 - 0.776Y_3 - 0.242Y_4 - 0.827Y_5 - 0.379Y_6 - 0.551Y_7}} \quad (3)$$

由以上回归结果可知,影响房地产行业信贷风险的主要指标包括:偿债能力指标:每股经营活动净现金流量和流动比率;盈利能力指标:销售利润率、净资产收益率和每股收益;经营能力指标:净资产增长率;宏观经济指标:对美元汇率、中长期贷款利率、房地产销售价格指数、全国土地交易价格指数和 GDP 变动率。

由于 Logistic 回归时不提供标准化的回归系数,因此需要将上述回归结果得到的回归系数进行标准化,计算公式如下:

$$\beta_i = \frac{b_i \times s_i}{\pi \sqrt{3}} \quad (4)$$

其中, β_i 为第 i 个自变量的标准化回归系数; b_i 为第 i 个自变量的非标准化回归系数; s_i 为第 i 个自变量的标准差,可以通过 SPSS 的 Statistics 模块中的 Summarize 命令得到; $\pi \sqrt{3}$ 是标准 Logistic 分布的标准差,近似等于 1.8138。根据式 (4) 得到六个变量的标准化回归系数分别为: $\beta_1 = -0.58$, $\beta_2 = 0$, $\beta_3 = -0.91$, $\beta_4 = -0.47$, $\beta_5 = -1.05$, $\beta_6 = -0.72$, $\beta_7 = -0.51$ 。因此,系数标准化后的 Lo-

gistic 回归函数为:

$$P = \frac{1}{1 + e^{-1.799 - 0.58Y_1 - 0.91Y_3 - 0.47Y_4 - 1.05Y_5 - 0.72Y_6 - 0.51Y_7}} \quad (5)$$

由式 (5) 可知,系数标准化后,宏观经济指标中 Y_1 ,即对美元汇率和中长期贷款利率的作用被减弱了,而 Y_5 ,即盈利能力指标中的销售利润率、总资产增长率和每股收益的作用变得明显了,其他指标的作用均有相应的加强, Y_1 对房地产行业信贷风险的影响依然明显,不是所有指标中最弱的。

综上所述,对房地产企业贷款信用风险影响最大的指标为盈利能力指标中的销售利润率、总资产增长率和每股收益;其次为房地产销售价格指数和全国土地交易价格指数;而宏观经济指标中的中长期贷款利率、对美元汇率和 GDP 增长率,财务指标中的每股经营活动净现金流量、流动比率和净资产增长率的作用也不容忽视。其次,回归结果证明,宏观经济指标是影响我国房地产公司信用风险的重要因素。

模型估计完成后,我们需要对模型的拟合优度进行检验,以评价其是否有效描述了变量以及模型匹配观测数据的程度。评价 Logistic 回归模型拟合优度的方法有很多,例如皮尔逊 χ^2 (Pearson χ^2) 检验、偏差 (Deviance) 检验以及 Hosmer-Lemeshow

拟合优度检验。其中 Hosmer Lemeshow 拟合优度指标(通常简称为 H-L),是由 Hosmer 和 Lemeshow 在 1989 年提出的一种专门针对 Logistic 模型拟合优度检验的方法,其检验效果明显由于其他检验方法,因此本文采用 Hosmer-Lemeshow 拟合优度检验,结果如表 4 所示。

表 4 Hosmer-Lemeshow 检验结果

Step	Hosmer and Lemeshow Test		
	Chi-square	df	Sig.
1	23.521	8	0.008

表 4 显示了 Hosmer-Lemeshow 的检验结果,其 HL 指标值为 23.521,自由度等于 8,这一指标与 χ^2 分布比较得到的概率值为 0.008,说明统计检验不显著,因此,可以认为模型很好地拟合了数据。模型对训练样本的判定结果如表 5 所示。

表 5 模型对训练样本的判定结果

Observed		Predicted		
		Y		Percentage Correct
Y	0	1167	91	94.3
	1	53	129	70.9
Overall Percentage				91.3

在数理统计中,通常犯的统计错误有两种:第一类错误是弃真,在本文中是指将可能发生信贷违约的企业误判为不会发生信贷违约的企业;第二种错误是纳伪,在本文中是指将不会发生信贷违约的企业误判为会发生信贷违约的企业。由表 5 可知,上述 Logistic 模型评价的总体准确率为 91.3%,其中犯第一类错误的概率为 29.1%,而犯第二类错误的概率仅为 5.7%,说明 Logistic 模型将低违约风险企业误判为高违约风险企业的概率较低,远远低于将高违约风险企业误判为地违约风险的概率。

模型的预测准确性如表 6 所示。

表 6 模型对测试样本的预测结果

Observed		Predicted		
		Y		Percentage Correct
Y	0	637	57	91. 8
	1	16	48	75. 0
Overall Percentage				90. 4

由表 6 可知,上述 Logistic 模型预测的总体准确性为 90.4%,其中犯第一类错误的概率为 25.0%,而犯第二类错误的概率为 8.2%,说明该

Logistic 模型对我国上市房地产公司的信贷风险预测整体准确度较高,但对 ST 公司的预测准确度偏低;同时也证明了 Logistic 模型比较稳定,推广能力较强。

3 研究结论

本文利用 97 家上市房地产公司 2003 年 1 季度到 2008 年 4 季度的财务数据以及相应的宏观经济数据,建立 Logistic 回归模型,对我国房地产企业信贷违约风险进行预测,得出以下结论:

- 1) 信贷风险模型 Logistic 对我国上市房地产公司的信贷风险预测整体准确度较高,但对 ST 公司的预测准确度偏低。
- 2) 对于房地产企业而言,决定其信贷风险大小的不仅仅是公司的财务状况,宏观经济指标例如房价、GDP 增长率、汇率和中长期贷款利率以及 GDP 增长率等也起到了重要作用。
- 3) 在公司的财务指标中,销售利润率、总资产增长率和每股收益是决定公司财务状况最重要的指标。

参考文献

[1] NABARRO R,KEY T. The real estate market measurement and real estate credit risk[J]. Journal of Banking and Finance,2002(3):18-26.

[2] 王坤,王泽森. 银行业房地产信贷风险成因及对策[J]. 金融改革,2007(5):18-20.

[3] 张婷,嵇玲竹. 房地产价格和商业银行房地产信贷风险的关系研究[J]. 价格理论与实践,2006(6):63-64.

[4] MARTIN D. Early warning of bank failure: a logit regression approach[J]. Journal of Banking and Finance,1977(3):249-276.

[5] OHLSON J. Financial ratios and the probabilistic prediction of bankruptcy[J]. Journal of Accounting Research,1980,18:109-131.

[6] 乔卓,薛锋. 上市公司财务困境预测 Logit 模型实证研究[J]. 湖南工程学院学报,2002(10):103-104.

[7] 张宗益,朱小宗. 传统信用风险度量模型的实证比较与适用性分析[J]. 预测,2005(2):55-59.

[8] 宋素荣,于丽萍. 上市公司财务危机预警的 Logistic 模型[J]. 财会月刊,2006(30):27-32.

[9] 庞素琳. Logistic 回归模型在信用风险分析中的应用[J]. 数学的实践与认识,2006(9):129-137.

[10] 李萌. Logit 模型在商业银行信用风险评估中的应用[J]. 管理科学,2005(2):33-38.

(下转第 78 页)

性杠杆效应有所增加,主要表现为坏消息对于市场波动性的影响力要大于等量好消息市场波动性的影响力。

参考文献

- [1] 杨星. 股指期货[M]. 广州:广东经济出版社,2002:6-9.
- [2] 娄群,步妍. 股指期货推出对现货市场影响的实证研究——基于台湾加权股价指数的实践[J]. 甘肃金融,2008(9):16-19.
- [3] 李华,程婧. 股指期货推出对股票市场波动性的影响研究——来自日本的实证分析[J]. 金融与经济,2006(2):81-83.
- [4] 李卓. 股指期货对股票现货市场的影响研究[D]. 长春:吉林大学,2007.
- [5] 王沛英. 香港股指期货市场的发展及其启示[J]. 统计与决策,2004(1):37-40.
- [6] 孙敬水. 计量经济学[M]. 北京:清华大学出版社,2004:335-337.

Empirical Study on Influence of Stock Index Futures on Stock Spot Market Fluctuation

Xie Lei ,Wang Yecheng

(School of Management ,Guilin University of Technology ,Guilin Guangxi 541004 ,China)

Abstract : In this paper ,the day closing price of Hongkong Hengsheng stock index futures from January 1984 to September 2008 is chosen as the original data to establish GARCH model and EGARCH model to make the empirical study. The conclusion is drawn that the introduction of Hengsheng stock index futures reduces the fluctuation of Hongkong stock spot market in some degree.

Key words : stock index futures ;stock market fluctuation ; GARCH model

(上接第 68 页)

Research on Credit Risk in Real Estate Industry Based on Logistic Model

Xiao Bing ,Li Chunhong

(Economics and Business Administration School ,Chongqing University ,Chongqing 400030 ,China)

Abstract : Based on the Logistic model ,this paper uses the maximum likelihood estimation method to estimate the parameters ,and assesses and predicts the credit default risk in real estate industry ,and makes a goodness-of-fit test to this model. The conclusion is that the real estate industry has special credit risk features.

Key words : real estate industry ;credit risk ;Logistic regression analysis ;principal component analysis