

上市公司房地产企业财务风险识别 ——基于 lasso 模型的实证研究

王君铭 吴艳辉 何慧

厦门大学

摘要

在宏观经济紧缩和严厉调控政策的影响下,我国房地产企业已经出现了显著的财务风险,对房地产企业财务风险的分析与防范,有利于帮助企业及时发现问题,尽可能降低财务风险。本文基于 2018 年 122 家房地产上市公司的相关数据,分析房地产企业财务风险会受到公司哪些财务指标的影响。首先对样本数据进行简单统计分析;接着用 lasso 方法进行变量选择,选出对公司财务风险影响的系数较大的财务指标;最后进行 logit 模型回归,得出资产负债率、总资产净利率和总资产报酬率三个财务指标和房地产公司财务风险显著相关,且总资产净利率越大,财务风险越低。

1 引言

近年来,随着国民经济的发展以及城市化进程的加快,房屋销售价格上涨明显,房地产业也进入到一个飞速发展的阶段。房价高涨为国民住房的民生问题带来了挑战,政府出台多项针对房地产企业的调控政策,以制止各种炒房、投机性购房行为,使房地产企业面临需求购买力的透支、过度的房地产投资、较高的待售面积,以及较为谨慎的货币政策等风险;同时,房地产企业项目开发周期长、资金占用金额高、资金回拢时间长、行业内部相互关联性高等特点,使其面临的财务风险进一步加大。同时,房地产企业始终在中国经济的发展中占据着重要的地位。房地产行业的任何波动都会带动中国经济的波动,同样其他行业的变动也会引起房地产行业的变动,因此,房地产行业面临财务风险的机率更高,这就要求房地产企业加强对财务风险的管理,有效进行财务风险评估与防范,以避免或降低财务风险。因此,对于房地产企业财务风险的研究极为有意义。

2 文献综述

2.1 财务风险评价模型研究综述

2.1.1 国外研究现状

国外学者试探了多种评价模型来预测企业的财务风险和信用风险，包括 Z 计分模型、Logistic 模型、Probit 模型、KMV 模型、神经网络模型等。Altman[1] 最早提出了 Z 计分模型，文章通过筛选出的 5 个具有代表性的指标进行了多元线性回归，并计算出一个总体的判别值-Z 值，根据 Z 值的取值来判断企业财务风险的高低。Zmiewski[2] 使用了 Probit 模型对企业进行财务风险和信用评价。KMV 模型最早由 KMV 公司于 1997 年提出，最早的 KMV 模型在企业风险评价的预测能力不强，经由国外学者不断的改进才提升了模型的评估效力，如 HCWu[3] 结合了 KMV 模型和随机森林算法进行企业信用评价，其预测效果显著。神经网络算法和支持向量机模型在财务风险评价领域也得到了推广。Back[4] 较早的将神经网络思想应用到风险评价的研究中。Ollhson[5] 首次使用条件 Logit 模型进行财务风险和信用评价，研究发现 Logit 模型适合对企业的违约概率进行准确估计。Psillakiabc[6] 使用了包括 Logistic 模型在内的多种预测模型进行对比分析；后来 Pereira[7] 等对 Logistic 模型做了诸多改进，由此可见 Logistic 模型在财务风险评价领域得到了广泛的应用。

2.1.2 国内研究现状

国内的许多学者借鉴现有指标体系，改进成熟的评价模型，通过多模型的对比分析，试图寻找适合当前中国企业的财务风险评价模型。周首华、杨济华和王平 [8] 认为 Z 计分模型的指标基本都是收益类指标，指标类型过于局限，不能真实反映企业财务状况。他们将对企业产生重要影响的现金流量类指标加入模型中，通过对 Z 模型的修正得到了 F 分数模型，并将真实的企业数据带入 F 分数模型中，最终得到财务风险 F 分数临界点。吴世农 [9] 应用 Fisher 线性判定分析、多元线性回归分析和 Logistic 回归分析三种方法，分别建立三种预测财务困境的模型，得出了相对同一信息集而言，Logistic 预测模型的误判率最低的结论。张爱民 [10] 借鉴 Altman 的 Z 值模型，结合主成分分析法构造了预测上市企业财务失败的主成分预测模型，模型预测能力显著增强；赵天任 [11] 则将 Z 计分模型与 KMV 模

型结合起来评价中小企业的财务风险，得到了两种方法结合可以提高预测效果的结论；也有众多学者采用了与 Altman 模型相结合的方法，其中郑享清 [12] 采用了主成分分析与 Altman 模型结合的方法，方匡南 [13] 提出了网络结构的 Altman 模型进一步提升了风险识别模型效力。此外也有很多新方法运用到财务风险评价中，如赵亚 [14] 引入了随机森林算法，郑治平 [15] 则基于因子分析法提取了企业财务风险特征，提出了基于神经网络的风险识别模型，构建了一种新的企业财务风险的分类方法。

2.2 基于财务指标的财务风险评价研究综述

国外研究现状 Fitzpatrick[16] 通过对成功的和失败的企业的一系列财务数据进行分析对比，发现“净利润/股东权益”和“股东权益/负债”这两个指标在所有指标中对财务风险的判别能力最高，引起了学者对财务风险实证研究的关注。William Beaver[17] 提出了单一判别变量模型，即使用个别财务指标对财务风险进行识别，这也是首次在财务风险评价中使用个别财务指标变量进行风险识别。Beaver 发现“现金流量/总负债”比率可以有效地判别财务风险困境。Altman[1] 选择 33 组破产企业 and 非破产企业，按照分层随机配对的原则将这两组进行配对，对 20 多个可能有用的变量进行分析，最终选出“营运资本/总资产”、“留存收益/总资产”、“息税前利润/总资产”、“权益市值/账面总负债”和“收入/总资产”作为分析企业财务风险的变量组合，通过分析得到了经典的 Z 计分模型。Blum[18] 从现金流量的角度对财务风险进行分析，根据行业分类、净销售收入和员工规模三个方面为 115 家破产企业配对运转正常的企业，利用多变量判别分析法得到了一个财务危机评价模型，变量包含“净现金流量/总负债”这一指标。Ohlson[19] 对 105 家倒闭企业和 2058 家正常运营企业在一段时间内陷入财务风险概率进行分析，发现“总资产对数”、“负债总额/资产总额”、“营运资本 /资产总额”、“流动负债/流动资产”、“净利润/资产总额”等指标对财务风险识别最显著。

2.2.1 国内研究现状

张玲 [20] 将选取的上市企业分为 2 组，每组样本中都包含有 ST 企业和非 ST 企业，并从 15 个财务指标中最后推导出资产负债比率、营运资金/总资产比率、总资产利润率、留存收益/资产总额比率 4 个变量建立判别函数来判别企业的财务风险状况，模型具有提前 4 年预测效果。杨保安、季海 [21] 将短期偿债能力比率、长期偿债能力比率、盈利能力比率、营运能力比率共 4 类 15 项财务指标纳入到前向三层 BP 网络模型中，对贷款企业的

财务状况进行分析,为商业银行的贷款方案选择提供依据。李素红、陈立文 [22] 筛选出 50 家房地产上市企业,对这些企业 6 个方面的财务能力共 20 个财务指标进行因子分析,得到各个企业的综合因子得分,并进行排序,发现多数房地产上市企业的财务风险较高,应该引起注意,最后根据排序将财务风险划分为 4 个等级。杜运潮、徐凤菊 [23] 初步选取企业盈利能力、资产运营能力、偿债能力、成长能力以及现金流量能力 5 个方面共 30 个指标,经过回归分析最终筛选出“营业利润”、“流动比率”、“净利润/营业总收入”3 项指标建立财务风险评价模型,分析房地产上市企业的财务状态。张煌 [24] 强采用熵权法与综合评价法相结合的方法对房地产企业进行研究,通过 10 个财务指标对所选企业的财务风险进行计算排序,发现存货周转率、销售收入增长率、资产负债率、净资产收益率对房地产财务风险的影响最为显著。赵振鲁 [25] 针对我国上市房地产企业的实证研究,提出了运用 COX 生存模型,分别从企业的盈利能力、营运能力、偿债能力、现金流量状况和公司治理共五方面综合选取相关财务指标和非财务指标等 23 个风险预警变量,通过定性与定量相结合的方法对财务风险进行了分析与预测。

2.2.2 文献综述评述

纵观国内外文献,数理统计方法在企业财务风险评价中得到了很大重视,甚至人工智能的方法也得到了一定的应用,使得企业对财务风险的管理更加科学和精确。在模型的选择上,国内外学者运用了多种模型研究财务风险,并进行了大量的对比分析,结论表明没有绝对准确的评价模型,都需要对模型进行适当的修正才能提高评价的准确率;在评价财务风险的指标选取上,国内外学者尝试了多种财务指标的比较分析,极大地丰富了财务风险评价相关理论和实证研究。国内已有研究大多侧重于对回归模型的改进以及不同模型间的比较分析,同时模型所使用的指标体系大多借鉴国外文献,鲜有对于指标筛选方法的对比分析研究。本文在前人研究的基础上,采用 2018 年 122 家房地产上市公司的财务数据作为样本数据,首先使用 LASSO 指标筛选的方法进行筛选,以筛选得到的指标作为自变量建立 Logistic 回归模型,通过回归模型预测房地产企业财务风险。

3 相关概念定义与介绍

3.1 财务风险的定义

财务风险主要分两类：一、广义的财务风险，广义的财务风险是公司在日常经营活动中，从整体各个环节中所遇到的会给公司的经营带来阻碍，造成公司财务指标背离最初订立的财务目标的情形，而带来相应的财产流失，同时损害公司利润的情况。广义的财务风险对公司来说是不可避免的，虽不可避免但是可以寻找一定的规律进行推测，公司可以运用这一方面特点采取有效措施对财务风险进行控制，严重的财务风险可能会导致公司破产。因此，为了可持续的发展，公司应重视财务风险；二、狭义财务风险狭义的财务风险指公司进行短期借款但由于资不抵债，公司到还款日没有足够的资金进行款项的归还，由此产生的财务风险。公司可以没有狭义的财务风险，因为只要公司不进行借款活动，就不会出现到期无法还款的情况，也就没有了狭义的财务风险。

3.2 Lasso 方法简介

Lasso(Least Absolute Shrinkage and Selection Operator) 是 Tibshirani 在 1996 年提出的一种能够减少指标的方法，该方法能够在参数估计的同时实现变量的选择，它的基本思想是在回归系数的绝对值之和小于常数的约束条件下，使残差平方和最小化，从而能够产生某些严格等于 0 的回归系数，达到筛选指标的目的，同时得到解释能力强的模型。借鉴 Tibshirani (1996) 的理论，Lasso 方法的基本原理简介如下：

假如有数据变量 $(x^i, y_i), i = 1, 2, \dots, n$ 其中， $X^i = (x_{i1}, x_{i2}, x_{i3} \dots x_{ip})$ 和 y_i 分别是解释变量和被解释变量的观测值。在一般的回归模型中，常常认为观测值彼此独立或者说被解释变量 y_i 在给定解释变量的条件下是相互独立的。同时，假设 x_{ij} 是标准化的，即 $\frac{1}{n} \sum x_{ij} = 0$ $\frac{1}{n} \sum x_{ij}^2 = 1$ 。令 $\hat{\beta} = (\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_p)^T$ ，Lasso 方法的估计量 $(\hat{\alpha}, \hat{\beta})$ 定义为：

$$(\hat{\alpha}, \hat{\beta}) = \operatorname{argmin} \left\{ \sum_i^n \left(y_i - \alpha - \sum_j \hat{\beta}_j x_{ij} \right)^2 \right\} \text{ s.t. } \sum_j |\beta_j| \leq t \quad (1)$$

其中 $t \geq 0$ 是调和参数，此时对任何的 t ，有 α 的估计 $\hat{\alpha} = \bar{y}$ 。同时为

不失一般性，假定 $\bar{y} = 0$ ，这样可上式整理如下：

$$\hat{\beta} = \operatorname{argmin} \left\{ \sum_i \left(y_i - \sum_j \hat{\beta}_j x_{ij} \right)^2 \right\} \text{ s.t. } \sum_j |\beta_j| \leq t \quad (2)$$

对调和参数 t 的控制将会使回归系数总体变小。若令 $\hat{\beta}_j^0$ 是回归参数的最小二乘估计值， $t_0 = \sum_j |\hat{\beta}_j^0|$ ， $t \leq t_0$ 。就会使一些回归系数缩小并趋于 0，有些甚至会等于 0。例如当 $t = t_0$ ，计算的结果就会使不为 0 的回归系数的数目减少到大约为 $\frac{p}{2}$ 个。上述式的表达式还可以用如下式的惩罚函数的形式表达。

$$\hat{\beta}^L = \operatorname{argmin}_{\beta} \left\{ \sum_{i=1}^N \left(y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^p |\beta_j| \right\} \quad (3)$$

上式的第一部分表示模型拟合的优良性，第二部分表示参数的惩罚。调和系数 $\lambda \in [0, \infty]$ 越小，模型的惩罚力度就越小，保留的变量就越多； λ 越大，模型的惩罚力度就越大，保留的变量就越少。Lasso 回归的特色就是在建立广义线型模型的时候，这里广义线型模型包含一维连续因变量、多维连续因变量、非负次数因变量、二元离散因变量、多元离散因变量等。除此之外，无论因变量是连续的还是离散的，Lasso 都能处理。总结来说，Lasso 方法对于数据的要求是极其低的，所以应用范围较广；此外，Lasso 还能够对变量进行筛选和对模型的复杂程度进行降低，这是 Lasso 模型区别于其它方法的独特优势。这里的变量筛选是指不把所有的变量都放入模型中进行拟合，而是有选择的把变量放入模型从而得到更好的性能参数。

3.3 Logistic 模型简介

Logistic 回归模型是最成熟也是应用最广泛的分类模型，它主要用来计量离散因变量与一组解释变量之间的关系。Logistic 模型按因变量类型可分为二分类 Logistic 模型、多分类无序 Logistic 模型，多分类有序 Logistic 模型三种。借鉴何晓群（2004）的相关理论，结合本文内容需要，下面对二分类 Logistic 回归模型基本原理做简要介绍。

假设因变量 Y 是只取 0 和 1 的随机变量，自变量为 X_1, X_2, \dots, X_p ，一般的线性回归模型已经无法刻画自变量和因变量的关系，因为此时因变量只有两个取值。将二分类因变量的取值转换成发生的概率，在因素 (X_1, X_2, \dots, X_p) 下事件 $(Y = 1)$ 生的概率记为 P ，则事件 $(Y = 0)$ 不发生的概率为 $1 - P$ ，

这样就把因变量与自变量的非线性关系转变成了概率值与自变量的线性关系，得到以下形式的模型

$$\text{logit}P = \beta_0 + X_1\beta_1 + X_2\beta_2 + \dots + X_p\beta_p \quad (4)$$

其中 $\text{logit}P \equiv \ln\left(\frac{P}{1-P}\right)$ 为 Logit 变换。 $\frac{P}{1-P}$ 为履约概率与违约的概率之比，也称为优势比，其值越大表示发生违约的概率就越小， $\text{logit}P$ 为优势比的对数。对上式进行转化之后可得：

$$P = \frac{1}{1 + e^{-\beta_0 + X_1\beta_1 + X_2\beta_2 + \dots + X_p\beta_p}} \quad (5)$$

这是 Logistic 模型的另一种表达式，通过 Logistic 模型可以对事件是否发生以及发生的可能性进行预测。一般选定 $P=0.5$ 作为判定的临界值。当预测概率比临界值小时，认为该事件出现的几率较小，当预测概率与临界值相比较小时，则认为事件发生的几率是比较大的。

4 研究设计

我国证监会将上市公司分为 ST 公司和非 ST 公司，ST 指 special treatment，ST 公司则是指公司的财务情况或其他情况出现异常，其流通的股票被进行特别处理。连续三年亏损的上市公司，其股票会被进行退市风险警示，因此公司是否被 ST 是公司是否存在财务状况的重要指标。本文选取房地产行业 2018 年的财务数据作为研究样本。将房地产公司是否为 ST 作为公司是否存在财务危机的判断标准，从公司的偿债能力、营运能力、盈利能力、发展能力和现金流指标五个方面进行财务指标的选取，将这些指标作为解释变量，来预测公司是否存在财务危机。在对指标的处理过程中采用 lasso 模型对这些指标进行分析和选择，选出最能影响企业财务状况的指标构建企业的财务风险预警模型。本文数据主要来源于国泰安数据库中公司研究系列中的财务报表和财务指标数据库。本文的实证分析工具选用的是 stata15。

表 1: 房地产企业财务风险预警指标构建

	财务指标	编号	定义
偿债能力指标	流动比率	A1	流动资产/流动负债
	速动比率	A2	速动资产/流动负债
	现金比率	A3	现金及现金等价物期末余额/流动负债
	经营活动现金比率	A4	经营活动产生的现金流量净额/流动负债
	资产负债率	A5	总负债/总资产
	长期资本负债率	A6	非流动负债合计/(所有者权益合计 + 非流动负债合计)
盈利能力指标	总资产报酬率	B1	(利润总额+财务费用)/平均资产总额
	总资产净利率	B2	净利润/总资产平均余额
	净资产收益率	B3	净利润/股东权益平均余额
	营业毛利率	B4	(营业收入-营业成本)/营业收入
	营业净利率	B5	净利润/营业收入
	应收账款周转率	C1	营业收入/应收账款平均占用额
营运能力指标	存货周转率	C2	营业成本/存货平均占用额
	流动资产周转率	C3	营业收入/流动资产平均占用额
	固定资产周转率	C4	营业收入/固定资产平均净额
	总资产周转率	C5	营业收入/平均资产总额
	总资产增长率	D1	(期末总资产-期初总资产)/期初总资产
	净资产收益率增长率	D2	(期末净资产收益率-期初净资产收益率)/期初净资产收益率
发展能力指标	净利润增长率	D3	(期末净利润-期初净利润)/期初净利润
	营业收入增长率	D4	(期末营业收入-期初营业收入)/期初营业收入
现金流量指标	营业收入现金含量	E1	商品、劳动收到的现金流入/营业收入

5 实证分析

本文选择 2018 年的 122 家房地产上市公司的相关财务指标作为本次实证分析的样本数据。本文被解释变量为房地产企业是否是高财务风险企业，变量名定义为 ST，若该房地产企业存在高财务风险，则 ST 取值为 1，否则取值为 0。本文的解释变量共包括 21 个财务指标。由于相对样本数量来说 p 值较大，可能会出现数据维度较高，导致稀疏模型，使得变量的方差过大，模型预测能力较差，因此本文先采用 lasso 模型进行进行变量筛选，更好地进行财务风险指标的选择，最后对选出的财务指标采用 logit 模型进行回归分析。

5.1 描述性统计分析

本文首先在 stata 中，对样本数据进行描述性统计，结果如下图 1 所示。

Variable	Obs	Mean	Std. Dev.	Min	Max
st	122	.0409836	.1990698	0	1
a1	122	2.040829	1.993496	.479494	20.93767
a2	122	.9858859	1.96678	.127488	20.73162
a3	122	.4602959	.8728453	.004227	7.443999
a4	122	.1158164	.5875384	-1.583276	5.414872
a5	122	.6453816	.1950579	.118612	.983857
a6	122	.0690124	.2495695	-.643236	1.421061
b1	122	.0447299	.0523912	-.331097	.158394
b2	122	.0253883	.0516494	-.392047	.163308
b3	122	.0761855	.2034959	-1.848934	.415964
b4	122	.3597373	.1513195	.018676	.793203
b5	122	.0269492	1.417091	-14.64651	4.382133
c1	122	303.9499	1235.058	.60226	9754.212
c2	122	82.42583	756.2116	.023212	8271.233
c3	122	.3529517	.391818	.008472	2.803563
c4	122	48.90973	114.4405	.249298	915.437
c5	122	.2345653	.196754	.004594	1.793533
d1	122	.127574	.2505131	-.490301	1.032811
d2	122	-1.452202	17.97714	-192.793	24.37816
d3	122	.409911	5.439057	-22.00968	42.656
d4	122	.2501947	.9663278	-.841448	7.109095
e1	122	1.347046	1.037614	.376932	10.4634

图 1: 描述性统计结果

从图中可以看出这个指标的平均值，最小值和最大值。例如其中 A5 指标为总资产负债率，其样本均值为 0.6454，最小值和最大值依次为 0.1186 和 0.9839。

5.2 lasso 回归

接着，本文用 lasso 模型进行变量选择。本文用 lasso2 进行 lasso 估计，并画出了整个解的路径。下面图 2 给出了根据参数 λ 的变化，lasso 估计量筛选出变量进入模型的情况。

Knot	ID	Lambda	s	L1-Norm	EBIC	R-sq	Entered/removed
1	1	20.92075	1	0.00000	-388.75544	0.0000	Added _cons.
2	2	19.06221	2	0.00208	-383.69408	0.0084	Added d4.
3	6	13.13884	4	0.02252	-384.01643	0.1050	Added a6 c4.
4	11	8.25158	5	0.09936	-393.27683	0.2108	Added d1.
5	13	6.85060	6	0.13999	-392.35261	0.2435	Added a5.
6	18	4.30238	7	0.22359	-394.64521	0.2937	Added d3.
7	21	3.25459	8	0.29341	-392.30763	0.3151	Added b2.
8	22	2.96546	9	0.33867	-387.41461	0.3218	Added b4.
9	25	2.24326	10	0.46215	-383.92715	0.3361	Added c5.
10	29	1.54619	11	1.27086	-379.58974	0.3456	Added b1.
11	31	1.28368	12	3.09358	-376.10330	0.3594	Added b5.
12	32	1.16964	14	3.90277	-364.55127	0.3627	Added a1 c3.
13	33	1.06573	15	4.67713	-357.38146	0.3570	Added a4.
14	34	0.97106	16	5.43487	-380.33448	0.4932	Added c2.
15	35	0.88479	17	6.26425	-378.16263	0.5092	Added b3.
16	39	0.60985	18	10.01603	-386.96209	0.5656	Added a2.
17	41	0.50631	19	11.43636	-386.51661	0.5852	Added a3.
18	43	0.42035	20	12.73028	-385.09165	0.6008	Added d2.
19	47	0.28973	19	14.80800	-397.26577	0.6202	Removed d3.
20	51	0.19970	20	16.15269	-393.95820	0.6288	Added e1.
21	52	0.18196	21	16.44233	-388.38501	0.6303	Added d3.

图 2: 微调参数 λ 模型变量进入情况

从图 2 可以看出, 当微调参数 λ 由大到小变化时, lasso 估计量也筛选出越来越多的财务指标变量进入模型。当 λ 的值大于 20.92075 时, 由于惩罚力度过大, 所有变量的系数都为 0。当 lambda 等于 20.92075 时, 常数项首先进入模型; 而当 λ 降为 19.06221 时, 变量 d4 首先进入模型, 以此类推。当 λ 的值为 6.8506 时, 变量 A5 进入模型。

下图 3 给出了 lasso 估计整个解的路径。

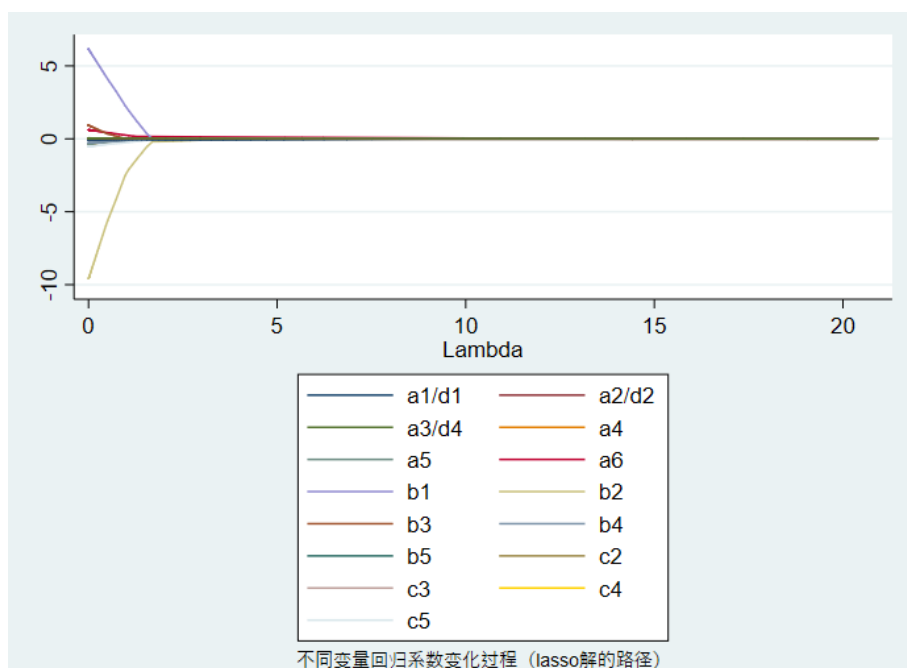


图 3: lasso 估计解的路径图

从图 3 中可以看出，当 λ 等于 0 时，即图中的最左边，不存在惩罚项，此时的 lasso 等价于 ols。当 λ 值越来越大时，惩罚力度加大，进入模型的变量减少。当 λ 的值很大时，图 3 中的最右边，由于惩罚力度过大，所有变量系数都归于 0。根据 EBIC，lasso 估计选出了 λ 值为 0.21917 时的模型，lasso 选出的变量结果如下图 4 所示。

Use lambda=.2191690827982403 (selected by EBIC).

Selected	Lasso	Post-est OLS
a1	-0.0329644	-0.0482580
a2	0.0188283	0.0242572
a3	0.0313812	0.0574632
a4	-0.2882560	-0.3622909
a5	-0.2944292	-0.3493897
a6	0.5363888	0.6238872
b1	5.2713791	6.1642085
b2	-7.8559005	-9.5882266
b3	0.6782282	0.9390305
b4	-0.2388236	-0.2714967
b5	0.0188143	0.0234511
c2	0.0002478	0.0002968
c3	-0.0100964	-0.0137607
c4	0.0005421	0.0005393
c5	-0.4610366	-0.5293445
d1	-0.0755651	-0.0824585
d2	-0.0009015	-0.0016167
d4	0.0447758	0.0460460
Partialled-out*		
_cons	0.3209287	0.3815323

图 4: lasso 变量选择结果

图 4 中第一列给出了 lasso 估计选择出来的变量，第二列是 lasso 所估计的变量系数。从表中可以看出，除了常数项以外，还有 18 个变量的系数是非零的；其他 3 个变量的系数为零，不进入模型。进入模型的变量包括 A1—A6, B1-B5, C2-C5, D1,D2 和 D4，其中变量系数较大的几个变量为 B1,B2,B3,A5,A6,A4,C5 等，表明这些变量对企业的信用风险影响相对较大。

另外，考虑到作为收缩估计量的 lasso 存在偏差 (bias)，图 4 中第三列给出了“post lasso”估计量的结果，即只使用 lasso 进行变量选择，然后扔掉 lasso 的回归系数，再对筛选出来的变量进行 OLS 回归。从图中可以看出，用 lasso 进行变量筛选以后，得出的 lasso 回归系数和 OLS 回归系数方向完全一致，绝对值大小接近，说明本次 lasso 变量选择结果合理。

5.3 logit 回归

在上述 Lasso 模型选择变量的基础上, 本文继续对房地产企业的信用风险进行分类和识别。本文采用 logistic 模型对样本数据进行回归, 房地产企业信用风险 ST, 1 表示高违约风险, 0 表示低违约风险, 并且各个财务指标同样进行了标准化处理, 因此在回归模型中系数的绝对值大小具有可比性。本文根据 lasso 选择出来的变量和变量系数的大小, 进行了五次 logistic 模型回归, 回归的结果如下所示。

	m1	m2	m3	m4	m5
st					
a1	-40.9905				
	()				
a2	41.6802				
	()				
a3	-38.8621				
	()				
a4	-27.0663	-0.0892	-0.0892	0.7033	
	()	(-0.06)	(-0.06)	-1.02	
a5	-322.0215	-16.7394 [*]	-16.7394 [*]	-11.7050 ^{**}	-13.0536 ^{***}
	()	(-1.91)	(-1.91)	(-2.06)	(-2.65)
a6	162.1196	0.1706	0.1706	-0.2605	
	()	-0.07	-0.07	(-0.13)	
b1	5829.0509	109.3459 ^{**}	109.3459 ^{**}	78.8654 ^{**}	81.3067 ^{***}
	()	-2.07	-2.07	-2.31	-2.65
b2	-8397.6304	-179.5003 ^{**}	-179.5003 ^{**}	-90.9817 ^{**}	-94.2694 ^{***}
	()	(-2.18)	(-2.18)	(-2.37)	(-2.72)
b3	752.6919	13.7259	13.7259		
	()	-1	-1		
b4	-317.7855				
	()				
b5	16.694				
	()				
c2	0.3782				
	()				
c3	-211.7646				
	()				
c4	0.2037				
	()				
c5	-452.3144	0.7608	0.7608		
	()	-0.17	-0.17		
d1	-18.1471				
	()				
d2	2.3245				
	()				
d4	50.668	0.6417	0.6417		
	()	-1.57	-1.57		
cons	291.5469	2.0793	2.0793	0.5772	1.4239
	()	-0.79	-0.79	-0.29	-0.92
N	122	122	122	122	122
t statistics in parentheses					
* p<.10, ** p<.05, *** p<.01					

图 5: logit 模型回归结果

模型 1(m1) 到模型五 (m5), 逐步选出对当地产企业信用风险影响较大

的财务指标变量。从图 5 中可以看出，A5 资产负债率、B1 总资产报酬率、B2 总资产净利率这三个指标，在模型 2 到模型 5 中对企业信用风险的影响都是显著的；并且当变量指标数量减为三个时，A5、B1、B2 三个指标都是在 99% 的水平上显著。

其中，A5 资产负债率和 B2 总资产净利率这两个指标与房地产企业信用风险显著负相关，B1 总资产报酬率与房地产企业信用风险显著正相关。B2 总资产净利率与房地产企业信用风险显著负相关，说明总资产净利率越大，企业信用风险越低，这完全符合财务管理理论中的实际情况。A5 资产负债率与房地产企业信用风险显著负相关，因为房地产企业具有一定的特殊性，其负债中有很大部分是来自于预收账款，预收账款在建筑竣工并且交房以后会转入营业收入，并且预收账款越多说明房地产企业营销收入越高，可用现金越多，从而企业财务信用风险越低，这与本文实证分析的结果一致。B1 总资产报酬率与房地产企业信用风险显著正相关，其中总资产报酬率等于利润总额与财务费用的加总和再除以平均资产总额，其中房地产企业的财务费用越多也会造成信用风险越大，故而这也与本文的实证结果一致。

6 结论

房地产行业具有项目开发周期长、投资规模大、涉及面广的特点，房地产企业在项目运营中面临着高风险，所以对房地产企业的信用风险进行评估非常重要。本文基于 2018 年 122 家房地产企业上市公司的财务数据建立评价指标体系，首先用 lasso 模型回归对信用风险评价指标进行筛选，然后采 logistic 模型对房地产企业的信用风险进行分类和识别，实证分析结果表明，资产负债率、总资产净利率和总资产报酬率对房地产企业信用风险影响显著，并且资产负债率和总资产净利率与房地产企业信用风险显著负相关，总资产报酬率与房地产企业信用风险显著正相关。

本文采用 lasso 方法不仅实现了财务指标体系的精简，且在可操作性和计算量方面具有一定优势；通过 logistic 财务风险识别模型的判定结果与实际基本相符，其对企业财务风险预测的准确程度也较高。模型的回归结果表明，运用 lasso 和 Logistic 模型可以有效地预测和检测样本房地产上市公司的财务风险，同样其他房地产公司也可以运用此模型建立符合自身特点的风险评估模型，从而有效地制定风险预警机制，保证企业在市场中拥有核心竞争力和持续竞争优势。

参考文献

- [1] Altman, E.I.. Financial Ratios, Discriminant Analysis and The Prediction Of Corporate Bankruptcy [J]. The Journal of Finance, 1968(23):568-609.
- [2] ME Zmijewski.Methodological Issues Related to the Estimation of Financial Distress Prediction Models[J].Journal of Accounting Research,1984,221:59-82.
- [3] HC Wu,YT Wu.Evaluating credit rating prediction by using the KMV model and random forest[J]Kybernets,2016,4510:1637-1651. B Back,T Laitinen,K S.Neural net works and genetic algorithms for bankruptcy predictions[J]Expert Systems with Applications.1996,11(4):407-413.
- [4] Ohlson, James A. Financial Ratios and The Probabilistic Prediction Of Bankruptcy [J]. Journal Of Accounting Research, 1980(1):109-131.
- [5] M Psillakiabc.Evaluation of credit risk based on firm performance[J].European Journal of Operational Research,2010,201(3):873-881
- [6] JM Pereira,M Basto,AFD Silva.The logistic lasso and ridge regression in predicting corporate failure[J].Procedia Economics & Finance,2016,39:634-641
- [7] 周首华, 杨济华, 王平, 论财务危机的预警分析——F 分数模式 [J]., 会计研究, 1996 (8): 8-11。
- [8] 吴世农, 卢贤义, 我国上市公司财务困境的预测模型研究经济研究 [J]., 2001, (6): 46-55。
- [9] 张爱民, 祝春山, 上市公司财务失败的主成分预测模型及其实证研究 [J], 金融研究, 2001, (3): 10-25。
- [10] 赵天任, 周涵俊, 基于 Z 计分与 KMV 模型的我国中小企业信用风险评价研究 [J], 经营管理者, 2010, (23): 40-40。
- [11] 郑享清, 王见, 吴莉君, 主成分分析法下企业信用风险评估研究 [J], 财会通讯, 2010, (23): 140-141。

- [12] 方匡南, 范新妍, 马双鹤, 基于网络结构 Logistic 模型的企业信用风险预警 [J], 统计研究, 2016, 3 (4): 50-55。
- [13] 赵亚, 李田, 苑泽明, 基于随机森林的企业信用风险评估模型研究 [J]. 财会通讯, 2017, (29)。
- [14] 郑治平, 基于模式识别的企业财务风险预测方法 [J]. 商业会计, 20156, (6): 25-28。
- [15] Fitzpatrick, P.J..A Comparison of Ratios of Successful Industrial Enterprises with those Failed Firm [J]. Certified Public Accountant,1932:598-605.
- [16] Beaver, W.H.. Financial Ratios as Predictors of Failure [J]. Journal of Accounting Research, 1966(4):71-111.
- [17] M. Blum. Failing Company Discriminant Analysis [J]. Journal of Accounting Research 1974(12):1-25.
- [18] 张玲, 财务危机预警分析判别模型, 数量经济技术经济研究 [J], 2000 (3): 49-52。
- [19] 张玲. 财务危机预警分析判别模型. 数量经济技术经济研究 [J], 2000 (3): 49-52.
- [20] 杨保安, 季海, 基于人工神经网络的商业银行贷款风险预警研究 [J]., 系统工程理论与实践, 2001 (5): 70-75。
- [21] 李素红, 陈立文, 基于因子分析法的房地产上市公司财务风险评价 [J], 河北工业大学学报, 2011, 40 (6): 101-106.
- [22] 杜运潮, 徐凤菊., 房地产上市公司财务风险识别——基于 Logistic 模型的实证研究 [J]. 财会月刊, 2013 (9): 19-21.
- [23] 张煌强. 房地产企业财务风险评价研究 [J]. 广西社会科学, 2015 (7): 81-85
- [24] 赵振鲁, 基于 COX 模型的上市房地产公司财务风险影响机制研究 [J]. 金融理论与实践, 2017(12): 43-48.