

上市公司财务风险预测——基于 logistic 模型

经济学院 武子泉

指导教师 陈思好

内容提要：本文以 A 股上市公司的各项财务指标作为研究样本，选取 2015-2020 的财务数据，借鉴已有的理论模型和研究经验，引入了财务变量及非财务变量，建立了用于预测上市公司财务风险的 Logistic 回归模型，并应用倾向得分匹配方法予以改进。通过对模型进行分析，可以发现：（1）T-1 年的变量数据包含了最多的公司财务状况信息，利用其建模效果最好；（2）以盈利能力（例如净资产收益率）、成长能力（例如营业利润同比增长率）等指标为主的财务指标所构建的因子模型达到了 78% 的总体预测度，表现出了较好的预测性；（3）非财务指标变量，例如外部审计意见，对于模型的预测起到了一定的作用；（4）运用倾向得分匹配方法对样本进行 1：1 匹配可以提升模型的预测效果，模型的总体预测率相较于 1：1 随机匹配有着 5% 的提升。

关键词：上市公司 财务风险 预测模型 Logistic 回归模型 倾向得分匹配

中图分类号：F272.1

Financial Risk Prediction for Listed Companies - Based on Logistic Model

Abstract: This research takes the financial indicators of Chinese listed companies as the research samples, selects the financial data from 2015-2020, draws on the existing theoretical models and research experience, introduces financial variables and non-financial variables, and establishes a logistic regression model for predicting the financial risk of listed companies, and applies the Propensity Score Matching method to improve it. Through the analysis of the model, it can be found that: (1) the variable data of year T-1 contains the most information about the financial situation of the company, and the modeling effect is best by using it; (2) the financial indicators constructed in terms of cash flow, operating capacity, solvency, profitability and growth capacity have great predictability in the model; (3) the non-financial indicator variables play a certain role in the prediction of the model; (4) the propensity score matching method is used to predict the financial risk of the sample companies. (4) using propensity score matching method to match samples 1:1 can improve

the prediction of the model.

Keywords: Listed companies, Financial risk, Predictive model, Logistic regression model, Propensity Score Matching

JEL: C53, D29, L29

目录

一、引言	1
(一) 研究背景及意义	1
(二) 研究内容及研究方法	3
二、文献综述	6
(一) 国外学者对于财务风险预测的研究概述	6
(二) 国内学者对于财务风险预测的研究概述	8
(三) 文献简要评述	10
三、理论基础与研究设计	12
(一) 财务风险相关理论	12
(二) Logistic 模型的介绍	15
(三) 倾向得分匹配简介	16
(四) 研究设计	16
四、建模与实证分析	18
(一) 样本数据选取	18
(二) 模型变量选取	18
(三) 实证分析	20
(四) 模型解释	32
五、研究总结及展望	33
(一) 研究结论	33
(二) 研究局限	33
(三) 未来研究展望	34
参考文献	35
后记	37

一、 引言

（一） 研究背景及意义

1. 研究背景

随着我国经济的持续发展，金融市场不断扩大对外开放，经济体制改革不断深入，各类企业作为市场经济的细胞在社会主义市场经济中占据了越来越重要的地位，贡献了经济发展的诸多力量。总体来说，企业在经济发展的过程中发挥着相当重要的作用。我国经济发展、供给侧结构性改革及产业升级、科技创新及大众就业等多方面的发展都离不开各行业大型国有及中小型民营企业。一方面，企业的发展推动资源合理配置，新兴产业部门从无到有地发展壮大，不断增多的就业岗位容纳了劳动力，带来了新的经济增长点，推动市场的持续繁荣；另一方面，企业的竞争性质持续推动企业创新发展，新科学与新技术得以应用并创造实用价值，生产效率的提升促使产业调整与升级，赋予经济发展新动力。

在市场经济中，企业的最终目标是创造利润，因此财务状况的好坏是衡量企业发展健康与否的重要标准，财务健康是企业的生命线所在。我国各行各业公司数量众多，在企业的经营发展中，不可避免地存在财务造假、披露虚假信息、对外担保违规、现金流断裂及资不抵债等各类财务问题，部分企业甚至面临着可能导致破产的较大规模的财务风险。因此，公司的良性经营，健康发展是当前经济学，金融学研究的核心主题之一。

从外部来看，日趋激烈的市场竞争与外部营商环境的不确定性等问题也在威胁着企业的生存。近年来不少企业由于经营不善等其他原因，屡屡出现现金流断裂乃至发生财务风险的情况。例如，在二零二零年开始的新冠疫情中，绝大多数企业，尤以小微企业为甚，都受到了不小的冲击。由于疫情显著影响了消费，外贸及投资，市场供需出现下滑，不少企业增长乏力，甚至难以维持日常经营。根据天眼查的数据显示，2021 年我国企业注册注销比由 2019 年的 11% 下降到 4%，处于历史低点。疫情不断的反复，经济复苏有限，上游涨价，企业面临前所未有的经营压力。在这种情况下，企业很可能爆发财务危机，影响经营者与投资者的利益，乃至市场整体的平稳。如何及时发现及识别财务风险是有效避免财务风险爆发的前提。

近年来，不少研究聚焦于公司的财务风险研究上，结果表明，大多数企业陷入财

务困境通常不是突发的，而是一个渐进的过程。多数企业表现出由财务状况正常，经历经营危机后财务逐步恶化，最后发展为财务风险并爆发，最终经历重组及破产。因此，企业的财务风险从先验的角度来说是可以进行预测的。因此，对公司财务风险的预测，无论是从企业微观角度看，还是从宏观经济角度看，都有着不同寻常的必要性。

2. 研究意义

财务风险从定义上看，是指企业在各项财务活动中，由于经营过程中难以预料和突发因素，企业在一定时期、一定范围内所获取的最终财务结果与财务预期发生偏差，影响企业正常收益的可能。财务风险从定义的外延看，有广义与狭义之分。风险的狭义定义是指部分学者将风险定义为损失的不确定性这一种定义。而广义的定义指出风险在损失的不确定性之上，还包括盈利的不确定性。公司在经营活动中常常出现较为严重的财务危机与风险的情况，便可称之为发生了财务风险。企业发生破产也通常是由于财务风险的发生。财务风险可以分为市场风险、信用风险、流动性风险和操作风险等四大类。从现实经济环境来说，财务风险是客观存在的，企业不可能完全消除财务风险，只能采取财务及非财务措施将财务风险对企业的危害降到最低。因此，对公司发生财务风险的概率进行预测有重要意义。

从宏观上看，企业及公司的重要性决定了公司出现财务风险，其影响往往波及众多方面。维持企业正常的经营活动，化解企业可能隐含的风险，对政府管理部门政策制定及经济的平稳向好发展，资源的合理分配及防范整体性金融风险具有重要的意义。

从微观上看，公司对于隐含的财务风险的识别与处理，既关系到公司自身的存亡和战略发展，也关乎股东、投资者和债权人的利益。对于公司的经营管理者来说，正确预测公司的财务风险有利于全面了解公司的财务健康状况，防范财务隐患和动态调整经营策略，及早做出应对措施以避免财务风险的爆发。对于外部投资者来说，及时预测所投资企业的财务风险有助于调整投资策略，理性选择投资的公司，减小可能的损失，创造最大的经济利益。同时，及时发现财务风险也有利于监管部门实时监督上市公司运营状况，纠正不合规行为。因此，探究公司财务风险的预测有着重要的现实意义。

但由于我国市场经济仍处于上升发展阶段，企业普遍存在各类问题，例如管理结构不尽合理，财务信息疏漏，管理方式不科学，缺乏完善内控制度，风险管理薄弱等

诸多问题。因此，当内部出现经营不善或外部环境恶化时，企业没有足够的准备来应对危机。因此借用统计学模型预测企业财务风险，积极实施措施预防风险，是十分必要的。

综上所述，通过建立财务风险预测系统，企业可以在财务风险到来之前，通过模型预测其概率，以便企业可以提早采取措施，可以为企业自身减小潜在的损失，为外部投资者及监管者带来正向的外部性，同时为经济发展及市场平稳提供微观保障。

（二） 研究内容及研究方法

1. 研究内容

进行实证研究之前，首先要确定具体的研究对象。由于考虑到研究的便利性及相关财务数据的可获得性，文章选择上市公司作为研究对象。而在财务风险的界定上，文章选择 A 股上市公司中被实施“ST”及“*ST”的公司作为财务风险公司的研究样本。

一九九八年四月二十二日，沪深交易所宣布，从即日起将对出现财务状况或其它状况异常的上市公司股票采取特别处理，这类股票称为 ST 股。目前，上交所及深交所仍然沿用 ST 规则。在 ST 规则中，财务状况异常主要指以下几种情况：最近一个会计年度的审计结果显示股东权益为负值；扣除非经常性损益后的净利润为负值；最近一个会计年度的审计结果显示其股东权益低于注册资本，即每股净资产低于股票面值；注册会计师对最近一个会计年度的财务报告出具无法表示意见或否定意见的审计报告；最近一个会计年度经审计的股东权益扣除注册会计师、有关部门不予确认的部分，低于注册资本；最近一份经审计的财务报告对上年度利润进行调整，导致连续一个会计年度亏损；经交易所或中国证监会认定为财务状况异常。如果出现更为严重的情况，例如最近三年连续亏损（以最近三年年度报告披露的当年经审计净利润为依据），因财务会计报告存在重大会计差错或者虚假记载等等情况，还会被除以更加严厉的退市风险警示，即*ST 股票。对于这两类股票，上交所及深交所对其交易进行了限制，主要有股票报价日涨幅限制为 5%，跌幅限制为 5%，上市公司披露的中期报告必须审计等措施。¹

选用 ST 及*ST 公司作为财务风险公司的典型代表是有其合理性的。在我国金融市场中，由于 A 股上市制度极为严格，各类公司“借壳上市”的情况较多，真正破产

¹ 《上海证券交易所股票上市规则（2022 年 1 月修订）》及《深圳证券交易所股票上市规则（2022 年修订）》

重组较少，因此很难向外国学者一样使用破产企业作为财务风险代表公司；其次，破产公司的财务数据并不向外公布，这也是原因之一。其次，从上文可以看到，被采取 ST 特别处理的公司一般都在盈利能力上出现了不小的风险，这类持续亏损的公司往往经营不良，管理能力不足，导致其偿还外部债务能力薄弱，负债滚雪球式上升，这又进一步恶化营收，形成滚雪球式的恶性循环，最终导致财务危机发生。因此，可以认为在我国被实施 ST 的公司即是财务风险公司。因此，本文选择 ST 及*ST 公司作为财务风险公司研究对象。

在确定研究对象后，本文内容主要聚焦于在借鉴先前部分学者对企业财务风险及企业财务风险预测模型的研究经验的基础上，采用 Logistic 回归估计的统计分析模型，将各类财务及非财务指标作为回归变量，运用统计检验的方法变量进行筛选，确定模型建立的方法，通过实证研究，分别构建基于随机样本匹配和倾向得分匹配的预警模型，并通过观察，比较及改进模型，得到较高的模型预测准确度，并最终得出结论。

本文共有五个章节：

第一章：引言。这一章介绍本文的研究背景、研究意义，研究的具体对象及内容，研究所使用的主要方法及本文研究可能的创新点；

第二章：文献综述。这一章具体概况了国外学者对于财务风险预测的研究现状，及国内学者对于财务风险预测的相关研究，并在对研究的概况中提取出可以借鉴的研究经验；

第三章：理论基础及研究设计。这一章具体引入了公司财务研究领域对公司财务风险的相关研究理论，并以此作为实证统计模型的研究依据。其次介绍了文章主要使用的 Logistic 统计模型，倾向得分匹配方法的基本概念及简介；

第四章：实证分析。这一章集中进行了统计模型的建模与分析。研究首先选择及划分样本，选择变量，对样本进行显著性及相关性检验，因子分析，建模回归及回归模型分析。在考察了 1：1 样本随机匹配模型的预测准确度后，建立 1：1 倾向得分样本匹配模型，检验其是否对于模型的有效性有显著的提升效果；

第五章：研究总结及展望。这章对于实证研究的模型做了经济学上及公司财务上的解释，总结了实证模型的研究结果，并对研究不足进行了评价与展望。

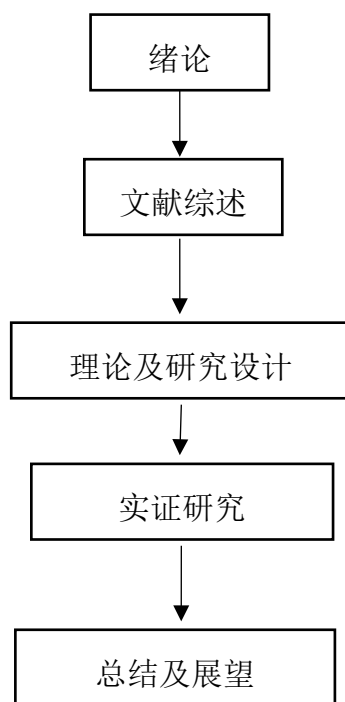


图 1.1 本文流程图

2. 研究方法

本文研究采用文献研究与实证建模研究相结合的方法。

本文首先参阅国内外相关研究，分析其建模方法，分析方法，研究可取之处及需要改进的缺陷，从而对公司财务风险预测研究的研究现状形成基本的认识框架，寻找进一步研究的方向。

其次，通过对上市公司财务数据进行建模分析，用逻辑回归分析方法建立财务风险预测模型，进行实证分析，建立一个建模完善、准确度高的统计模型。

3. 研究创新点

文章的创新点主要如下：

(1)应用倾向得分匹配方法处理原始样本，剔除行业与市值所造成的样本偏差，使模型更加有效；

(2)选择 t-1、t-2、t-3 年的样本分别进行建模，以更好地考察指标对风险的预测的前瞻性；

(3)加入非财务指标，丰富了模型的内涵，使得模型的预测更加精确。

二、 文献综述

(一) 国外学者对于财务风险预测的研究概述

1. 财务风险预测研究的初步发展

财务风险预测模型的初步研究起源于二十世纪初。二十世纪初资本主义公司制度趋于成熟，因此对公司的财务状况做进一步的精细化研究成为了现实需求。伴随着公司治理学说的不断完善和财务管理学的发展，国外学者较早开始了定量分析公司财务与使用统计模型来进行公司财务危机的研究。

Fitzpatrick(1932)首先提出了对财务风险进行预测，并做了单变量检验。Fitzpatrick 选择了 20 对匹配公司的数据，并讨论了将各类会计指标比率作为破产指标的问题。对 13 组变量进行检验后，检验结果证明净利润与股东权益的比值具有最好的识别效果。

Beaver(1966)对单变量模型进行了进一步的研究。Beaver 选取了 79 组财务风险及非风险公司作为研究样本，5 个财务指标变量为一组，共选取了 6 组变量进行回归。最终发现公司现金流/公司总负债的判别效果要优于其他变量。同期 Wilcox(1971)受到 Beaver 这项研究的启发，也同样对财务风险单变量指标预测进行了研究，得到了更深入的结论。Wilcox 对 Beaver 的模型进行了进一步研究和改进，提出净现金流入量和净现金流出量之间的相关性是一个预测财务风险的良好指标。

由此，Altman(1968)首次尝试使用多变量线性模型来预测财务风险。Altman 使用 33 组公司财务破产的公司及没有破产的健康公司作为样本，使用 33 个财务变量作为备选变量，并将其进行线性组合建立 Z-Score 模型。模型根据预测率最高的原则，选取了 EBITDA / 总资产、留存收益/总资产及主营业务收入/总资产等一系列指标来进行建模。模型的基本思想是，将企业 Z 值得分的临界值确定为 2.675，如果企业 Z 值大于 2.765，则企业的财务十分健康，出现财务问题的可能性较小；反之，企业的得分小于 2.765，企业很可能出现财务恶化的问题，需要多加关注。1977 年，Altman、Narayanan 和 Haldeman 进一步改进了 Z-Score 模型，提出了著名的 Zeta 模型。Zeta 模型使用七个财务指标作为预测变量，包括总资产、EBITDA/总资产、营收波动率、流动比率、利息保障倍数、留存收益及 PE。Zeta 模型评估简便，预测准确度高，因此在商业评估上得到了广泛的应用。

2. 财务风险预测研究的发展

由于经济的不断发展,企业的组织模式及财务结构不同于以往,Zeta 模型的预测准确率也逐年下降。除此之外,Zeta 模型还面临着统计学本身的一系列限制,例如,财务变量数据常常是未知分布的,其均值,方差等都不符合正态性的要求。因此,利用更为先进的数学模型来进行财务风险预测成为了重要的研究课题。

Martin(1977)最早应用 logistic 模型进行财务指标建模。在选取了 58 家破产的银行作为样本后,Martin 检验了 25 个财务变量的有效性,并最终选取了总资产净利率,现金流量等指标建立了 logistic 模型进行预测,其准确性优于 Z-Score 模型及 Zeta 模型。同期 Ohlson(1980)运用极大似然估计建立了 Probit 模型,研究发现资本结构、公司规模、盈利能力和发展能力四类财务变量显著影响公司发生财务风险的可能性。

Collins 和 Green(1982)对多元线性模型、Logistic 模型及其他概率模型进行了比较及总结。研究认为,Logistic 模型的有效性在面对不服从正态分布的真实样本时,其有效性是最强的。而其他模型往往会有更强的假设,因此在处理样本时不是一致估计,预测准确度相对更差。

Zmijewski(1984)建立了多元 Probit 模型,并通过研究发现多元 Probit 模型相较二元 Probit 模型其在降低样本的选择偏差上具有更好的效果。通过对样本进行更细致的分类,可以显著降低样本构建时出现的误差。随着相关研究的不断深入,Logistic 模型的可靠性逐渐被研究者所发现,成为了建模的主流选择。

Wansley(1986)第一次使用 Cox 模型进行财务风险建模。在对 130 组银行样本进行建模分析后,发现 Cox 模型也有着较好的预测效果。Bandopadhyaya(1994)使用 Cox Weibull 模型分析了企业退出破产法第 11 章保护的概率,并指出对于公司来说,其产能利用率和利息负担与公司发生财务风险的可能性呈现正相关关系。Henebry(1996)对其他学者的银行破产风险 Cox 模型进行了进一步探讨,主要聚焦于添加相关的变量是否可以完善模型的预测效果。研究发现在模型中加入净现金流量等指标,可以改善模型的预测表现。

随着机器学习及人工神经网络的发展,经济学者开始利用神经网络进行更加精细化及更加复杂的模型构建。最早引入神经网络模型的是 Odom 和 Sharda(1990)有关于公司信评的研究。研究将信用贷款数据作为样本,搭建了人工神经网络模型来分析银

行的财务风险概率,并且发现这个模型的有效性要高于传统的信评模型。Michael(1999)针对神经网络的预测有效性做了理论方面的解读。研究详细阐述了神经网络模型进行预测的数学原理、前提假设及模型衍生应用。研究还对神经网络,传统多元模型及 Logistic 模型进行了横向对比,指出神经网络模型要优于其他模型。进入二十一世纪之后,神经网络模型由于其巨大的潜力,日益成为研究的主要关注对象。

(二) 国内学者对于财务风险预测的研究概述

从总体来看,由于经济体制改革与市场经济模式在我国起步较晚,国内学者对于公司财务风险的研究也要晚于国外学者。不同于国外学者的实证模型,国内学者的模型大多用于判别企业是否存在财务风险,而不是选择破产企业进行研究,这更适合于我国国情。

周首华、杨济华、王平真(1996)最早利用 Z-Score 模型对我国的企业进行了评估。研究选用了流动比率、负债率与现金流量率等指标进行建模,在对约 4000 家样本公司进行分析后,发现流动比率和负债率的预测准确度是最高的。这是首次有学者运用多元回归分析模型的方法对我国企业财务风险进行研究。

陈晓、陈治鸿(2000)是最早应用 Logistic 模型分析我国公司财务危机的学者。他们在研究中对于 1995-1997 年三年的财务数据进行了建模,构建了由总资产周转率、营运资本 / 总资产、流动比率等财务指标所组成的 Logistic 模型。根据模型的预测结果,模型对 T-3 年的财务风险有着 73% 的预测精度,效果较好。

张玲(2000)对 A 股上市公司进行了财务风险建模研究。样本选取了 120 家公司,从营运能力、偿债能力、盈利能力、资本结构四类财务指标中筛选建模,形成了准确率较高的 Logistic 模型。模型对于 T-2 及 T-3 年的财务数据样本的预测正确度为 87% 及 70%,而对时间间隔更远的财务数据样本进行预测的表现则较差。

吴世农、卢贤义(2001)的研究对几种不同的模型进行了全面的比较分析,在这一领域的研究中具有重要的地位,被诸多学者引做参考。吴世农、卢贤义选取了一九九八年至二零零零年中的 70 家被 ST 处理的上市公司作为财务组公司,1700 家正常公司作为对照组,分别建立了 Fisher 线性分析模型、Logistic 回归模型与多元线性回归模型,并对模型效果进行了比较。研究结果表面,Logistic 模型的效果要显著好于其他两个模型。在财务变量上,模型中加入净资产收益率(Return on Equity, ROE)等

成长变量会对模型有比较好的提升效果。

张爱民(2001)运用主成分分析法,对财务风险预测模型中的变量进行了主成分分析及新变量提取,认为对模型运用主成分分析法可以显著减小模型的预测失误概率。

舒惠好(2005)运用了样本划分的方法,将样本公司按照不同所在行业进行划分分组,再进行研究。对建立的 Z-Score 模型进行分析后,研究得出按照模型分析,我国上市公司的 Z 值呈现出行业性差异。研究进一步细化了多元判别模型的研究方法。

彭大庆(2006)等利用时间序列方法,对静态的财务指标进行了动态化处理,结合主成分分析和逻辑回归分析模型等,构建出了动态预测模型。研究结果表明,动态预测模型的预测准确率在统计上显著高于静态预测模型。

王克敏、姬美光(2006)首先将非财务指标引入到财务风险建模当中。研究利用 Logistic 模型,引入了公司治理结构相关的财务变量,检验其对财务风险发生概率的影响。研究重点检验了前十大董事持股比例,公司高管数量及结构,公司代理成本,资金占用等相关指标对公司发生财务风险的影响。研究同时从财务理论方面定性阐述了公司治理结构出现问题对公司财务状况的负面影响,得出了例如“拥有国家股控股股东的公司相对其他公司,更易于出现财务风险问题”等较为新颖的结论。黄善东(2007)同样进行了公司治理结构相关指标的研究。其研究选择了两千余家上市公司作为样本,构建了逻辑回归模型。研究发现,国家参股比例等指标比较显著,加入非财务变量后,企业财务风险预警模型的回归效果和预测准确度可以发现明显上升。

陈磊(2008)以 71 家上市公司的季度公布财报数据作为研究样本,使用指数加权移动平均控制图作为研究方法进行建模,构建了一个多阶段财务风险预警模型。根据模型的实证结果,模型对上市公司财务风险具有较好的预测效果。

在同时,学者也试图探索多种模型在预测方面的应用。宋新平(2009)使用遗传算法模型、支持向量机模型,建立了优化输入变量的财务困境预测方法。于谦龙、顾晓敏(2011)建立了最小二乘支持向量机模型,对于高校的财务风险进行了专项研究。研究发现偿债能力、学生欠费状况、支付能力等 14 个指标所构建的支持向量机模型对高校的财务决策具有良好的支持作用。

随着机器学习相关研究的不断进步,越来越多新模型被用于预测当中。冯龙飞、王凤洲(2013)选用了 A 股制造业上市公司 2006-2013 年的财务数据作为研究对象,

采用 C5.0 决策树模型进行财务风险预测的实证分析。实证结果表明，盈利能力指标与现金流量在财务困境预警中信息含量最高。整体上看，模型具有很好的预测性。王智（2017）使用 Bayes 判别分析，具体分析了上市公司财务风险预测中模型的应用。王甜（2019）构建了 Probit-SVM 组合预警模型进行实证分析，证明组合预警模型相比单一模型而言拥有更高的分类准确度和稳健性，模型的预测精度也更高。

（三）文献简要评述

从公司财务风险预测模型的研究轨迹来看，模型主要经历了由单变量向多变量发展、由多元回归模型与逻辑回归模型向机器学习模型发展，由单一财务指标向加入非财务指标的方向发展这几方面的不断进化。模型逐渐由相对简洁且易于理解向专业化，复杂化的方向发展。

尽管从上述研究来看，我国对财务风险预测的研究取得了丰硕成果，但仍面临着一定的问题。我国资本市场、证券市场的发展仍处于初级阶段，与国外的较为成熟的资本市场体制仍然存在不小的差异，国外模型中某些假设的因素并不符合我国的实际情况，因此需要借鉴但无法照搬国外的研究经验。同时，财务风险预测指标体系不甚完善，指标之间往往存在着内生性与相关性。另外，影响企业发生财务风险的其他因素，例如非财务风险，并没有很清晰的研究结论。

综上所述，本文借鉴郑芳云（2013）、刘小龙（2014）及卢永艳等（2016）较为前沿的研究，主要考虑以下几个方面进行研究设计：

（1）以往文献对于样本的处理存在财务风险公司与非财务风险公司数量相差过大，如果直接使用原始样本，会造成两个问题：一是从现实角度讲，健康公司的数量和总体规模要远远大于出现财务风险的公司，两者数量差异较大，存在不平衡抽样问题；二是企业的规模与行业作为内生的变量，对于企业陷入财务风险的概率可能有较大的影响，如果直接用原始样本，可能存在控制组与处理组差异较大的问题，存在样本偏差可能会影响模型的有效性。本文针对性应用 1:1 倾向得分匹配方法处理原始样本可以有效解决这两个问题。使用 1:1 匹配控制了控制组和处理组的变量数量，可以使样本更加均匀；使用市值与行业两个变量进行相似的样本匹配，可以尽可能地剔除行业与市值所造成的样本偏差；

（2）由以往文献的整理分析可以看出，影响企业财务危机的因素不仅仅有财务因

素，还与外部整体的宏观因素，外部对企业的评价，以及企业内部的管理结构、管理成本等诸多因素有关。但目前的研究主要集中于依据财务指标变量进行建模，对非财务因素涉及较少，因此模型相对单一，缺乏新意。在指标的选择上，本文研究尝试加入非财务指标，可以使得模型的内涵更加丰富，以期模型的判别力更好，模型的预测能够更加精确。

(3)根据以往的研究表明，Logistic 回归模型有着模型简洁，效果较好的优点，是最为广泛的应用模型之一，因此本文同样应用 Logistic 回归模型。

(4)从以往的文献来看，大多数文献都集中在使用某一年的财务变量数据进行建模，没有进行纵向对比，模型说服力不足。为了考察不同年份的变量数据所形成的模型判别力是否有优劣之分，本文研究将其他针对单一年份样本数据的研究进行结合，尝试对 t-1、t-2、t-3 年的样本分别进行建模，以更好地考察不同年份的指标所含的预测信息是否有差别，从而筛选出对风险预测的预测能力最强的年份样本数据。

三、 理论基础与研究设计

(一) 财务风险相关理论

1. 财务风险的界定

财务风险从定义上看,是指企业在各项财务活动中,由于经营过程中难以预料和突发因素,企业在一定时期、一定范围内所获取的最终财务结果与财务预期发生偏差,从而形成的使企业蒙受经济损失,影响企业正常收益的可能。也就是说财务风险通常就是指企业的财务状况困难。通常情况下,发生财务风险的企业资产债务比值相对较高,现金流不足以覆盖经营支出,且往往出现收入周期与债务周期不匹配的问题。

财务风险从所界定的定义来看,有广义的定义和狭义的定义。决策理论学家把风险定义为损失的不确定性,这是风险的狭义定义。而广义的定义则认为,风险不主要是指损失的不确定性,而且还包括盈利的不确定性。公司在经营活动中常常出现较为严重的财务危机与风险的情况,便可称之为发生了财务风险。企业发生破产也通常是由于财务风险的发生。财务风险可以分为市场风险、信用风险、流动性风险和操作风险等四大类。¹

对于财务风险的具体形式,国内外有着多样的看法。Beaver(1966)对财务风险首先进行了研究,提出财务风险主要有财务企业宣告破产、财务账户入不敷出、债券违约等几种情况。Altman(1968)在他的研究中,将财务风险的爆发划分为经营失败、无偿债能力、债务整体违约和破产这四个连续的阶段,并把公司破产作为财务风险的标志。Deakin(1972)和 Lau(1987)的分类方法也大致相似。总体来看,国外学者对于财务风险的发生表现形式大致都可以归类为这几类:(1)现金不足,公司的正常经营受限,没有能力获得充足的现金;(2)流动性不足,公司的流动资产小于流动负债,企业在履行短期偿债义务的时候遇到压力;(3)资产不足,财务上长期偿债能力下降,留存收益由盈余转为赤字,总资产小于总负债;(4)上述事态逐渐发展为无法正常偿债,出现大规模欠债或贷款无法偿还,最终企业宣布破产。

国内学者对于企业财务风险的发生主要从四个角度来界定,即技术破产(无法偿还债务),会计破产(资不抵债),法定破产(进入法定破产程序),企业失败(正式破

¹ 理查德 A. 布雷利、斯图尔特 C. 迈尔斯:《公司财务原理》,方曙红译,机械工业出版社 2008 年版。

产) 这四个方面。¹在这四个阶段之下企业财务风险问题由浅入深, 具有非常高的可度量性。但是由于我国资本市场的特殊性, 大多数上市公司存续的时间不长, 证监会关于公司上市的核准与注册也采取了较为严格的监控措施, 市场上融资需求也相对较大, 这些原因使得可能出现财务风险的上市公司仍有一定的价值, 即“壳资源”。因此, 即便有上市公司进入严重财务风险面临退市风险, 也会存在借壳公司注入资产上市。因此, 我国相关研究的界定都采取被实行“特别处理”(ST)的上市公司定义为财务风险公司。

2. 财务风险的成因与管理

企业陷入财务风险的原因具有多样性, 可以产生于筹融资、投资、经营以及利润分配与再生产等多个过程。根据我国公司现实情况分析, 主要有以下几个成因理论:

(1) 经营管理问题。公司管理中很多问题容易暴露在不健全的经营管理制度下。常见的问题例如企业内部治理结构不完善, 财务报表制度出现漏洞, 内控制度不完善, 没有合规风控意识, 这些都会导致管理问题。企业的日常管理和经营决策大多依赖于经营者的管理经验, 管理水平以及以此为基础的系统管理制度。如果经营管理出现问题, 企业的日常生产, 产品研发, 市场策略制订都会出现波动, 企业盈利能力会出现下滑, 抵御市场风险的能力下降, 出现财务危机的可能性大增;

(2) 融资问题。企业融资渠道有直接融资和间接融资之分, 而融资能力的高低很大程度上影响了企业的现金流水平与财务健康。我国资本市场还处于发展阶段, 债务融资与权益融资的核准较为严格, 限制条件较多, 大部分企业难以通过这种方式进行融资。而信贷融资渠道需要担保或抵押, 融资风险相对更高。由此, 企业现金流的保证很大程度上依赖内源融资, 有一定可能导致企业出现财务上的风险;

(3) 企业生命周期。企业生命周期理论是指企业的发展与成长的动态轨迹。企业的生命周期可以划分为包括发展、成长、成熟、衰退四个阶段。在不同阶段, 企业的组织管理形式, 产品生产策略, 发展模式等都有不同的特点, 每个阶段企业的盈利能力、偿债能力, 现金流水平等也各有不同。在企业的发展与成长阶段, 企业的资金、技术储备与人力资源相对缺乏, 负债率较高, 流动资产与流动负债比相对较小, 容易发生财务风险。而企业的生命周期也与宏观经济周期相关。例如在“美林时钟”的“经

¹ 赵国忠:《上市公司财务困境研究》, 北京大学出版社 2009 年版。

济下行，通胀上行”阶段，经济疲软对企业盈利会造成不小的冲击，产业政策的收紧与融资环境的紧缩也会对企业的财务状况造成威胁。例如在历次金融危机中，企业发生财务危机乃至破产的概率大大上升。

（4）其他原因，例如企业投资失败，闲置资金利用效率低下，资产流动性不足等原因，也会造成企业的资本运作不畅通，盈利下降、经营效益低下、净现金流量下滑，这时企业债务问题如果集中爆发，没有足够的资金用于偿债，若不能采取适当措施加以补救，便会导致财务风险。

综上所述，可以看出，企业财务风险的管理是企业管理中一项重要的课题。财务风险管理是指经营主体对其经营过程中存在的各种财务风险进行识别、度量和评价的行为。财务风险管理主要由风险识别、风险度量和风险控制等环节组成的，其中核心是风险的识别与度量问题，是为了达成目标是降低财务风险，减少风险损失的核心目标。¹

在财务风险管理中，财务风险的识别与度量是一个企业避免财务风险的重中之重。其中，风险识别是指在财务出现问题之前，对相关财务风险所作的定性判别。在市场经济环境中，财务风险并不都具有可预见性。在财务风险识别阶段，相关信息的收集、分类方法，风险的合理处理程序，风险的汇总、辨识，风险发展走向的监测都是必要的。风险识别是风险管理程序的基础。而风险度量更为核心，决定了企业对风险的策略方向和决策结果。通常情况下的风险度量都指可以运用数学方法，统计模型进行计算，并通过相关指标的大小判断风险发生的可能性与风险损失大小的系统处理方法。

从上述的角度来看，企业财务风险的识别与度量，是一个渐进的阶段。在这一阶段中伴随着盈利减少、入不敷出、大规模债务违约等现象，企业的财务健康状况不断恶化，最终陷入无法挽回的困境中。从上文的研究，可以看出，财务风险的识别与度量，可以以不同阶段企业财务的不同财务问题特征，来选择适宜的风险识别与度量指标。例如，在企业财务风险初期，企业盈利能力下降，现金流入无法满足资金需求，此时选择盈利指标，例如企业净现金流量等，可以较好进行风险识别；同时，对于企业债务的高低，风险识别也需要选择相应的指标，例如现金债务比。在财务风险继续发展的阶段，企业正常营运受阻，资产周转不畅，已经进入了可能资不抵债的状况，

¹ 张继德：《企业财务风险管理》，经济科学出版社 2015 年版

此时资产周转状况的衡量指标同样可以纳入到风险识别与度量的模型当中。除此之外，对企业的成长能力的衡量同样可以在企业财务风险的早期识别中起作用。例如当企业的营业收入同比增长率或利润同比增长率逐渐下降，表明企业很可能在后续没有足够的资金去进行研发与再生产，相对于其他企业陷入财务风险的可能性更大。

综上所述，对于财务风险的正常界定，识别与管理是财务风险预测模型的理论基础，同时也是模型设计所要考虑的重要因素。下文中的研究设计与财务指标的选择，分析也建立在这个理论基础上。

(二) Logistic 模型的介绍

Logistic 回归模型是一种对数几率模型，又被称为逻辑模型、评定模型、分类评定模型。Logistic 回归模型属于离散选择法模型之一，隶属于多元变量分析范畴，是社会学、生物统计学、临床、数量心理学、计量经济学、市场营销等统计实证分析的常用方法。Logistic 回归模型主要用于说明，描述和推断两分类或多分类的因变量与一组自变量的量化关系。该模型用于研究时，具有诸多的优点。首先，Logistic 回归模型属于概率型回归，在回归中的被解释变量和解释变量之间的关系不一定要是线性的，与现实情况更加契合。其次，Logistic 回归不需要方差不变的前提，自变量之间也不必假设符合正态分布，符合实际数据的情形。Logistic 回归的因变量是分类变量，表明不同事件发生的可能性大小，因此可以较好地应用于事件发生可能性的度量。

遵从上述的理由，本文选用 Logistic 回归模型作为建模所用的模型，利用所选的财务指标与非财务指标作为自变量，企业财务风险状态作为因变量。模型假设企业发生财务风险的概率为 P （发生财务风险因变量取 1，没有发生财务风险因变量取 0）。同时，假定反映 P 在 0, 1 之间变化幅度的 $F(P)=\ln[P/(1-P)]$ 可以用财务变量的线性组合来表示，可以得到 Logistic 回归模型的表达式：

$$Z_i = \ln\left(\frac{P_i}{1-P_i}\right) = \alpha_i + \sum_{j=1}^n \beta_j X_{ij} + \varepsilon_i$$

即可以写作概率函数形式：

$$P_i = \frac{1}{1 + e^{-Z_i}}$$

其中 X_{ij} 为所选择的财务指标与非财务指标， β_j 为待估计的参数， P_i 为在条件 $(X_{i1}, X_{i2}, X_{i3}, \dots, X_{in})$ 下企业发生财务风险的概率。其中，如果设定某一个临界值 P^*

作为判断事件发生的标准，并比较 P_i 与临界值 P^* ，便可以对事件发生的概率，即企业发生财务风险的概率进行预测。这里运用主流方法选择 $P^*=0.5$ 。

（三） 倾向得分匹配简介

倾向得分匹配（Propensity Score Matching，简称 PSM）是一种统计学方法，用于处理观察研究时所用样本的数据。在观察研究中，由于现实中种种原因，数据偏差和混杂变量较多，样本中可能存在较多没有考虑到的因素影响，其可能会有内生性等问题。倾向评分匹配的方法正是为了减少这些偏差和混杂变量的影响，以便对实验组和对照组进行更合理的比较，防止样本选择偏差的影响。

倾向得分匹配方法的基本思想是依据每个样本的“倾向得分值”去匹配相似的样本，使得样本间能够尽量“相近”，从而尽可能去除内生变量对于样本的影响。其基本方法是：

- （1）采用 Logistic 回归，选择适当的混杂因素作为自变量，计算倾向得分值；
- （2）得分匹配，常常采用最邻近匹配、半径匹配、核匹配等方法，来对样本进行配对；
- （3）基于新样本进行多元分析。

本文尝试利用倾向得分匹配，选用公司市值与行业进行倾向得分匹配，减少样本选择偏差的影响。

（四） 研究设计

（1）根据所制订的研究计划与研究目标，选择 A 股公司财务数据作为样本，并提取适当的财务指标与非财务指标作为研究变量；

（2）对样本数据进行摘要统计及简单描述，掌握数据的大致分布情况，并对其进行简单分析；

（3）对整理好的数据进行变量的独立性与相关性检验，剔除掉不显著的及相关性强的变量；

（4）对整理后的样本变量数据进行因子分析。首先进行 KMO 检验和 Bartlett's 球状检验以检验是否适合因子分析；如可以进行因子分析则提取公共因子，将公共因子作为最后进入 logistic 回归模型的自变量；

（5）构建 Logistic 回归模型，对模型的显著性等做出评价，考察模型的效果，并检

验模型的预测效果是否达到理想状态；

(6)使用倾向得分匹配方法对样本进行匹配，并再次进行 Logistic 回归，观察回归效果并与随机匹配进行对比，检验所构建模型的预测能力；

(7)对模型进行总结，得出结论。

四、 建模与实证分析

（一） 样本数据选取

本文研究数据选取自 Wind 股票数据库，研究所用样本上市公司主要基于在中国 A 股市场主板及创业板的上市公司。选取 A 股上市公司，一方面便于数据搜集以及整理，另一方面，选取中国本土上市公司可以更好地反映在中国当前环境下的公司财务风险现状，对现实更具有指导意义。因此，样本的选取主要依据以下原则：

(1)上市公司财务数据选用公司年报所公布数据，样本区间为 2015 年至 2021 年；

(2)剔除期间退市，即数据不完整连贯的公司；

(3)ST 及*ST 股票选取上，选取“财务状况出现异常”的公司作为样本，剔除“因其它状况出现异常”的公司财务数据；

公司财务数据进行整理后，共有样本 2354 个。在进行研究时，将研究样本分为两组，即财务风险组和非财务风险组。财务风险组共有样本 119 个，非财务风险组共有样本 3395 个。在进行数据集划分时，将所有 18-19 年被实施 ST 及*ST 的财务风险公司及对应匹配的正常公司作为样本内测试集，20-21 年被实施 ST 及*ST 的财务风险公司及对应匹配的正常公司作为样本外验证集。

（二） 模型变量选取

由于本文研究的对象主要为经营连续两年亏损，被进行特别处理的 ST 股票及连续三年亏损，被退市警示的*ST 股票，因此变量需要有足够的显著性来体现企业的财务特征。从其他学者进行研究的经验来看，统计模型所选取的财务变量需要满足以下条件：

(1) 变量指标较为简洁，可获得性强。在诸多财务变量中，能够反映上市公司财务状况的指标可以从各个方面获得，但需要进行筛选剔除难以取得数据、可计量性不足的变量。

(2) 变量需要从多个角度反映公司的整体财务状况。从现代公司财务理论来看，企业的财务状况主要取决于企业的现金流量、营运能力、偿债能力、盈利能力和成长能力等等。通常来讲，企业所遭遇的财务风险，都可以归结为“营运能力不足，持续亏损无法盈利”、“现金流断裂”、“无法偿还债务”等几类常见原因。因此，通过分析这几类指标，公司的财务状况可以得到清晰，整体的反馈，财务风险发生的可能性也

能够最大限度地得到预示。

综上所述，通过结合财务理论分析及参考部分学者的研究思路，本文共选取以下 16 个财务指标作为备选财务指标。

表 4.1 备选财务指标

指标分类	指标名称	符号	指标含义
盈利指标	净资产收益率 ROE	A1	净利润/平均净资产
	总资产报酬率 ROA	A2	经营收益/总资产
	销售净利率	A3	净利润/销售收入
	EBITDA/营业总收入	A4	EBITDA（反推法）/营业总收入
偿债指标	流动比率	B1	流动资产/流动负债
	速动比率	B2	流动资产中扣除存货部分/流动负债
	现金到期债务比	B3	经营活动产生的现金流量净额 / （短期借款+一年内到期的非流动负债+应付票据）
	资产负债率	B4	负债合计/归属母公司股东权益
营运指标	应收账款周转率	C1	应收账款转为现金的平均次数
	营运资本周转率	C2	一年中营运资本周转的次数
	存货周转率	C3	营业成本 / [(期初存货净额+期末存货净额) / 2]
	总资产周转率	C4	营业总收入 / [(期初资产总计+期末资产总计) / 2]
现金流量指标	经营活动产生现金流量净额/营业收入	D1	经营活动产生的现金流量净额(TTM) / 营业收入(TTM)
	现金营运指数	D2	会计收益和现金净流量的比例关系
	净利润现金含量	D3	经营活动产生的现金流量净额/归属于母公司所有者的净利润
成长指标	营业收入同比增长率	E1	营业收入相对上年同期增长百分比
	营业利润同比增长率	E2	营业利润相对上年同期增长百分比
	经营活动产生的现金流量净额同比增长率	E3	经营活动产生的现金流量净额相对上年同期增长百分比
	净资产收益率同比增长率	E4	ROE 相对上年同期增长百分比
	总资产同比增长率	E5	总资产相对上年同期增长百分比

数据来源：Wind 数据库

现代公司财务理论认为，公司的财务健康状况除了使用传统财务报表中可见的指标去衡量，也可以使用一些非财务指标来衡量。虽然其与财务数据并不直接相关，可测量性比财务指标差，但可以综合反映企业微观层面的经济状况。借鉴其他学者的研究思路，本文研究在选取非财务指标方面主要考虑了以下两个方面：

(1) 审计意见。在公司管理中，外部审计是一项较为重要的监督制度。外部注册会计师对企业财报的可信与否出具的意见即审计意见，是企业重要的外部监控之一。专

业注册会计师能良好地判别企业当前的财务状况并出具审计意见，因此审计意见类型对企业财务危机的发生与否具有一定的预测能力。审计意见主要有四种：标准无保留意见，带强调事项的无保留意见，保留意见及无法表示意见。本文将审计意见引入为虚拟变量，标准无保留意见归类为 0，其他归类为 1；

(2)管理成本。按照代理成本理论的定义，管理成本指委托人（股东）与代理人（经理）之间的利益冲突。具体内容有委托人的监督成本、代理人费用和剩余价值的损失等等。管理成本是公司资本结构的重要影响因素，资本结构的选择使得公司代理成本可以影响公司的治理与营运，进而影响财务状况。例如，公司管理费用较高的话，通常说明公司代理机制不合理，可能存在代理冲突，公司有发生财务风险的可能性。

综上所述，本文选取以下 2 个非财务指标作为备选指标。

表 4.2 备选非财务指标

指标名称	符号	指标含义
审计意见类型	F1	虚拟变量，0 为标准无保留意见，1 为其他
管理费用率	F2	管理费用/营业总收入

数据来源：Wind 数据库

（三） 实证分析

1. 数据摘要统计

首先对全样本下的公司财务数据进行摘要统计，以 T-1 年为例：

表 4.3 T-1 年变量数据摘要统计

Variable	Mean	Std. Dev.	Min	Max
A1	-82.8023	1240.445	-19098.3	871.5032
A2	-0.77825	17.02133	-132.408	48.4201
A3	-28.3946	255.3104	-3697.59	78.1504
A4	-17.3546	251.8668	-3640.63	128.216
B1	1.455432	0.944391	0.0318	8.4235
B2	1.120577	0.847979	0.0177	8.1255
B3	547.2057	7460.165	-3458.31	114904.5
B4	4.046419	14.52093	-30.4603	169.1467
C1	28.77644	135.3018	0	1507.334
C2	12.74083	62.8261	0	771.8677

<i>C3</i>	100.533	1321.962	0	20375.67
<i>C4</i>	0.721124	0.63556	0.0264	4.1899
<i>D1</i>	-4.01349	63.78502	-691.78	212.9
<i>D2</i>	0.395901	7.728928	-50.2782	59.791
<i>D3</i>	-183.182	2818.374	-42661.8	2476.541
<i>E1</i>	106.9143	793.8117	-95.3288	8947.864
<i>E2</i>	-212.835	1244.941	-14499.8	2738.359
<i>E3</i>	64.1684	2566.912	-27077.9	24885.28
<i>E4</i>	-5419.51	58738.31	-799123	420.3995
<i>E5</i>	21.26397	72.49882	-77.0804	864.2326
<i>F1</i>	0.159664	0.367066	0	1
<i>F2</i>	16.58058	47.88895	0.2624	664.8727

从上面数据可以看出，公司之间的部分变量所代表的财务状况差距较大，并反映在了变量中。例如，变量 B3（现金到期债务比）、变量 E3（经营活动产生的现金流量净额同比增长率）都有较大的标准差，说明不同公司之间偿债能力与经营成长能力有者较大的差距。与之相反的是部分指标公司之间的差距则较小，例如 C4 变量（总资产周转率）标准差只有 0.63，表明公司之间总资产周转速度没有较大的差距。同时，也可以看出各项数据在量级上存在差距，因此首先需要对数据进行标准化处理。

2. 显著性检验

对于服从正态分布的变量，可以利用 T 检验来检验其显著性；而对于位未知分布的变量来说则不适用。由于本文研究所使用的变量均未知其分布，所以采用 Kruskal-Wallis 秩检验的方法。Kruskal-Wallis 秩检验的基本思想是检验 n 组样本数据集中每个样本的秩及组间的平均秩次，由此判断是否可以认为 n 组数据出自同一样本。因此，使用 Kruskal-Wallis 秩检验可以判断模型选取的各项变量在 ST 组及非 ST 组间是否存在显著性差异，从而做到变量的清洗筛选。

表 4.4 T-1 年变量进行 Kruskal-Wallis 秩检验结果

变量	<i>chi-squared</i>	<i>probability</i>
<i>A1</i>	69.975	0.0001
<i>A2</i>	68.253	0.0001
<i>A3</i>	48.186	0.0001
<i>A4</i>	28.85	0.0001

<i>B1</i>	6.601	0.0102
<i>B2</i>	7.511	0.0061
<i>B3</i>	9.659	0.0019
<i>B4</i>	2.269	0.132
<i>C1</i>	3.57	0.0588
<i>C2</i>	38.74	0.0001
<i>C3</i>	11.177	0.0008
<i>C4</i>	48.894	0.0001
<i>D1</i>	2.92	0.13
<i>D2</i>	0.16	0.689
<i>D3</i>	6.099	0.0135
<i>E1</i>	20.842	0.0001
<i>E2</i>	35.215	0.0001
<i>E3</i>	0.193	0.6602
<i>E4</i>	33.93	0.0001
<i>E5</i>	40.35	0.0001
<i>F1</i>	18.12	0.0001
<i>F2</i>	18.911	0.0001

从上表可以看出，在 $P=0.05$ 的显著性水平下，大多数 T-1 年的变量通过 Kruskal-Wallis 秩检验，表明其在财务风险组与非财务风险组有着明显的差异。而 E3（经营活动产生的现金流量净额同比增长率）、D2（现金营运指数）、D1（经营活动产生的现金流量净额/营业收入）、B4（资产负债率）、C1（应收账款周转率）没有通过显著性检验，说明这 5 个指标在不同公司间的差异并不能对公司是否发生财务风险起到预示作用，因此在模型中予以剔除。

表 4.5 T-2 年变量进行 Kruskal-Wallis 秩检验结果

变量	<i>chi-squared</i>	<i>probability</i>
<i>A1</i>	44.344	0.0001
<i>A2</i>	42.852	0.0001
<i>A3</i>	18.813	0.0001
<i>A4</i>	2.132	0.1442
<i>B1</i>	0.009	0.9228
<i>B2</i>	1.423	0.0239
<i>B3</i>	18.005	0.0001
<i>B4</i>	0.009	0.9228
<i>C1</i>	1.78	0.1822
<i>C2</i>	69.865	0.0001
<i>C3</i>	8.24	0.0041
<i>C4</i>	36.762	0.0001
<i>D1</i>	7.735	0.0054
<i>D2</i>	5.786	0.0162
<i>D3</i>	6.611	0.0101

<i>E1</i>	0.627	0.4285
<i>E2</i>	4.357	0.0369
<i>E3</i>	0.815	0.3666
<i>E4</i>	1.882	0.1701
<i>E5</i>	15.097	0.0001
<i>F1</i>	7.845	0.0001
<i>F2</i>	8.327	0.0039

分析同上，在 $P=0.05$ 的显著性水平下，大多数 T-2 年的变量通过 Kruskal-Wallis 秩检验，表明其在财务风险组与非财务风险组有着明显的差异。而 E4、E3、E1、C1、B4、B1、A4 这 7 个指标没有通过显著性检验，予以剔除。

表 4.6 T-3 年变量进行 Kruskal-Wallis 秩检验结果

变量	<i>chi-squared</i>	<i>probability</i>
<i>A1</i>	47.729	0.0001
<i>A2</i>	38.366	0.0001
<i>A3</i>	18.683	0.0001
<i>A4</i>	0.19	0.6629
<i>B1</i>	1.686	0.1942
<i>B2</i>	0.123	0.7262
<i>B3</i>	18.969	0.0001
<i>B4</i>	3.461	0.0628
<i>C1</i>	3.563	0.0591
<i>C2</i>	49.184	0.0001
<i>C3</i>	15389	0.0001
<i>C4</i>	46.04	0.0001
<i>D1</i>	11.835	0.0006
<i>D2</i>	7.134	0.0076
<i>D3</i>	3.656	0.0559
<i>E1</i>	12.115	0.0005
<i>E2</i>	6.974	0.0083
<i>E3</i>	2.045	0.1527
<i>E4</i>	3.51	0.061
<i>E5</i>	7.327	0.0068
<i>F1</i>	2.824	0.0928
<i>F2</i>	8.447	0.0037

分析同上，在 $P=0.05$ 的显著性水平下，大多数 T-3 年的变量通过 Kruskal-Wallis 秩检验，表明其在财务风险组与非财务风险组有着明显的差异。而 F1、E4、E3、D3、C1、B4、B2、B1、A4 这 8 个指标没有通过显著性检验，予以剔除。

从三年的各个变量显著性检验也可以看出，随着时间接近公司发生财务危机的 T 年，ST 组与非 ST 组之间财务状况的差异不断扩大，表现即为 T-1 年卡方值均要大于前两年，没有显著性的指标数量也有所下降。同时也可以看出财务风险公司与非财务

风险公司在盈利能力与偿债能力的差距要更为显著，表明公司在出现财务危机之前，很可能已经陷入入不敷出的财务陷阱中，随着公司财务的不断恶化，并最终导致危机。

综上所述，通过显著性检验，在 T-1、T-2、T-3 的样本中分别剔除了不显著的变量，并留下显著的变量进入模型做进一步的检验。

3. 因子分析

由于变量选择自公司财务指标中特定的方面，因此很可能存在多重共线性。为了避免变量之间存在较大的相关性从而影响统计结果，因此可以对变量进行因子分析。因子分析的基本思想是从分为不同组的变量中提取公共因子，使因子可以代表较多的同组变量的信息。一方面因子分析可以很好地避免变量之间较大的相关性造成统计结果的失真，另一方面也可以精简模型，减小数据量。

在进行因子分析之前，首先对变量进行 KMO 检验和 Bartlett's 球状检验，保证同组变量之间存在一定的相关性，可以提取有效的公共因子。KMO 检验用于检查变量间的相关性和偏相关性。KMO 统计量取值在 0~1 之间，KMO 统计量在 0.6 以上时效果比较好；当 KMO 统计量在 0.5 以下时则表明因子分析不适用。Bartlett's 球状检验用于检验变量的相关系数矩阵是否是单位阵。如果检验概率 P 值小于某一显著性水平，则可以认为变量的相关系数矩阵不是一个单位阵；反之相关系数矩阵可能是一个单位阵，不适合使用因子分析法。

表 4.7 T-1、T-2 与 T-3 年变量 KMO 检验和 Bartlett's 球状检验结果

Determinant of the correlation matrix			
T-1	Det = 0.000		
T-2	Det = 0.000		
T-3	Det = 0.000		
Bartlett test of sphericity			
	Chi-square	Degrees of freedom	p-value
T-1	4591.498	136	0.000
T-2	4776.601	105	0.000
T-3	2215.467	78	0.000
Kaiser-Meyer-Olkin Measure of Sampling Adequacy			
T-1	KMO = 0.657		
T-2	KMO = 0.666		
T-3	KMO = 0.621		

说明：H0 of Bartlett test: variables are not intercorrelated

从以上分析可以看出，三年的数据 KMO 统计量均大于 0.6，均通过了 KMO 检验，同时也拒绝了 Bartlett's 球状检验的零假设。因此两项检验都可以支持数据进行因子分析。

首先对 T-1 年的数据进行因子分析，得到如下的因子特征值及贡献率。

表 4.8 T-1 年因子贡献率

<i>Factor</i>	<i>Eigenvalue</i>	<i>Difference</i>	<i>Proportion</i>	<i>Cumulative</i>
<i>Factor1</i>	4.51385	2.55371	0.4583	0.4583
<i>Factor2</i>	1.96014	0.16923	0.199	0.6573
<i>Factor3</i>	1.79091	0.93158	0.1818	0.8391
<i>Factor4</i>	0.85933	0.3178	0.0872	0.9263
<i>Factor5</i>	0.54153	0.26149	0.055	0.9813
<i>Factor6</i>	0.28003	0.12442	0.0284	1.0097
<i>Factor7</i>	0.15562	0.04104	0.0158	1.0255
<i>Factor8</i>	0.11457	0.05564	0.0116	1.0372
<i>Factor9</i>	0.05893	0.03962	0.006	1.0431
<i>Factor10</i>	0.01931	0.02046	0.002	1.0451
<i>Factor11</i>	-0.00114	0.00315	-0.0001	1.045
<i>Factor12</i>	-0.0043	0.01873	-0.0004	1.0445
<i>Factor13</i>	-0.02302	0.04043	-0.0023	1.0422
<i>Factor14</i>	-0.06345	0.01319	-0.0064	1.0358
<i>Factor15</i>	-0.07664	0.0203	-0.0078	1.028
<i>Factor16</i>	-0.09693	0.08177	-0.0098	1.0181
<i>Factor17</i>	-0.1787	0	-0.0181	1

可以看出，前四个因子的特征值大于 0.8，累计贡献率达到了 92%。特就是说，若选取前 4 个因子进行建模，可以包含原有 17 个财务变量 92%的信息量。如果按照本文设定的 90%的提取原则，前 4 个因子可以纳入到模型当中。因此综合考虑，提取前 4 个因子作为新变量，命名为 Factor1-4。为了对因子所包含的具体信息做出解释，采用正交旋转法，得到如下的因子载荷矩阵。

表 4.9 T-1 年因子载荷矩阵

<i>Variable</i>	<i>Factor1</i>	<i>Factor2</i>	<i>Factor3</i>	<i>Factor4</i>
<i>A1</i>	0.4671	0.6883	-0.5456	-0.0876
<i>A2</i>	0.7456	0.1744	-0.0182	0.401
<i>A3</i>	0.9578	-0.2159	0.0038	-0.1507
<i>A4</i>	0.9532	-0.2206	-0.0069	-0.1531
<i>B1</i>	0.1671	0.5698	0.7386	-0.112
<i>B2</i>	0.1819	0.5396	0.7475	-0.1424
<i>B3</i>	-0.4443	-0.6784	0.5706	0.1163
<i>C2</i>	0.0401	-0.0341	-0.0549	0.0982
<i>C3</i>	-0.0169	-0.0335	-0.0441	-0.0861

<i>C4</i>	0.1903	0.0496	0.0503	0.1925
<i>D3</i>	-0.0154	-0.0262	-0.0373	-0.0106
<i>E1</i>	-0.02	0.043	-0.018	0.3029
<i>E2</i>	0.2623	0.0167	0.1241	0.461
<i>E4</i>	0.8008	-0.4199	0.1366	-0.0201
<i>E5</i>	0.1557	0.0017	0.0107	0.2766
<i>F1</i>	-0.3637	-0.1267	-0.1123	-0.3453
<i>F2</i>	-0.866	0.2909	-0.0831	0.2091

从上图的因子载荷矩阵可以看出：

(1)因子 Factor1 在 A1、A2、A3、A4 变量上载荷均大于 0.4，载荷较大。A1 等指标都为盈利指标变量相关，因此该因子可以解释为主要包含了公司盈利能力的信息；

(2)因子 Factor2 在 A1、B1、B2 上载荷较大。该因子可以解释为包含了公司一定偿债能力的信息；

(3)因子 Factor3 在 B3、E2、E4 上载荷较大。E2 与 E4 为公司成长能力指标，说明该因子可以反映公司未来成长能力方面的较多信息；

(4)因子 Factor4 在 E5 以及 F2 上载荷较大，说明该因子可以反映公司总资产同比增长率以及管理费由等财务指标的信息。

对 T-2，T-3 年的变量同样进行因子分析，结果如下表。

表 4.10 T-2 与 T-3 年因子贡献率

	<i>Eigenvalue</i>		<i>Difference</i>		<i>Proportion</i>		<i>Cumulative</i>	
	<i>T-2</i>	<i>T-3</i>	<i>T-2</i>	<i>T-3</i>	<i>T-2</i>	<i>T-3</i>	<i>T-2</i>	<i>T-3</i>
<i>Factor1</i>	4.05114	3.06076	2.79497	1.72475	0.6069	0.5864	0.6069	0.5864
<i>Factor2</i>	1.25617	1.33602	0.12356	0.59137	0.1882	0.256	0.795	0.8424
<i>Factor3</i>	1.13261	0.74464	0.52516	0.23959	0.1697	0.1427	0.9647	0.9851
<i>Factor4</i>	0.60745	0.50506	0.33303	0.34925	0.091	0.0968	1.0557	1.0818
<i>Factor5</i>	0.27441	0.1558	0.19012	0.0955	0.0411	0.0299	1.0968	1.1117
<i>Factor6</i>	0.0843	0.0603	0.07066	0.03672	0.0126	0.0116	1.1094	1.1232
<i>Factor7</i>	0.01364	0.02358	0.01373	0.01897	0.002	0.0045	1.1115	1.1278
<i>Factor8</i>	-0.0001	0.00461	0.00011	0.01146	0	0.0009	1.1115	1.1286
<i>Factor9</i>	-0.00021	-0.00684	0.00112	0.04184	0	-0.0013	1.1114	1.1273
<i>Factor10</i>	-0.00133	-0.04868	0.04564	0.07752	-0.0002	-0.0093	1.1112	1.118
<i>Factor11</i>	-0.04697	-0.1262	0.05093	0.10855	-0.007	-0.0242	1.1042	1.0938
<i>Factor12</i>	-0.0979	-0.23476	0.03867	0.0202	-0.0147	-0.045	1.0895	1.0488
<i>Factor13</i>	-0.13657	-0.25496	0.09007	0.0000	-0.0205	-0.0488	1.0691	1
<i>Factor14</i>	-0.22664	-	0.00779	-	-0.034	-	1.0351	-
<i>Factor15</i>	-0.23444	-	0.0000	-	-0.0351	-	1	-

表 4.11 T-2 与 T-3 年因子载荷矩阵

T-2 Variable	Factor1		Factor2		Factor3		T-3 Variable
	T-2	T-3	T-2	T-3	T-2	T-3	
A1	0.0442	0.0365	0.0302	0.1637	0.4417	0.1782	A1
A2	0.237	0.2932	0.0596	0.5078	0.7124	0.4009	A2
A3	0.9984	0.9953	-0.023	-0.0331	-0.0527	-0.0001	A3
B2	0.0567	0.0656	-0.0352	0.1201	0.1137	0.1429	B3
B3	0.9964	0.0291	-0.0274	0.0632	-0.0555	0.0616	C2
C2	0.0203	0.0259	0.0886	0.0451	0.0898	0.0609	C3
C3	0.0161	0.1141	0.7589	0.1049	-0.1332	0.1802	C4
C4	0.0901	0.9868	0.7481	-0.0683	0.136	-0.0378	D1
D1	0.9964	-0.0126	-0.0251	-0.584	-0.0552	0.4324	D2
D2	-0.0076	-0.1159	-0.2888	0.2999	0.1452	0.3734	E1
D3	0.0078	0.0463	-0.0645	0.728	0.0717	-0.2801	E2
E2	0.0102	-0.0027	0.0585	0.2146	0.1986	0.2866	E5
E5	0.0653	-0.9871	0.0528	0.0842	0.2965	0.0583	F2
F1	-0.0025		0.1034		-0.4537		
F2	-0.9971		0.0251		0.0661		

可以看出, T-2 与 T-3 年的数据在满足 90%因子贡献率的阈值上, 均取了前 3 个因子进入模型。因子信息分析过程同上, 因此不再赘述。

4. 基于 1: 1 随机匹配样本下 logistic 回归

对财务风险组与非财务风险组进行 1: 1 随机匹配, 得到用于回归的样本后, 对 T-3、T-2、T-1 三年的数据, 依照截面数据处理方式分别进行建模, 得到了如下的 Logistic 模型。

表 4.12 T-1 年 Logistic 模型

Logistic regression		Number of obs =		238		
		LR chi2(4) =		108.67		
		Prob > chi2 =		0.0000		
Log likelihood =		-110.63541		Pseudo R2 =		0.3594
Y	Coef.	Std. Err.	z	P>z	[95% Conf. Interval]	
Factor1	-4.37987	2.808379	-1.56	0.019	-9.88419	1.124456
Factor2	-0.06972	2.831601	-0.02	0.098	-5.61956	5.480118
Factor3	0.023202	2.250509	0.01	0.592	-4.38772	4.434119
Factor4	-2.39037	0.813406	-2.94	0.003	-3.98462	-0.79613
_cons	0.711693	0.438117	1.62	0.104	-0.147	1.570386

表 4.13 T-2 年 Logistic 模型

Logistic regression				Number of obs =		238
				LR chi2(4) =		71.33
				Prob > chi2 =		0.0000
Log likelihood =		-129.306		Pseudo R2 =		0.2762
Y	Coef.	Std. Err.	z	P>z	[95% Conf. Interval]	
Factor1	4.729958	7.955727	0.59	0.0752	-10.863	20.3229
Factor2	-0.43716	0.282341	-1.55	0.122	-0.99054	0.116219
Factor3	-2.64834	0.680793	-3.89	0.000	-3.98267	-1.31401
_cons	-0.12108	0.561877	-0.22	0.829	-1.22234	0.980176

表 4.14 T-3 年 Logistic 模型

Logistic regression				Number of obs =		238
				LR chi2(4) =		16.53
				Prob > chi2 =		0.0009
Log likelihood =		-156.70652		Pseudo R2 =		0.1901
Y	Coef.	Std. Err.	z	P>z	[95% Conf. Interval]	
Factor1	0.079451	0.249719	0.32	0.12	-0.40999	0.568891
Factor2	-0.73947	0.311526	-2.37	0.018	-1.35005	-0.1289
Factor3	-0.30979	0.32533	-0.95	0.341	-0.94742	0.327846
_cons	0.022948	0.136898	0.17	0.867	-0.24537	0.291262

从三张表所显示的回归结果来看，三个模型都达到了一定的预测效果。

(1) 从回归模型的 P 值来看，三个模型的 P 值都均小于 0.01，即在 99% 的水平上显著。从 Log likelihood 来看，T-1 年的似然函数绝对值最小，为-110，表明其效果好于 T-2 年及 T-3 年模型。LR 统计量上，T-1 年模型的 LR 统计量也大于另两个模型。回归结果的伪 R 平方值（Pseudo R²）上看，T-1 年的伪 R 平方值达到 0.35，说明模型四个变量能够解释公司财务危机 35% 的发生原因，同样大于 T-2 及 T-3 年的模型；

(2) T-1 年模型的四个变量（不包含常数项）都在 P=0.05 的水平上显著，说明四个因子对财务风险的解释能力都较强；

综合以上的数据分析，为了达到最理想的拟合效果，可以依照 T-1 年数据建立 Logistic 模型，如下：

$$\ln \frac{p}{1-p} = -4.3799Factor1 - 0.0697Factor2 + 0.0232Factor3 - 2.3904Factor4 + 0.7117$$

根据 T-1 年样本建立的 Logistic 模型，可以计算出下表的 Logistic 模型预测准确度：

表 4.15 T-1 年 Logistic 模型预测准确度

TRUE			
Classified	D	~D	Total
+	87	16	103
-	32	103	135
Total	119	119	238
Classified + if predicted $\Pr(D) \geq .5$			
True D defined as $Y \neq 0$			
Sensitivity	$\Pr(+ D)$		73.11%
Specificity	$\Pr(- \sim D)$		86.55%
Positive predictive value	$\Pr(D +)$		84.47%
Negative predictive value	$\Pr(\sim D -)$		76.30%
FALSE + rate for TRUE ~D	$\Pr(+ \sim D)$		13.45%
FALSE - rate for TRUE D	$\Pr(- D)$		26.89%
FALSE + rate for classified +	$\Pr(\sim D +)$		15.53%
FALSE - rate for classified -	$\Pr(D -)$		23.70%
Correctly classified			77.83%

从上表数据可以看出，模型对于该样本的整体预测准确度达到了约 80%。其中，对于财务危机公司的正确预测占比达到 73%，对于非财务风险公司的预测也能达到 86.55% 的准确度。可以看出模型对于非财务风险公司的预测准确度较高，而对于财务风险公司的预测准确度稍低，可能仍然存在提升的空间。

5. 基于 1:1PSM 匹配样本下 logistic 回归

在已有研究的基础上，将前文有所介绍的倾向得分匹配方法，依次应用于 T-1、T-2、T-3 三年的样本中，从而考察是否可以提高模型的拟合优度预测准确度。

这里对三年的样本公司新引入两个指标，并依据这两个指标进行倾向得分匹配：

(1) 行业指标。行业分类取自 Wind 数据库中 Wind 行业一级分类，共将样本公司划分为采矿业、交通运输业、电力、热力、燃气及水生产和供应业、租赁和商务服务业等 19 类公司，将行业分类作为虚拟变量参与匹配；(2) 企业规模指标。企业规模指标选用上市公司 18 年年底总市值数据，将总市值数据作为变量参与匹配。

在匹配方法上，本文研究借鉴已有的学者的相关研究经验，使用最邻近匹配法，对样本计算倾向得分并进行匹配，得到样本后，将三年的数据进行 Logistic 建模，得到如下的结果：

表 4.16 T-1、T-2 与 T-3 年 logistic 模型

	Log likelihood	Number of obs	LR chi2(4)	Prob > chi2	Pseudo R2
T-1	-97.2283	231	125.57	0.0000	0.3924
T-2	-110.842	236	75.3	0.0000	0.2466
T-3	-187.423	231	24.66	0.0007	0.2163

从三个模型的对比来看，与 1:1 随机匹配样本模型的情况类似，在 1:1 倾向得分匹配样本下进行建模，同样是 T-1 年的表现更优：虽然三个模型均在 99% 的显著性水平下显著，但 T-1 年模型的 Log likelihood 值小于 T-2 与 T-3 年，伪 R 平方也大于另两个模型，表明其有更好的拟合优度。T-1 年 Logistic 模型的具体情况如下表：

表 4.17 1:1PSM 匹配下 T-1 年 Logistic 模型

Logistic regression				Number of obs =		231
				LR chi2(4) =		108.67
				Prob > chi2 =		0.0000
Log likelihood =		-97.228393		Pseudo R2 =		0.3924
Y	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
Factor1	-1.78293	0.904669	-1.97	0.119	-9.88419	-0.00981
Factor2	-1.24854	0.390522	-3.2	0.98	-5.61956	-0.48313
Factor3	-1.83677	0.577895	-3.18	0.992	-4.38772	-0.70412
Factor4	-0.21071	0.521524	-0.4	0.003	-3.98462	0.811463
_cons	0.711693	0.438117	1.62	0.104	-0.147	1.570386

由 T-1 年数据建立 Logistic 回归模型，由上表可以看出，最终建立的 Logistic 模型其函数表达式为：

$$\ln \frac{p}{1-p} = -1.783Factor1 - 1.249Factor2 - 1.837Factor3 - 0.211Factor4 + 0.712$$

同时，通过下表的模型预测准确度可以发现模型较为有效：

表 4.18 1:1PSM 匹配下 T-1 年 Logistic 模型预测准确度

TRUE			
Classified	D	~D	Total
+	91	20	111
-	28	92	120
Total	119	112	231
Classified + if predicted Pr (D) >= .5			
True D defined as Y != 0			
Sensitivity		Pr (+ D)	78.47%
Specificity		Pr (- ~D)	82.14%
Positive predictive value		Pr (D +)	81.98%
Negative predictive value		Pr (~D -)	76.67%

FALSE + rate for TRUE ~D	$\Pr(+ \sim D)$	17.86%
FALSE - rate for TRUE D	$\Pr(- D)$	23.53%
FALSE + rate for classified +	$\Pr(\sim D +)$	18.02%
FALSE - rate for classified -	$\Pr(D -)$	23.33%
Correctly classified		79.22%

从上表可以看出,使用倾向得分匹配后的模型在总体准确度上基本持平于 1:1 随机匹配下的 Logistic 模型,达到了基本相同的效果,表明其总体不输于原模型。但从表格中可以看出,运用 1:1 随机匹配下的 Logistic 模型在预测财务风险公司时,准确度为 73%,而运用 1:1 倾向得分匹配方法的 Logistic 模型,在预测财务风险公司时的准确度为 78%,准确度有 5%的提升。因此可以发现,运用 1:1 倾向得分匹配后,模型有着更为良好的表现。在现实生活中,该财务风险预测模型多用于对公司的财务风险进行预测,因此其正向分类的准确度就成为了一个重要的指标。从上表可以看出,1:1 倾向得分匹配模型的应用价值得到了进一步的提升。

使用 20-21 年数据进行样本外验证,该模型有着较为理想的准确度。

表 4.19 20-21 年 Logistic 模型预测准确度

TRUE			
Classified	D	~D	Total
+	124	39	163
-	48	131	179
Total	172	170	342
Classified + if predicted $\Pr(D) \geq .5$			
True D defined as $Y \neq 0$			
Sensitivity		$\Pr(+ D)$	72.27%
Specificity		$\Pr(- \sim D)$	77.31%
Positive predictive value		$\Pr(D +)$	76.11%
Negative predictive value		$\Pr(\sim D -)$	73.60%
FALSE + rate for TRUE ~D		$\Pr(+ \sim D)$	22.69%
FALSE - rate for TRUE D		$\Pr(- D)$	27.73%
FALSE + rate for classified +		$\Pr(\sim D +)$	23.89%
FALSE - rate for classified -		$\Pr(D -)$	26.40%
Correctly classified			74.79%

从上表可以看出,该模型在样本外预测时也有着较好的准确度,财务风险企业预测准确度达到约 75%,其中对财务风险企业(ST 企业)达到 72%,对非财务风险企业(非 ST 企业)达到 77%。综上所述,该模型有一定的实际价值,基本可以满足统计学上的应用要求。

(四) 模型解释

对所建立的统计模型进行复盘及解释，可以发现该统计模型有以下几方面的经济含义：

(1) 无论是以 1:1 随机匹配方法所处理的样本建立模型，还是以 1:1 倾向得分匹配方法进行建模，都可以发现利用 T-1 年的样本进行建模，相对于 T-2 及 T-3 年都可以得到预测效果更加良好的 Logistic 模型，主要表现在模型的伪 R 平方值都更低，及预测准确度更好。其中的经济含义即企业在出现财务风险问题的前一年的各项财务指标最具有时效性，包含了最多企业当前的隐含风险相关的信息。这也说明企业的财务风险是一个由小变大，渐进的过程。在 T-3 及 T-2 年中，部分财务指标作为模型中的变量在财务风险组与非财务风险组之间没有显著性差别，表明此时财务风险仅仅反映在了公司财务某些方面，公司只有局部的财务风险；而在 T-1 年出现异常的财务指标显著多于前两年，说明公司即将出现整体的财务风险。这其中的过程是渐进的。同时也说明，T-1、T-2、T-3 年模型需要综合应用，从而公司管理者，外部监督及投资者可以尽早预判到可能出现财务风险的公司，就能更为及时进行调整策略，防范风险；

(2) 从随机匹配及 PSM 匹配的因子载荷矩阵来看，在四个因子载荷较高的原始变量中，盈利能力变量及成长能力变量都比较显著，说明这两方面的财务问题能够显著影响到企业发生财务风险的可能性。盈利能力是企业财务的核心关注点，企业的债务偿还，研发投入，利润分配，发展战略等等都与盈利相关。盈利可以直观，清晰地反映企业目前的财务健康状况与未来发展前景。而在盈利之上企业才能不断发展，主营业务才能不断扩大，企业的发展才有保障。因此，模型包括了较多的盈利能力方面与成长能力方面变量的信息是合理的，这也符合实际情况；

(3) 非财务指标同样有一定的预警作用。变量指标是构建财务风险预警模型的核心问题，而指标的选择如果过于局限，则会显著影响建模效果。从所选的管理费用率与审计意见这两个指标来看，其在模型中均起到了显著性水平以上的预测效果。这也是符合经济规律的。审计意见出自外部专业的审计机构，其对公司财务的熟悉程度较高，专业性较强，在审计时很可能注意到一些财务上潜藏的风险诱因。而在管理费用率上看，管理费用比率过高则明显说明公司在管理上存在内生性问题，并最终影响公司的日常运行和战略决策。

五、 研究总结及展望

(一) 研究结论

本文研究基于 A 股上市公司的财务数据，借鉴已有的学者的文献、理论模型和研究经验，将 ST 公司以及非 ST 公司作为处理组及控制组，引入了五个类别的财务变量及非财务变量建立了 Logistic 回归模型，并应用 1:1 倾向得分匹配对模型进行改进，建立了一个清晰简洁，较为有效的预测模型，并通过对模型的检验，证明了模型的应用价值较高。

本文主要有以下的研究结论：

(1) 从 Logistic 模型的回归结果来看，T-1，T-2 及 T-3 年的财务变量数据及非财务变量数据进行对比，可以看出 T-1 年的变量数据包含了最多的公司财务状况信息，使用 T-1 年的财务变量数据及非财务变量数据进行 Logistic 建模，可以达到相较 T-2 与 T-3 年更好的效果；这也说明企业陷入财务危机的过程是一个动态的过程；

(2) 以现金流量、营运能力、偿债能力、盈利能力和成长能力等方面构建的财务指标普遍对企业财务风险具有较强的解释性与预测性，表明财务风险的发生将首先反映在财务指标上，其中的信息非常具有预示意义；在五个指标中，盈利能力的财务变量指标普遍表现更优，原因可以解释为企业没有良好的盈利能力，其资不抵债进而发生财务风险的可能性相对较大；

(3) 非财务指标通过了显著性检验，在模型所形成的因子中也具有一定的解释效力，表明加入非财务指标变量对于模型的预测是有效果的；

(4) 运用倾向得分匹配方法对样本进行 1:1 匹配进而建立的模型，其预测效果优于运用 1:1 随机匹配样本的模型，这个方法可以显著提升模型的统计学价值。

(二) 研究局限

(1) 本文选择的样本为 A 股上市公司前三年的财务数据作为样本，但是根据研究后的回顾，财务风险公司，即 ST 公司的各项财务数据在被进行特别处理的前四年中变化较大，有较大的波动，可能会对模型产生有效性上的影响；

(2) 本文研究建模所用的数据，大部分都来自于 Wind 数据库中上市公司按照要求所披露的财务报告，研究数据局限在了上市公司范围内；

(3) 本文从现金流量、营运能力、偿债能力、盈利能力和成长能力等五个角度选取

了 16 个财务指标变量及 2 个非财务变量，涵盖的范围可能相对较少，其中非财务指标没有考虑公司治理结构等方面的变量，范围较小；

(4) 企业年报等财务报表只能反映出公司在某一阶段的财务状况，即在过去一个会计年度内的盈利、债务及现金流量的情况，对企业财务状况的反映具有一定的滞后性；采用该数据进行建模，可能模型的判断的及时性会受到影响；

(5) 在诸多财务变量中，同一类别下的财务指标很可能出现内生性问题，影响模型的判别效果；对于变量中可能出现的内生性问题，虽然进行倾向得分匹配可以减小一部分内生性，但是无法做到较为彻底的解决。

(三) 未来研究展望

如何合理、高效地判断企业的财务状况所处的状态，判定公司是否有发生财务风险的可能性及发生财务风险的概率，对于公司的日常管理及公司的未来发展都具有重要意义。本文依据上市公司财务数据建立了 Logistic 模型，来对企业财务风险进行预测和分析。综上所述，本文的研究仍需要做一定量的改进。在今后的研究中可以从以下几方面进行深入研究，提高模型的预测效果：

(1) 从现金流量、营运能力、偿债能力、盈利能力和成长能力等方面选取更多的财务变量指标，同时从股权结构，董事会结构，公司代理等方面选取更多的非财务指标加入到模型中，以期取得更加精准的预测度；

(2) 尝试从多个渠道获取非上市公司财务数据，如果能够将准确、真实的非上市公司的有关财务数据加入到建模之中，模型可以预期达到更广泛的适用性；

(3) 选用合适的工具变量对财务指标变量间的内生性进行处理，以期达到更加具有独立性的财务变量进而提升模型的有效性。

参考文献

- [1] 陈芳、吴杰:《中小企业财务危机预警模型比较研究——基于因子分析与 Logistic 回归模型的对比》,《财会通讯》,2017 年第 5 期,第 106-108 页。
- [2] 陈磊、任若恩、曹汉平:《公司多阶段财务危机动态预警研究》,《系统工程理论与实践》,2008 年第 11 期,第 29-35 页。
- [3] 陈晓、陈治鸿:《企业财务困境研究的理论、方法及应用》,《投资研究》,2000 年第 6 期,第 29-33 页。
- [4] 黄善东、杨淑娥:《公司治理与财务困境预测》,《预测》,2007 年第 2 期,第 63-67 页。
- [5] 刘小龙:《基于 Logistic 模型的中小企业财务危机预警研究》,东华大学,Y2506243,2014 年。
- [6] 卢永艳、王维国:《基于面板 logit 模型的上市公司财务困境预测》,《数学的实践与认识》,2010 年第 5 期,第 37-43 页。
- [7] 彭大庆、陈良华、陈春苗:《上市公司财务困境动态预测模型的实证研究》,《工业技术经济》,2006 年第 10 期,第 147-150 页。
- [8] 舒惠好:《财务预警系统研究第行业差异变量模型假说》,《财会通讯》,2005 年第 8 期,第 77-81 页。
- [9] 王克敏、姬美光:《基于财务与非财务指标的亏损公司财务预警研究——以公司 ST 为例》,《财经研究》,2006 年年第 7 期,第 63-72 页。
- [10] 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》,2001 年第 6 期,第 46-55 页。
- [11] 于谦龙、顾晓敏:《基于最小二乘支持向量机的高校财务困境预警研究》,《统计与决策》,2011 年第 8 期,第 142-144 页。
- [12] 张爱民、祝春山、许丹健:《上市公司财务失败的主成分预测模型及其实证研究》,《金融研究》,2001 年第 3 期,第 10-25 页。
- [13] 张玲:《财务危机预警分析判别模型》,《数量经济技术经济研究》,2000 年第 3 期,第 49-51 页。
- [14] 郑芳云:《我国上市公司财务困境预测研究——基于有序三分类 Logistic 回归》,浙江大学,Y2400296,2013 年。
- [15] 周首华、杨济华、王平:《论财务危机的预警分析——F 分数模式》,《会计研究》,1996 年第 8 期,第 8-11 页。
- [16] Altman, Edward I., 1968, Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *The Journal of Finance (New York)*, 23(4), pp. 589-609.

- [17]Altman, Edward I, 1977, ZETA Analysis: A New Model to Identify Bankruptcy Risk of Corporations. *Journal of Banking & Finance*, 1(1), p. 29.
- [18]Bandopadhyaya, and Arindam., 1994, An Estimation of the Hazard Rate of Firms Under Chapter 11 Protection. *The Review of Economics and Statistics*, 76(2), pp. 346-350.
- [19]Beaver, and William H., 1966, Financial Ratios as Predictors of Failure. *Journal of Accounting Research*, 4, pp. 71-111.
- [20]Collins, Robert A, and Richard D Green., 1982, Statistical Methods for Bankruptcy Forecasting. *Journal of Economics and Business*, 34(4), pp. 349-354.
- [21]Fitzpatrick. P. J, 1932, A Comparison of Ratios of Successful Industrial Enterprises With Those of Failed Firms. *Certified Public Account*, 6, pp. 727-731.
- [22]Hamilton, and Martin A., 1977, Estimating the Logistic Curve from Grouped Data. *Journal of Statistical Computation and Simulation*, 5(4), pp. 279-301.
- [23]Henebry, and Kathleen L., 1996, Do Cash Flow Variables Improve the Predictive Accuracy of a Cox Proportional Hazards Model for Bank Failure. *The Quarterly Review of Economics and Finance*, 36(3), pp. 395-409.
- [24]Lane and William R., 1986, An Application of the Cox Proportional Hazards Model to Bank Failure. *Journal of Banking & Finance*, 10(4), pp. 511-531.
- [25]Michael Y. Hu, G. Zhang, and B. Eddy Patuwo., 1999, Artificial Neural Networks in Bankruptcy Prediction: General Network and Cross-Validation analysis. *European Journal of Operation Research*, 116(1), pp. 16-32.
- [26]Odom. M. D, and R. Sharda. 1990, A Neural Network Model for Bankruptcy Prediction. IJCNN International Joint Conference on Neural Networks.
- [27]Ohlson J. A., 1980, Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*, 18(1), pp. 109-131.
- [28]Wilcox, and Jarrod W. A., 1971, Simple Theory of Financial Ratios as Predictors of Failure. *Journal of Accounting Research*, 9(2), pp. 389-395.
- [29]Zmijewski, and Mark E., 1984, Methodological Issues Related to Estimation of Financial Distress Prediction Models. *Journal of Accounting Research*, 22(2), pp. 59-82.

后记

光阴荏苒，本科的四年时间一晃而过，转眼间，我也到了迎接毕业季的时候。回首在经济学院进行学习的日子，尽管只是弹指一挥间，已经足够让我思考很多，积累很多，沉淀很多。无论是对学术的态度，还是对人世间的态度，与四年前那个对什么都一知半解的我相比，现在的我已经焕然一新。

四年前，当我高考填志愿时，我仅仅是抱着对金融高薪职业的憧憬与幻想，选择了经济学院。但随着学习的深入，我逐渐明白了，经济不仅仅是一个“教人赚钱”的学问。还记得王弟海老师在《微观经济学》第一节开宗明义：“除非你有比别人更多的信息，否则从长远看，你没有办法比别人赚到更多的钱。”是的，这就是经济学的规律。经济学就是这样一门学问：借助分析的手，去拨开繁复万千的现象，探寻隐藏在深处的规律。经济学又是一门实践的学问，一门需要“情怀”的学问。他促使我们将目光从黑板上，书本中和软件里移开，去关注这片大地上的人来人往与川流不息。经济数字的背后是什么？经济运行为谁带来了好处？是否大家的福利都因此增多？每一位经济学的研究者，经济的从业者，经济规则的制定者，都必不可少的就是透过纸面，感受真实，“置身事内”的情怀。本科的四年，正是树立人生理想的四年。“穷则独善其身，达则兼济天下”，未来千变万化，谁也难以在多变的时代做一个不变者，但当东风吹散迷雾，我坚信我会发挥自己的热量，“摆脱冷气，只是向上走”，用自己的工作为他人带来美好与希望。四年学习的成功与否，我相信我可以用未来的实绩进行回答。

纵观我四年的学习，我最想感谢的就是我的父母，身边的老师，同学，朋友以及一切给予我帮助的人。感谢所有任课老师，将我领入经济学的大门。我的导师陈思好老师，在我的论文写作中给予了悉心的指导与细致的帮助，在此致以诚挚的感谢。

感谢与我朝夕与共的朋友们。“SOS 团”的群友，“混元形意门”的全体群友，以及其他所有带给我欢笑，与我知心相交的朋友以及网友们，谢谢你们。感谢所有来自 ACGN，出现在我精神世界的朋友们。

感谢我的父母，是你们默默支持了我，无私地为我奉献，你们让我时刻感到幸福和幸运。

虽然没有突出的成绩，亮眼的表现以及高光时刻，我也仍旧认为我在复旦的四年是一段不可磨灭的经历。在这里，我看到了一个更大的平台，我的视野不断开阔，心胸不断宽广。我衷心希望从这里启航，能够乘风破浪，一往无前。