



基于公司治理的 信息披露舞弊预警研究

蔡志岳, 吴世农

厦门大学 管理学院, 福建 厦门 361005

摘要: 以 2001 年~2005 年间的 192 家被认定为信息披露舞弊的 A 股上市公司及相应的 192 家配对公司为样本, 基于财务指标和治理指标, 分别运用 Logistic 回归分析和混合 BP 神经网络构建上市公司信息披露舞弊的预警模型。实证结果表明, 治理指标有助于提高信息披露舞弊预警模型的有效性, 混合 BP 神经网络模型的预测能力更强。

关键词: 公司治理; 信息披露舞弊; 混合 BP 神经网络; Logistic 回归

中图分类号: F830.91 文献标识码: A 文章编号: 1672-0334(2006)04-0079-12

An Early Warning Study on Disclosure Fraud Detection Based on Corporate Governance

CAI Zhiyue WU Shiong

School of Management Xiamen University Xiamen 361005 China

Abstract This paper collects a sample of 192 A-share listed companies committing disclosure fraud spanning from 2001 to 2005 and the corresponding 192 matched companies. Based on financial indicators and corporate governance indicators, the paper develops two disclosure fraud detection models by applying the Logistic regression and hybrid BP neural network method respectively. The empirical results show that corporate governance indicators help to improve the efficiency of the prediction model. Moreover, the hybrid BP neural network model dominates the Logistic regression model.

Key words corporate governance; disclosure fraud; hybrid BP neural network; Logistic regression

1 引言

上市公司信息披露是指上市公司将直接或间接地影响到投资者决策的重要信息以公开报告的形式提供给投资者, 它是上市公司内部人与外部投资者之间减少信息不对称、降低资本成本的重要方式, 也是市场监管者加强市场监管、提高资本市场配置效率的手段。在大股东掏空上市公司、内部人操纵股价以及公司迎合新股上市融资、配股和增发政策或公司避免特别处理和退市处理等目的的驱动下, 一

些上市公司通过未及时披露公司重大事项或定期报告、业绩预测结果不准确或不及时、信息披露虚假或严重误导性陈述以及未依法履行其他职责等信息披露舞弊手段逃避证券监管和法律制裁, 一旦东窗事发, 投资者、债权人等相关利益者将遭受巨大损失, 整个证券市场也将面临信任危机。

如何预防和发现上市公司的信息披露舞弊成为投资者、债权人乃至监管层的当务之急, 国外财务学者和实务界对信息披露舞弊特别是财务舞弊的预警

收稿日期: 2006-04-19

基金项目: 国家自然科学基金 (70372035)

作者简介: 蔡志岳 (1978-), 男, 福建晋江人, 厦门大学管理学院博士研究生, 研究方向: 证券与投资等。

研究做出了不懈努力。Coglitore和Berryman发现,简单的分析性程序可以揭示财务报告项目的显著变化和异常关系^[1]。美国审计标准报告第56号文件(1997)明确要求会计师在审计计划阶段和结束阶段必须实施分析性程序,但是财务舞弊行为日益复杂,简单的分析性程序越来越不能满足需要,财务学者们开始求助于统计判别模型。Peirson运用逐步Logistic回归模型成功判别了大部分的财务舞弊,并指出会计数据包含甄别财务舞弊的有效信息^[2]。Beneish根据8个财务指标建立了Probit回归预测模型,总体判别成功率达到75%^[3]。Bell和Carcello运用Logistic回归模型证明了基于财务数据的预测模型在判别舞弊公司方面比经验丰富的审计师表现出色^[4]。

由于舞弊样本信息和知识获取的困难,统计判别模型始终无法取得令人满意的预警效果,财务学者不得不探索更有效的复杂模型。Eining、Jones和Loebbecke根据审计师不情愿依赖于过度定量决策方法的特点,构建了基于审计师对话和Logistic模型的整合专家系统,实验表明这个系统可以帮助审计师大大提高对于财务舞弊的判别成功率^[5]。Green、Choi、Fanning和Cogger基于公开财务数据,以财务指标为输入变量,运用人工神经网络技术建立了财务舞弊预测模型,发现它的预测成功率显著高于统计预测模型,并建议审计师在审计初始阶段使用该模型^[6,7]。Spahis、Doupoupos和Zopounidis采用一种多标准辅助判别方法对10个财务指标建立预测模型,发现这种方法比传统的判别方法更有效^[8]。Lin、Hwang和Becker根据会计收益指标和会计收益趋势构建了基于模糊神经网络的财务舞弊判别模型,实证结果表明它可以有效地减少审计师的偏见或弥补审计师的不足^[9]。

对于信息披露舞弊的预警研究国内尚处于起步阶段。鹿小楠和傅浩在Beneish的研究基础上,以舞弊当年的9个财务指标为预警变量,分别建立多元判别模型和Logistic回归模型,但对舞弊公司的预测成功率都仅仅为60%,远远不能满足应用要求^[10]。

综上所述,国内外研究者将分析性程序、统计模型甚至人工神经网络等复杂方法应用于预警研究,取得了一些成果,但预警模型大多以财务指标作为输入变量,研究局限于财务视野,使模型的有效性始终没有实质性的突破。财务指标只是公司经营陷入困境的表面特征,易受时效的限制,这可能导致预警模型的不稳定性,而且财务信息的披露容易被管理层操纵,使财务指标传递的信息夹杂噪音甚至完全扭曲。例如安然公司运用了复杂的融资工具,即特殊目的主体和其他衍生工具,避免在资产负债表上披露债务,借以提高财务杠杆,若仅从三大报表的财务指标根本无法发现安然的问题。董事会是公司经营战略的制定者,而高管人员是既定战略的执行者,公司舞弊的根本原因在于公司的股权结构失衡、董事会机制失效、管理机制腐化和外部治理缺陷等微

观因素。本文在财务定量指标的基础上,引入治理指标等定性信息,运用因子分析和逐步Logistic回归分析进行预警变量的提炼和筛选,力图提高模型的有效性,尝试将经过混合算法修正的BP神经网络模型引入信息披露舞弊的预警研究,并与Logistic回归模型进行比较,探索人工智能技术应用于财务研究的可行性。

2 研究设计

2.1 研究样本的选择

2.1.1 样本的选择

从1999年上市公司开始对外披露公司治理信息,2000年12月财政部颁布了《企业会计制度》,并对会计制度实施了重大改革,为了保证样本数据的真实性和可比性,本文的研究期间为2001年~2005年。根据2001年~2005年间中国证监会、沪深交易所和财政部对上市公司的公开谴责、公开批评和公开处罚的公告,本文选取了192家非金融类舞弊A股上市公司为信息披露舞弊样本(剔除了力合股份、中农资源、潜江制药、ST吉光华和江苏琼花等5家数据不全的样本),对每一家舞弊公司按以下步骤选取配对公司。①行业,按证监会行业细分标准,选择与其细分行业相同的公司;②公司规模,在满足行业配对标准前提下,选择与其资产总额最接近的公司;③上市交易年限,在满足行业和规模配对标准前提下,选择与其上市交易年限最接近的公司;④公司声誉,在研究期间内,配对公司必须不在受监管层处罚的上市公司之列。

2.1.2 数据描述

研究数据主要来源于WIND资讯和CCER公司治理数据库,部分治理数据由作者根据年报补充整理。在研究期间内,对于只有一次舞弊的公司,将其被处罚的当年作为舞弊年度;对于连续两年或更多年舞弊的公司,将其舞弊的第一年作为舞弊年度。表1统计了2001年~2005年间的192个舞弊样本的行业分布,其中制造业的舞弊公司最多,高达105个,占样本比例为54.69%;综合类其次,舞弊公司为16个,占样本比例为8.33%;信息技术业、批发和零售贸易业的舞弊公司分别为15个和14个,占样本比例分别为7.81%和7.29%。

2.2 预测变量的设计和筛选

本文以哑变量FRAUD(0/1)作为度量上市公司信息披露舞弊的代理变量(因变量),即在研究期间内是否因信息披露舞弊行为被监管层公开谴责、批评或处罚,是为1,否为0。Feroz Park和Pastena的研究表明,公司舞弊的第一年同舞弊被发现之间的时间间隔平均为19个月^[11],因此记公司被发现舞弊的第一年为T年,舞弊前的最近两个会计年度分别为(T-1)年和(T-2)年,本文以(T-1)年的数据为主,以(T-2)年的数据为辅,基于财务视角和治理视角,选择可能包含揭示公司在T年是否舞弊的有效信息的预警变量。

表 1 舞弊样本行业分布

Table 1 Industry Distribution of Fraud Sample

行业类别	公司数目	占样本比例
A 农、林、牧、渔业	9	4.69%
B 采掘业	2	1.04%
C 制造业	105	54.69%
C0 食品、饮料	15	
C1 纺织、服装、皮毛	9	
C2 木材、家具	1	
C3 造纸、印刷	5	
C4 石油、化学、塑胶、塑料	19	
C5 电子	5	
C6 金属、非金属	12	
C7 机械、设备、仪表	30	
C8 医药、生物制品	7	
C9 其他制造业	2	
D 电力、煤气及水的生产和供应业	5	2.60%
E 建筑业	3	1.56%
F 交通运输、仓储业	5	2.60%
G 信息技术业	15	7.81%
H 批发和零售贸易	14	7.30%
J 房地产业	7	3.65%
K 社会服务业	5	2.60%
L 传播与文化产业	6	3.13%
M 综合类	16	8.33%
合计	192	100%

2.2.1 备选财务指标

作为上市公司经营业绩的最直观和最公开的表现形式, 财务指标一直是研究者进行舞弊预警的首选变量。本文在 Beneish 经验研究的基础上, 结合中国实际情况, 选择以下 8 个财务指标作为备选预警变量^[3], 见表 2。

(1) 应计利润率 (*ACTA*)。 *ACTA* 是指 ($T-1$) 年公司的应计利润与总资产的比值。公司的净利润可以分解为经营性现金流净额与应计利润两部分, 其中应计利润是指那些不直接形成当期现金流入或流出, 但按照权责发生制和配比原则应计入当期损益

的那些收入或费用 (例如折旧、摊销和应收账款增加额等)。应计利润是公司管理层进行盈余管理的主要手段, 高额的应计利润往往预示着管理层进行利润操纵^[12-13]。

(2) 财务杠杆指数 (*LEVI*)。 *LEVI* 是指 ($T-1$) 年公司的资产负债率与 ($T-2$) 年公司的资产负债率的比值。财务杠杆比例的恶化很可能是公司管理层舞弊的债务契约动因^[14]。

(3) 资产质量指数 (*AQI*)。 *AQI* 是指 ($T-1$) 年公司的资产质量比率与 ($T-2$) 年公司的资产质量比率的比值, 而资产质量比率是除固定资产之外的非流动资产与总资产的比值, 用以衡量总资产中潜在收益较不确定的那部分的比例。 *AQI* 是衡量资产实现风险变化的指标^[15], 若大于 1 意味着公司可能资本化或递延费用。

(4) 销售收入指数 (*SGI*)。 *SGI* 是指 ($T-1$) 年公司的主营业务收入与 ($T-2$) 年公司的主营业务收入的比值。基于美国数据的研究显示, 高增长的公司管理层往往面临更大的财务压力和资本市场压力^[16], 为迎合市场的增长预期和防止股价下跌, 在公司经营不如意的情况下, 管理层拥有强烈的舞弊动机。然而, 需要注意的是, 销售收入的大幅增加并不意味着公司一定进行财务舞弊。

(5) 应收账款周转率指数 (*DSRI*)。 *DSRI* 是指 ($T-2$) 年公司的应收账款周转率与 ($T-1$) 年公司的应收账款周转率的比值, 用来度量两个连续的会计年度中应收账款和销售收入能否保持平衡。应收账款周转率的大幅度下降很可能是因为公司虚增销售收入。

(6) 毛利率指数 (*GMI*)。 *GMI* 是指 ($T-2$) 年公司的毛利率与 ($T-1$) 年公司的毛利率的比值。毛利率水平的大幅度下降意味着公司的经营环境和赢利前景严重恶化^[17], 公司管理层迫于压力往往铤而走险, 以致公司舞弊。

(7) 三费比例指数 (*SGAI*)。 *SGAI* 是指 ($T-1$) 年公司的三费比例与 ($T-2$) 年公司的三费比例的比值, 其中三费比例是指管理费用、销售费用和财务费用之和占销售收入的比例。三费比例的大幅增长说明公司的经营竞争环境严重恶化^[17]。

(8) 折旧率指数 (*DEPI*)。 *DEPI* 是指 ($T-2$) 年公司的折旧率与 ($T-1$) 年公司的折旧率的比值。当折旧率指数大于 1 时, 固定资产的折旧速度下降, 这可能意味着公司管理层采取了盈余管理政策。

2.2.2 备选治理指标

现代公司治理理论认为, 良好的公司治理机制有赖于内部机制和外部机制的有机结合。内部机制包括股权集中度、董事会治理和高管激励等机制, 外部机制包括公司控制权市场、中小投资者法律保护、产品竞争市场、独立审计和机构投资者监督等机制。基于国外学者关于公司治理的理论和经验研究, 结合中国的具体治理实际, 本文以下列 8 个维度构建预警变量, 见表 2。

表 2 研究变量定义
Table 2 Definitions of Variables

类型	名称	符号	定义
因变量	信息披露舞弊	<i>FRAUD</i>	哑变量: 在 2001 年 ~ 2005 年期间内, 是否因信息披露舞弊被公开谴责、批评或处罚, 是为 1 否为 0
	应计利润率	<i>ACTA</i>	$\frac{(\text{净利润} - \text{经营性现金流量净额})_{(T-1)}}{\text{总资产}_{(T-1)}}$
	财务杠杆指数	<i>LEVI</i>	$\frac{\text{资产负债率}_{(T-1)}}{\text{资产负债率}_{(T-2)}}$
	资产质量指数	<i>AQI</i>	$\frac{\text{资产质量比例}_{(T-1)}}{\text{资产质量比例}_{(T-2)}}$ 其中, 资产质量比例 = $\frac{\text{总资产} - \text{流动资产} - \text{固定资产净值}}{\text{总资产}}$
备选 财务 变量	销售收入指数	<i>SGI</i>	$\frac{\text{主营业务收入}_{(T-1)}}{\text{主营业务收入}_{(T-2)}}$
	应收账款周转率指数	<i>DSRI</i>	$\frac{\text{应收账款周转率}_{(T-2)}}{\text{应收账款周转率}_{(T-1)}}$
	毛利率指数	<i>GMI</i>	$\frac{\text{毛利率}_{(T-2)}}{\text{毛利率}_{(T-1)}}$
	三费比例指数	<i>SGAI</i>	$\frac{\text{三费比例}_{(T-1)}}{\text{三费比例}_{(T-2)}}$, 其中, 三费比例 = $\frac{\text{管理费用} + \text{销售费用} + \text{财务费用}}{\text{主营业务收入}}$
	折旧率指数	<i>DEPI</i>	$\frac{\text{折旧率}_{(T-2)}}{\text{折旧率}_{(T-1)}}$
	第一大股东持股比例	<i>TOP1</i>	第一大股东持股比例
	董事会规模	<i>BSIZE</i>	$\ln(\text{董事会总人数})$
	独董比例	<i>OUTDIR</i>	$\frac{\text{独立董事人数}}{\text{董事会总人数}}$
	领导结构	<i>DUAL</i>	是否两职合一, 是为 1 否为 0
	审计委员会	<i>AUDCOMM</i>	是否设置审计委员会, 是为 1 否为 0
备选 治理 变量	董事会会议	<i>BMEET</i>	$\ln(\text{年度董事会会议次数})$
	高管激励	<i>TMP</i>	$\ln(\text{高管年薪总和})$
	控制权竞争程度	<i>HERF2_10</i>	第二至第十大股东的持股比例的 <i>Herfindahl</i> 指数
	法律保护程度	<i>PROTECT</i>	是否在其他市场挂牌上市, 是为 1 否为 0
	产品竞争程度	<i>PROCOM</i>	$\frac{\text{营业费用}}{\text{主营业务收入}}$
	审计意见类型	<i>AUDCLA</i>	审计意见是否为标准无保留, 是为 1 否为 0
	机构投资者监督	<i>FUNDHOLD</i>	基金持股占流通股的比例

表 3 备选财务预警变量逐步回归筛选结果
Table 3 Results of Stepwise Regression on Potential Financial Indicators

结束步骤	变量	系数	标准差	Wald 统计量	自由度	相伴概率
第 2 步	ACTA	1.473	0.367	16.119	1	0.000
	SGI	-0.280	0.127	4.846	1	0.028
	截距	0.125	0.116	1.149	1	0.284

(1) 股权集中度。股权集中度的变化将影响代理问题的性质, 从而影响公司行为。由于存在监督成本, 对于中小股东来说, 监督所带来的收益可能远远小于为此所付出的代价, 所以中小股东不愿意去监督管理层^[18], 只有大股东有足够的利益激励去实施监督^[19]。当股权过度分散时, 中小股东和管理层之间的利益冲突就产生了代理问题; 当股权过度集中时, 大股东和中小股东之间的利益冲突就会尤为突出。因此, 本文选择第一大股东持股比例 (TOP1) 作为股权集中度的代理变量。

(2) 董事会治理。董事会作为联结股东与管理层的纽带, 是公司治理机制的核心。董事会的治理效率直接关系到公司价值和股东利益, 如果董事会能有效地进行决策并监督管理层的经营行为, 公司的价值将提高, 股东利益也会得到有效保护。董事会的规模、董事会的独立性、董事会领导结构、是否设置审计委员会和董事会会议次数等董事会特征无疑将影响着董事会的监督效率, 进而影响着公司的信息披露行为^[20-26]。

(3) 高管激励。对管理层的激励有利于降低股东与管理层的代理成本, 确保管理层以股东利益最大化为经营目标^[27]。通常高管激励采取持股和年薪两种方式, 由于中国上市公司高管持股比例微乎其微, 激励作用极其有限, 所以本文以 ln (高管年薪总和) 作为高管激励 (TMP) 的代理变量。

(4) 公司控制权市场^[28]。尽管中国的公司控制权市场并不活跃, 但当其他大股东将投票权联合在一起时将成为控股股东实施隧道行为的制约力量, 而且这些股东的股权越集中, 越有可能去争取或协助外来者获得对公司的控制权, 本文以第二至第十大股东持股比例的 Herfindahl 指数作为公司控制权竞争程度的代理变量 (HERF2_10);

(5) 中小投资者法律保护^[29]。法律体系和监管标准是保护中小投资者的有效机制, 在 B 股上市和境外上市的公司, 由于成熟资本市场比较严格的监管标准和比较健全的法律体系的约束, 它们的信息披露行为相对于 A 股上市公司更加规范。本文以是否在其他市场挂牌上市作为法律保护程度的代理哑变量 (PROTECT), 是为 1 否为 0。

(6) 产品竞争市场。产品竞争程度的增加将有助于监督管理层^[29], 有助于限制控股股东的机会主义行为, 规范公司行为。本文以营业费用与销售收入的

比值作为度量产品竞争程度的代理变量 (PROCOM);

(7) 独立审计。独立审计是缓解代理问题的重要外部力量^[30,31], 注册会计师具有丰富的会计知识及审计经验, 能够对上市公司的财务状况及其变动进行深入分析, 形成高质量的审计意见, 非标准的审计意见意味着公司运营或信息披露行为不够规范。本文引入审计意见类型哑变量 (AUDCLA) 作为外部审计监督水平的代理变量, 当审计意见为标准无保留时为 1, 其他为 0。

(8) 机构投资者监督。机构投资者有动力采取积极态度去监督控股股东和管理层^[32], 即使它们没有实力去参与治理, 它们仍然有足够的去督促公司提高信息披露透明度, 规范披露行为^[33]。本文以上市公司的基金持股占流通股的比例作为机构投资者监督的代理变量 (FUNDHOLD)。

2.2.3 预警变量的筛选

为简化预警模型, 本文采用逐步 Logistic 回归分析对 8 个备选财务指标进行筛选 (结果见表 3), 并将与舞弊行为关系最密切的应计利润率 (ACTA) 和销售收入指数 (SGI) 作为模型的财务预警变量。

因为公司治理指标可能并非单一或简单联合起到治理作用, 更可能是相互交叉在一起共同达到治理效果, 因此本文采用因子分析法寻找治理指标的线性组合以最大化描述其变化情况, 并进行正交旋转, 提炼出 9 个治理因子变量, 见表 4 和表 5。

表 4 因子的特征值与解释方差

Table 4 Eigenvalues and Variance Explained Rate of Factors

主成分	特征值	解释方差 (%)	累积方差 (%)
Z ₁	1.732	14.435	14.435
Z ₂	1.420	11.837	26.272
Z ₃	1.061	8.846	35.117
Z ₄	1.060	8.834	43.951
Z ₅	1.057	8.805	52.756
Z ₆	1.018	8.483	61.239
Z ₇	1.009	8.411	69.650
Z ₈	1.006	8.384	78.034
Z ₉	1.005	8.377	86.410

表 5 旋转后的因子载荷矩阵
Table 5 Rotated Factorial Load ed Matrix

治理变量	Z ₁	Z ₂	Z ₃	Z ₄	Z ₅	Z ₆	Z ₇	Z ₈	Z ₉
AUDCOM	0.825	0.041	-0.180	-0.069	-0.200	-0.045	-0.037	0.104	0.030
OUTDIR	0.726	0.013	-0.009	0.388	0.019	-0.011	0.020	-0.052	-0.063
TMP	0.691	0.034	0.295	0.006	0.272	0.127	0.133	-0.005	-0.020
HERF2_10	0.024	0.860	0.004	-0.021	0.022	0.070	0.028	-0.064	-0.057
TOP1	-0.044	-0.816	0.035	-0.083	0.033	0.107	0.074	-0.069	-0.113
AUDCLA	0.007	-0.035	0.960	-0.090	-0.046	0.053	0.046	0.042	-0.005
BMEET	0.181	0.067	-0.092	0.941	0.019	-0.014	-0.029	-0.040	0.027
PROTECT	0.007	-0.012	-0.044	0.019	0.967	-0.029	0.001	0.036	0.035
DUAL	0.038	-0.036	0.060	-0.014	-0.025	0.987	-0.004	0.056	-0.042
FUNDHOLD	0.077	-0.049	0.053	-0.026	0.006	-0.004	0.990	0.013	0.007
BSIZE	0.050	0.002	0.042	-0.045	0.037	0.056	0.014	0.985	-0.076
PROCOM	-0.039	0.057	-0.007	0.018	0.035	-0.042	0.007	-0.075	0.987

表 6 备选治理预警变量逐步回归筛选结果
Table 6 Results of Stepwise Regression on Potential Governance Indicators

结束步骤	变量	系数	标准差	Wald统计量	自由度	相伴概率
	Z ₂	0.274	0.120	5.245	1	0.022
	Z ₃	-1.228	0.169	53.130	1	0.000
第 4步	Z ₄	0.430	0.121	12.588	1	0.000
	Z ₇	-0.638	0.212	9.042	1	0.003
	截距	0.101	0.132	0.577	1	0.448

Z₁ 为董事会的内控机制 (AUDCOM 和 OUTDIR), Z₂ 为股权结构的股权集中度和股权制衡度 (TOP1和 HERF2_10), Z₃ 为外部审计的监督程度 (AUDCLA), Z₄ 为董事会的会议强度 (BMEET), Z₅ 为中小投资者法律保护程度 (PROTECT), Z₆ 为董事会的领导结构 (DUAL), Z₇ 为机构投资者的监督程度 (FUNDHOLD), Z₈ 为董事会的规模 (BSIZE), Z₉ 为产品竞争市场的竞争强度 (PROCOM)。

采用逐步 Logistic 回归分析对 9 个治理因子变量进行筛选 (结果见表 6), 并将与舞弊行为关系最密切的 Z₂、Z₃、Z₄ 和 Z₇ 4 个治理因子变量作为治理预警变量。

2.3 研究方法的设计

考虑到上市公司信息披露舞弊行为的发生概率与公司所处的财务状况、治理水平之间的相互关系可能是线性的, 也可能是非线性的, 本文分别运用

Logistic 回归模型和混合 BP 神经网络对这一问题建模。

2.3.1 Logistic 回归模型

Logistic 回归模型判别的基本思想是: 设 X_i 为第 i 个舞弊公司的预测变量向量, 那么其舞弊的概率 P_i 和 X_i 之间存在如下回归关系, 即

$$P_i = \frac{1}{1 + e^{-y_i}} \quad (1)$$

其中,

$$Y_i = a + \sum b_j X_j \quad (2)$$

Logistic 回归模型根据样本数据使用最大似然法估计出各参数值, 进而通过 (1) 式计算求得公司舞弊的概率。Logistic 回归模型的曲线为 S 型, 且预测最大值趋近 1, 最小值趋近 0, 所以常被用于非线性 0-1 分类预测。本文选择 0.5 作为 Logistic 回归模型的分割点, 即如果通过模型计算出来的概率 P_i 大于或等于 0.5 那么该公司被判别为舞弊公司; 如果 P_i 小于 0.5 则被判别为非舞弊公司, 即正常公司。

2.3.2 混合 BP 神经网络模型

BP 神经网络具有很好的线性或非线性的输入-输出映射能力, 而且能够处理任意类型的数据, 信息容错能力强。但 BP 神经网络也存在着下列缺陷, 即容易陷入局部极小; 收敛速度慢和引起振荡效应; 鲁棒性不好, 网络性能对网络的初始设置比较敏感等。本文将 BP 神经网络和遗传算法结合起来, 形成一种混合训练算法, 达到优化 BP 神经网络的目的。训练首先用遗传算法对初始权值分布进行寻优, 在解空间中定位出一个较小的搜索空间, 然后采用 BP 算法在这个小的解空间搜索出最优解。混合 BP 神经网络的算法流程图和具体设计如图 1 所示。

(1) BP 神经网络配置

创建 BP 神经网络预测模型时需设计输入节点、输出节点、隐含层和激活函数。本文的建模方案具体如下。

① 输入节点的设计。根据逐步 Logistic 回归分析筛选所得的 2 个财务指标和 4 个治理因子作为输入变量, 变量个数 6 即为节点数目。

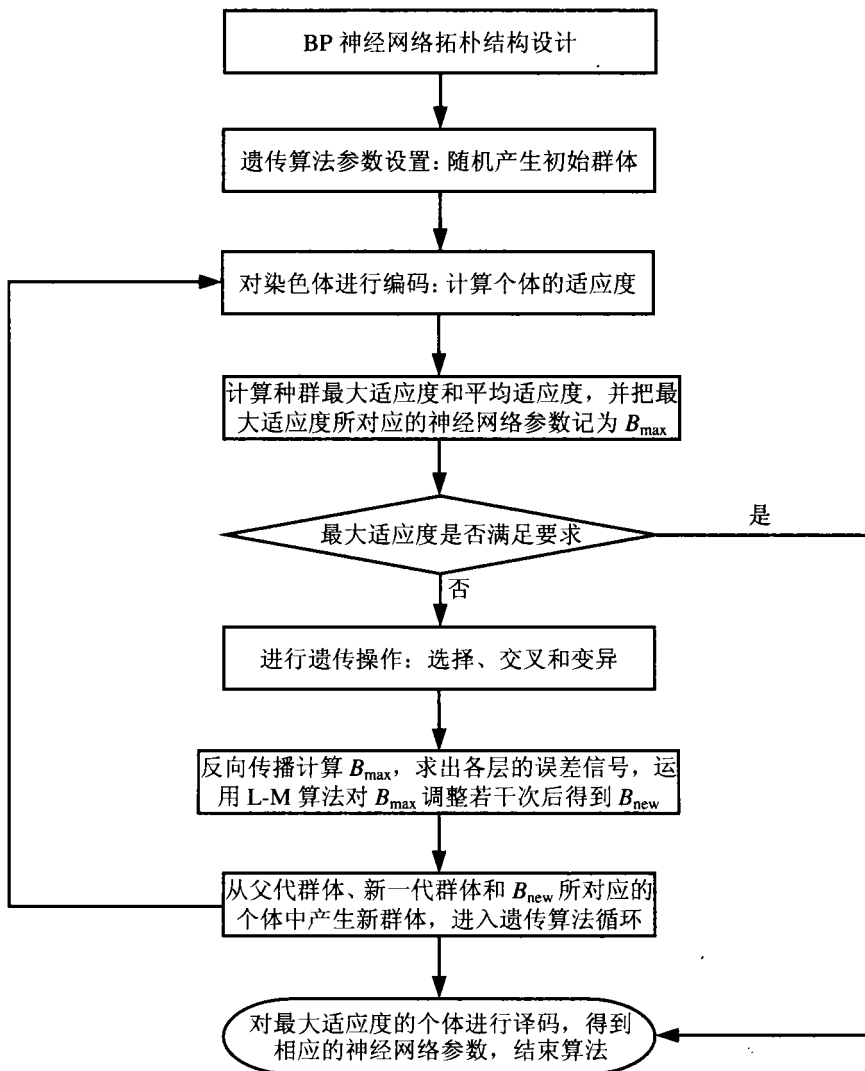


图 1 混合算法结构流程图

Figure 1 The Structured Flow Chart of Hybrid Algorithm

② 输出节点的设计。输出层的节点数取 1, 如果检验样本输入学习后的人工神经网络, 输出神经元的输出结果若大于或等于 0.5 则该公司为舞弊公司; 若小于 0.5 则为非舞弊公司。

③ 隐含层的设计。本文采用三层 BP 神经网络结构, 其中包含一个输入层、一个输出层和一个隐含层。隐含层节点数的选择对网络的性能影响很大, 遗憾的是, 至今仍没有具体的理论指导。本文参考经验公式得

$$n_1 = \sqrt{n+m} + a \quad (3)$$

其中, n_1 为隐含层节点数, n 为输入节点数, m 为输出节点数, a 为 1~10 之间的常数。根据不同的 n_1 值进行模型预测效果的弹性分析, 最终取 $n_1 = 7$ 。

④ 激活函数的选择。本文选择 sigmoid 函数作为激活函数。

(2) 自适应遗传算法设计

为了在保持群体多样性的同时保证算法的收敛性, 本文采用自适应遗传算法。具体参数设置如下。

① 群体规模与初始群体的设计。随机产生初始群体, 种群的大小取 160 最大进化代数取 50。

② 编码设计。本文采用浮点编码, 浮点编码的长度就是变量个数。

③ 适应度函数选择。本文采用误差平方和函数。

④ 遗传操作算子的设计。选择算子采用锦标赛法, 交叉算子采用浮点交叉, 变异算子采用浮点变异; 交叉概率 P_c 取 0.6~0.9 变异概率 P_m 取 0.001~0.1 并进行自适应调整。

3 实证研究结果及其分析

本文基于公司治理视角, 根据上市公司 ($T-1$) 年的财务指标和公司治理指标, 应用 Logistic 回归模型和混合 BP 神经网络构建公司在 T 年信息披露舞弊概率的预警模型, 并对两种模型的预警效果进行比较分析。

3.1 预警变量描述性统计

本文首先对舞弊公司(样本组)与其配对公司(控制组)的研究变量进行了描述性统计, 并对两组样本进行了单变量比较分析, 其中采用 t 检验比较均值差异, 采用 Wilcoxon 符号秩检验比较中值差异。统计结果见表 7。

表 7 变量描述性统计及单变量检验

Table 7 Descriptive Statistics and Univariate Tests of Variables

变量	组别	最大值	最小值	方差	均值	t 检验- p 值	中值	Wilcoxon检验- p 值
ACTA	样本组	14.287	-0.423	1.373	0.1714	0.001***	-0.120	0.000***
	控制组	1.114	-0.420	0.245	-0.1715		-0.240	
SGI	样本组	8.451	-2.636	1.186	-0.1366	0.005***	-0.260	0.000***
	控制组	5.027	-1.579	0.750	0.1363		-0.022	
Z_2	样本组	2.159	-1.638	0.955	0.0996	0.034**	0.225	0.027**
	控制组	3.106	-1.987	1.036	-0.1000		-0.226	
Z_3	样本组	1.381	-2.645	1.181	-0.4347	0.000***	0.223	0.000***
	控制组	1.191	-2.006	0.481	0.4349		0.504	
Z_4	样本组	2.887	-2.690	0.978	0.1602	0.000***	0.258	0.00***
	控制组	3.507	-2.518	0.998	-0.1599		-0.243	
Z_7	样本组	2.765	-0.682	0.427	-0.1707	0.001***	-0.268	0.156
	控制组	6.844	-0.690	1.328	0.1711		-0.269	

注: *为在 10% 的置信区间上显著, **为在 5% 的置信区间上显著, ***为在 1% 的置信区间上显著

显然, 对样本组与控制组而言, 不管是 t 检验还是 Wilcoxon 符号秩检验, 作为财务预警变量的应计利润率和销售收入指数以及作为治理预警变量的 Z_2 ($TOP1$ 和 $HERF2_{-10}$)、 Z_3 ($AUDCLA$)、 Z_4 ($BMEET$) 和 Z_7 ($FUNDHOLD$) 4 个治理因子变量都存在着显著差异, 这说明在舞弊被发现之前, 舞弊公司在财务状况或是治理水平上就与配对公司存在着显著的差异。为了更有效地提前将舞弊上市公司甄别出来, 本文构建了预警模型。

3.2 Logistic 回归预测

以筛选所得的 6 个自变量为输入变量, 以上市公司在 T 年信息披露舞弊的概率为输出变量, 应用 SPSS12.0 软件, 根据 130 对建模样本 ($T-1$) 年的财务指标和公司治理指标进行 Logistic 回归分析, 得到统计结果如表 8 所示。

从表 8 可以看出, 与理论假设相符, 应计利润率越高, 管理层越可能操纵利润, 公司舞弊的可能性也就越大; 销售收入指数越高, 公司越不大可能舞弊, 这与美国的经验研究结论不一样。一方面, 销售收入指数越高, 说明公司的经营前景越理想, 没有必要舞弊。另一方面, 这可能是因为美国高成长公司实行高额的薪金和股票期权激励, 管理层才会有强烈的舞弊动机; 而国内没有这样的高管激励方案, 在经营不佳时管理层缺乏动力去粉饰公司的高成长力。

与理论假设相符, 代表外部审计监督程度的 Z_3 和代表机构投资者监督程度的 Z_7 越大, 公司舞弊的可能性越小, 这充分说明了高质量的审计意见和机构投资者的积极态度有足够的影响力去监督和规范公司行为。代表董事会会议强度的 Z_4 越大, 公司越可能舞弊, 这说明董事会会议越多, 并不意味着董事越积极、越有效^[24], 而更可能是公司隐患越多^[19, 25], 使公司不得不舞弊, 铤而走险。与理论假设相矛盾, 代表股权结构的股权集中度和股权制衡度的 Z_2 越大, 即第一大股东持股比例越低, 第二至第十大股东的股权制衡力量越强, 公司越可能舞弊, 这可能意味着股权结构对于公司舞弊的影响较为复杂, Z_2 并不足以解释全部影响, 有待于进一步研究。

对所建模型进行 Chi square 统计检验和拟合优度检验, 发现 Chi square 在 1% 的置信区间上统计性显著, 意味着模型所包含的 6 个自变量对因变量有显著的解释能力; 而 Hosmer & Lemeshow 统计性不显著, 不能拒绝模型拟合数据很好的原假设。但是 $-2 \text{ Log Likelihood}$ 指标的值为 406.451 偏大, 即通过似然函数测量的所拟合的模型与饱和模型之间的差异程度较大; $Cox \& Snell R^2$ 和 $Nagelkerke R^2$ 指标的值分别为 0.28 和 0.373 偏小, 说明模型的拟合程度不是特别理想。

基于 6 个自变量, 构建信息披露舞弊的 Logistic 预警模型为

$$P_{-1} = \frac{1}{1 + e^{-(0.196 + 0.994CTA - 0.219SGT + 0.248Z_2 - 1.156Z_3 + 0.389Z_4 - 0.63Z_7)}}$$

最后, 利用其余 62 对检验样本对所构建模型进行检验, 预警模型对舞弊公司的判别成功率为 62.90%, 总体判别成功率为 69.35%, 预警结果并不非常理想, 具体见表 9。

表 8 Logistic 回归结果
Table 8 Results of Logistic Regression

类型	变量	预测符号	模型 I (0/1)
财务变量	$ACTA$	+	0.990 (5.603)**
	SGI	+	-0.219 (2.857)*
	Z_2	-	0.248 (4.165)**
治理变量	Z_3	-	-1.156 (45.697)***
	Z_4	-	0.389 (9.894)***
	Z_7	-	-0.630 (8.521)***
截距			0.196 (-1.89)
	$-2 \text{ Log likelihood}$		406.451
	$Cox \& Snell R^2$		0.280
	$Nagelkerke R^2$		0.373
	$Model Chi square$		125.886***
	$Hosmer \& Lemeshow Test$		6.174
观察值			260

注: *为在 10% 的置信区间上显著, **为在 5% 的置信区间上显著, ***为在 1% 的置信区间上显著

表 9 Logistic回归预测结果
Table 9 Predictive Results of Logistic Model

建模样本			检验样本				
实际类型	预测类型		总计	实际类型	预测类型		总计
	舞弊公司	正常公司			舞弊公司	正常公司	
舞弊公司	78	52	130	舞弊公司	39	23	62
正常公司	22	108	130	正常公司	15	47	62
舞弊判别成功率	60.00%			舞弊判别成功率	62.90%		
正常判别成功率	83.08%			正常判别成功率	75.81%		
总体判别成功率	71.54%			总体判别成功率	69.35%		

表 10 混合 BP神经网络预测结果
Table 10 Predictive Results of Hybrid BP Neural Network Model

建模样本			检验样本				
实际类型	预测类型		总计	实际类型	预测类型		总计
	舞弊公司	正常公司			舞弊公司	正常公司	
舞弊公司	100	30	130	舞弊公司	44	18	62
正常公司	6	124	130	正常公司	15	47	62
舞弊判别成功率	76.92%			舞弊判别成功率	70.97%		
正常判别成功率	95.38%			正常判别成功率	75.81%		
总体判别成功率	86.15%			总体判别成功率	73.39%		

3.3 混合 BP神经网络模型预测

根据上文同样的输入变量和输出变量,应用 MATLAB7.0软件,根据 130对建模样本 ($T-1$) 年的财务指标和公司治理指标进行网络训练,构建混合 BP神经网络预警模型如下。

输入层到隐含层的权值矩阵为

$$W_{ih} = \begin{bmatrix} -0.0243 & 0.0216 & 0.0198 & -0.0125 & 0.0314 & -0.0312 \\ 0.2735 & -0.0055 & -0.0072 & 0.0005 & -0.0022 & 0.0267 \\ 0.0173 & 0.0022 & 0.0031 & -0.0111 & -0.0021 & -0.0025 \\ 0.0284 & 0.0025 & 0.0036 & -0.0157 & -0.0030 & -0.0033 \\ -0.2051 & 0.4988 & 0.2040 & -0.1842 & 0.1909 & 0.2703 \\ 0.6942 & 0.1851 & 0.1105 & -0.1498 & -0.1158 & -0.0211 \\ 1.0799 & -0.4861 & 0.0307 & -0.1508 & 0.1275 & -0.1694 \end{bmatrix} \quad (5)$$

隐含层到输出层的权值矩阵为

$$w_{ho} = [11.6237 \quad -1.5206 \quad 80.1322 \quad -42.3162 \quad -1.7922 \quad -1.0994 \quad 1.3553] \quad (6)$$

输入层到隐含层的阈值矩阵为

$$\theta_{ih} = [-77.6010 \quad 294.5256 \quad 28.8199 \quad 42.7225 \quad 673.7422 \quad 892.0175 \quad 644.0104] \quad (7)$$

隐含层到输出层的阈值矩阵为

$$\theta_{ho} = [-26.8454] \quad (8)$$

利用其余 62对检验样本对所构建的预测模型进行检验,模型对舞弊公司的判别成功率为 70.97%,对正常公司的判别成功率为 75.81%,总体判别成功率为 73.39%,可以有效地提前甄别信息披露舞弊的上市公司,如表 10 所示。

3.4 两种预警模型比较

本文根据同样的 130对建模样本,采用同样的输入变量,分别应用 Logistic回归模型和混合 BP神经网络建立上市公司信息披露舞弊的预警模型,并利用同样的 62对检验样本对两种模型进行检验,结果如表 11 所示。

表 11 预测结果比较

Table 11 Comparison of Predictive Results

样本	模型	误判率		成功率		
		I 类误差	II 类误差	舞弊判别成功率	正常判别成功率	总体判别成功率
建模 样本	Logistic 回归	40.00%	16.92%	60.00%	83.08%	71.54%
	混合 BP 神经网络	23.08%	4.62%	76.92%	95.38%	86.15%
检验 样本	Logistic 回归	37.10%	24.19%	62.90%	75.81%	69.35%
	混合 BP 神经网络	29.03%	24.19%	70.97%	75.81%	73.39%

显然, 不论是基于建模样本还是基于检验样本, 混合 BP 神经网络模型的预测成功率均高于 Logistic 回归模型。因为因变量与自变量观测的数据结构可能导致 Logistic 回归在系数估计上发生偏差, 常见的数据结构问题有多元共线性、过离散、空单元、完全分离和异常值等, 而混合 BP 神经网络模型具有强大的输入-输出映射能力, 而且信息容错能力强。在对信息披露舞弊概率的预警研究中, 混合 BP 神经网络模型具有传统的统计判别模型所无法比拟的优势, 它不需要确认财务指标或公司治理指标与公司运营之间的关系是线性还是非线性, 对指标的分布或彼此之间是否高度相关也无任何要求。

需要指出的是, 混合 BP 神经网络模型的黑箱效应使它无法判断预警变量对信息披露舞弊的影响作用是正向还是负向, 而 Logistic 回归模型可以较直观地体现这种作用。本文建议投资者、债权人和监管层将这两种方法相结合, 先运用 Logistic 回归模型分析关键因素对上市公司信息披露舞弊的影响方向, 再运用混合 BP 神经网络模型计算舞弊概率, 最后根据风险预警结果制定科学决策。

3.5 本文研究与国内同类研究的结果比较

据作者了解, 国内关于信息披露舞弊的预警研究或类似研究仅有鹿小楠和傅浩的研究报告^[10], 他们在 Beneish 的研究基础上, 以舞弊当年的 9 个财务指标为预警变量, 分别建立多元判别模型和 Logistic 回归模型, 但对舞弊公司的预测成功率都仅为 60%, 而且该研究并没有将建模样本与检验样本分开, 可能导致所得的 60% 预测成功率仍偏高于其实际预警能力。

与他们的研究相比, 本文基于公司治理视角, 在财务定量指标的基础上引入治理指标等定性信息, 建立的 Logistic 回归预警模型和混合 BP 神经网络模型对舞弊公司的预测成功率分别为 62.90% 和 70.97%, 提高了预警模型的有效性。由此可见, 对于投资者, 公司治理数据与财务数据一样具有信息价值, 而且混合 BP 神经网络等人工智能技术在财务领域

具有广阔的应用前景。

4 结论和启示

本文基于财务视角和公司治理视角, 根据逐步 Logistic 回归分析和因子分析对备选财务指标和治理指标进行提炼与筛选, 再运用 Logistic 回归模型和混合 BP 神经网络模型构建上市公司信息披露舞弊预警模型, 发现在舞弊被发现的前一个会计年度, 舞弊公司在财务状况和治理水平上就与正常公司存在着显著的差异; Logistic 回归模型可以直观地体现出预警变量对于舞弊的影响作用, 混合 BP 神经网络模型对舞弊公司的判别成功率为 70.97%, 总体判别成功率为 73.39%, 可以有效地提前甄别舞弊的上市公司, 充分体现了人工智能技术应用于财务领域的可行性和优越性。

为保护投资者利益和规范市场秩序, 本文建议沪深交易所等监管层逐步建立上市公司信息披露舞弊的风险预警机制。通过对上市公司财务数据和公司治理数据的收集, 逐步建立起上市公司财务状况和公司治理水平的动态监管, 并运用 Logistic 回归模型等统计方法对影响舞弊的风险因素进行分析, 再运用混合 BP 神经网络等人工智能方法对上市公司的舞弊风险水平实行分类管理, 把高风险的上市公司甄别出来, 实行重点监管, 从而合理有效地分配有限的监管资源, 引导并规范上市公司信息披露行为, 有效地保护投资者权益。

参考文献:

- [1] Coglitore F, Berryman R G. Analytical Procedures A Defensive Necessity[J]. Auditing: A Journal of Practice & Theory, 1988, 7(2): 150-163.
- [2] Persons O. Using Financial Statement Data to Identify Factors Associated with Fraudulent Financial Reporting [J]. Journal of Applied Business Research, 1995, 11(3): 38-46.
- [3] Beneish M D. The Detection of Earnings Manipulation

- [] . *Financial Analysts Journal* 1999 155(5): 24 - 36
- [4] Bell T B, Carcello J V. A Decision Aid for Assessing the Likelihood of Fraudulent Financial Reporting[J]. *Auditing* 2000 19(1): 169 - 184.
- [5] Eining M M, Jones D R, Loebbecke J K. Reliance on Decision Aids: An Examination of Auditors' Assessment of Management Fraud[J]. *Auditing* 1997 16(2): 1 - 18.
- [6] Green B B, Choi J H. Assessing the Risk of Management Fraud through Neural Network Technology[J]. *Auditing* 1997 16(1): 14 - 28.
- [7] Fanning K, Cogger K. Neural Network Detection of Management Fraud Using Published Financial Data[J]. *International Journal of Intelligent Systems in Accounting Finance & Management* 1998 7(1): 21 - 41.
- [8] Spahis Ch, Doumpos M, Zopounidis C. Detecting Falsified Financial Statements: A Comparative Study Using Multicriteria Analysis and Multivariate Statistical Techniques[J]. *The European Accounting Review* 2002 11(3): 509 - 535.
- [9] Lin J W, Hwang M J, Becker J D. A Fuzzy Neural Network for Assessing the Risk of Fraudulent Financial Reporting[J]. *Managerial Auditing Journal* 2003 18(8): 657 - 665.
- [10] 鹿小楠, 傅浩. 中国上市公司财务造假问题研究[J]. *上证研究*, 2003 (2): 36 - 95
- [11] Ferioz E H, Park K, Pastena. The Financial and Market Effects of the SEC's Accounting and Auditing Enforcement Releases[J]. *Journal of Accounting Research* 1991 29(Supplement): 107 - 148.
- [12] Healey P M. The Effect of Bonus Schemes on Accounting Decisions[J]. *Journal of Accounting and Economics* 1985 7(3): 85 - 107.
- [13] Jones J J. Earnings Management during Import Relief Investigations[J]. *Journal of Accounting Research* 1991 29(2): 193 - 228.
- [14] Beneish M D, Press E. Costs of Technical Violation of Accounting based Debt Covenants[J]. *Accounting Review* 1993 68(2): 233 - 257.
- [15] Siegel J G. How to Analyze Businesses Financial Statements and the Quality of Earnings[M]. Englewood Cliffs, NJ: Prentice-Hall, 1991.
- [16] National Commission on Fraudulent Financial Reporting. Report of National Commission on Fraudulent Financial Reporting[R]. New York: AICPA, 1987.
- [17] Lev B, Thiagarajan S R. Fundamental Information Analysis[J]. *Journal of Accounting Research* 1993 31(2): 190 - 215.
- [18] Grossman S, Hart O. The Cost and Benefit of Ownership: A Theory of Lateral and Vertical Integration[J]. *Journal of Political Economy* 1986 94(4): 691 - 719.
- [19] Shleifer A, Vishny R W. Large Shareholder and Corporate Control[J]. *Journal of Political Economy* 1986 94(3): 461 - 488.
- [20] Jensen M C. The Modern Industrial Revolution, Exit and the Failure of Internal Control Systems[J]. *Journal of Finance* 1993 84(3): 831 - 880.
- [21] Beasley M. An Empirical Analysis of the Relation between the Board of Director Composition and Financial Statement Fraud[J]. *Accounting Review* 1996 71(4): 443 - 465.
- [22] Uzun H, Samuel H S, Vam R. Board Composition and Corporate Fraud[J]. *Financial Analysts Journal* 2004 60(3): 33 - 43.
- [23] Dechow P M, Sloan R G, Sweeney A P. Detecting Earnings Management[J]. *Accounting Review* 1995 70(2): 193 - 225.
- [24] Klein A. Company Performance and Board Committee Structure[J]. *Journal of Law and Economics* 1998 41(1): 275 - 303.
- [25] Lipton M, Lorsch J A. A Modest Proposal for Improved Corporate Governance[J]. *Business Lawyer* 1992 48(1): 59 - 77.
- [26] Vafeas N. Board Meeting Frequency and Company Performance[J]. *Journal of Financial Economics* 1999 53(1): 113 - 142.
- [27] Jensen M C, Murphy K J. Performance Pay and Top Management Incentives[J]. *Journal of Political Economy* 1990 98(2): 225 - 263.
- [28] 白重恩, 刘俏, 陆洲, 宋敏, 张俊喜. 中国上市公司治理结构的实证研究[J]. *经济研究* 2005(2): 81 - 91.
- [29] Hart O. The Market Mechanism as an Incentive Scheme[J]. *Bell Journal of Economics* 1983 14(2): 366 - 689.
- [30] Jensen M C, Meckling W H. Theory of the Company: Managerial Behavior, Agency Costs, and Ownership Structure[J]. *Journal of Financial Economics* 1976 3(4): 305 - 360.
- [31] Watts R, Zimmerman J. Agency Problems, Auditing and the Theory of the Company: Some Evidence[J]. *Journal of Law and Economics* 1983 26(3): 613 - 633.
- [32] Gillan S, Stuart L, Starks L. Corporate Governance Proposals and Shareholder Activism: The Role of Institutional Investors[J]. *Journal of Financial Economics* 2000 57(2): 275 - 305.
- [33] Ajinkya B, Bhojraj S, Sengupta P. The Association between Outside Directors, Institutional Investors and the Properties of Management Earnings Forecast[J]. *Journal of Accounting Research* 2004 43(3): 343 - 376. □