

高管个人特征与公司业绩^①

——基于机器学习的经验证据

陆 瑶, 张叶青, 黎 波, 赵浩宇

(清华大学经济管理学院, 北京 100084)

摘要: 在目前的公司治理文献中, 大部分的高管特征研究一方面仅关注单一的高管特征与公司业绩之间的关联, 缺乏全面的高管特征分析; 另一方面主要围绕因果推断进行研究, 缺乏从预测能力出发的系统定量的结论。本文首次采用机器学习算法中的 Boosting 回归树, 全面考察了多维度高管特征对公司业绩的预测性。以我国 2008 年~2016 年的上市公司为样本, 研究了高管的多维个人特征是否能预测公司业绩, 并进一步分析了对公司业绩预测能力较强的高管个人特征及其预测模式。研究发现: 1) 整体而言, 在我国公司 CEO 和董事长的特征对公司业绩的预测能力较弱; 2) 在众多高管个人特征之中, 高管持股比例和年龄对公司业绩的预测能力较强; 3) 高管持股比例和年龄与公司业绩之间的关联都呈现出非线性的特点, 与以往的理论较为吻合。本研究不仅利用机器学习方法从一个更为全面的视角对中国的高管特征进行了研究, 也为公司高管聘任和激励机制设计等方面提供了有益的启发。

关键词: 机器学习; Boosting 回归树; 公司治理; 公司业绩

中图分类号: C93; F27 **文献标识码:** A **文章编号:** 1007-9807(2020)02-0119-21

0 引 言

“高管特征与公司绩效的关联”一直以来是公司治理领域的重点研究问题。以往研究一方面探究了高管是否在公司决策中担任了至关重要的角色, 另一方面挖掘出了可能会影响公司绩效的一系列高管的个人特征^[1,2]。那么公司高管的个人特征能够帮助预测公司业绩吗? 在这些可能预测公司绩效的高管特征中, 哪些特征更为关键呢? 它们的预测机制又是怎样的? 以往研究主要围绕单一特征与公司业绩之间的因果关系推断进行研究, 而缺乏从预测能力出发的系统定量的结论。本文试图通过运用机器学习的方法来弥补这一空白。

目前, 大部分公司金融实证研究的目的是推

断变量之间的因果关系。因果解释性模型从假设性理论出发, 利用统计模型来验证理论中可观测变量之间的因果关系。这一类建模过程被称为解释性建模 (explanatory modeling), 其中最常用的模型是线性回归模型和结构方程模型。与解释性建模不同, 预测性建模 (predictive modeling) 的目的是准确预测未来的观测值。

预测性分析虽然较少出现在公司金融 (特别是公司治理) 的实证研究中, 但在学术研究中同样至关重要。首先, 预测性模型没有假设变量之间因果关系, 而大部分拟合效果好的模型也没有假设变量之间特定的函数形式 (比如线性关系、U 型关系、指数关系等); 相对地, 解释性模型预先对因果关系提出假设, 然后利用数据来进行检验。因此, 预测性模型能够发掘出数据中更为复

① 收稿日期: 2019-05-24; 修订日期: 2019-12-14。

基金项目: 国家自然科学基金优秀青年基金资助项目 (71722001); 清华大学自主科研计划课题项目 (2015THZWY09; 2019THZWJC11); 国家自然科学基金资助项目 (71490723; 71432004); 教育部人文社会科学重点研究基地项目 (16JJD630006)。

作者简介: 陆 瑶 (1978—), 女, 北京人, 副教授, 博士生导师。Email: luyao@sem.tsinghua.edu.cn

杂的规律,有助于提出并验证新的理论来解释数据中存在的规律,从而推动解释性模型和理论的进步。其次,预测效果提供了一个评估解释性模型的新角度^[3],衡量每个变量对预测表现的贡献度能刻画该变量的重要程度,进而检验以往理论的可靠性。第三,预测性模型帮助研究者比较不同的解释性模型。最后,预测能力的强弱程度能够反映理论解释现象的能力的强弱程度^[4]。如果某一个解释性模型的预测效果特别差,则说明其所依托的理论不足以解释现象,该理论仍存在较大的发展空间。但值得一提的是,虽然预测性模型和解释性模型分别从“预测”和“理解”的角度出发,但两者之间并非是完全对立的。本文**预测性模型的设计基于一系列相关经济金融学文献,而非单纯从数据中凭空构建**。因此,本文与以往理论和实证研究的关系密不可分、相辅相成。

绝大多数对于高管特征与公司业绩的研究仅仅关注了两者之间的因果关联,而忽视了高管特征对公司业绩的预测能力。解释性模型和预测性模型从概念和优化目标上存在很大差异^[5,6]②,对公司业绩有解释力的高管指标并不一定能准确预测公司业绩,因此无法从因果推断模型的估计结果中得出预测效果的结论。**本文的研究目标是通过直接探究高管特征对公司业绩的预测效果来弥补这一学术空白。**

那么,为什么要采用机器学习的方法来解答这一问题呢?首先,改革开放以来,我国经历了从中央高度集中的计划经济体制到中国特色社会主义市场经济的转轨阶段。从公司的股权结构来看,国有企业高管的选拔和任命往往出于政治性因素的考虑,而国企高管面临的非市场化的晋升考核体制也使得他们在任职期间的目标不是企业的经营业绩,而是政治诉求^[7-8]。在转型期经济环境下,国有企业的改革不断推进,中国特色社会主义市场经济体制逐步建立,经济环境出现明显波动。经济环境的波动可能导致高管特征与公司绩效之间的关联不是一成不变的,而是随经济环境波动而动态变化的。以往研究最常使用线性

拟合模型,该模型假定变量之间的相关性是恒定的,并在此基础上进行模型估计,因此难以在经济环境波动的情形下得到可靠准确的预测模型。第二,已有文献发现高管的个人特征与公司绩效之间的非线性关系以及这些特征之间的交互作用③,这些研究意味着传统的简单线性拟合模型难以清晰、准确地厘清变量之间的复杂关系。基于此,为了全面挖掘高管特征与公司绩效之间关联,本文引入处理复杂预测问题的机器学习模型,试图针对这一公司治理的传统问题提供一个更为全面的解答。

本文以2008年~2016年的上市公司为样本,实证评估了高管特征对公司绩效的预测能力,进一步挖掘了对公司业绩预测能力较强的高管的个人特征,并刻画了它们的预测机制。本文采用的机器学习方法为**“Boosting 回归树”**。**该方法的基本思想是:从初始训练集中得到一个基回归树,然后在当前预测误差的基础上训练新的基回归树,每次迭代都向损失函数负梯度的方向移动,从而达到损失函数随着迭代次数增加而逐渐减小的效果,最后加权结合多个基回归树得到回归函数。进一步地,基于训练模型分析了各个高管特征在预测公司业绩中的重要程度和预测效果。**研究结果发现:整体而言,在我国公司CEO和董事长的特征对公司业绩的预测能力较弱。而在众多高管特征之中,持股比例和年龄对公司业绩的预测能力较强,且它们与公司业绩之间的关系呈现出明显的非线性的特点。在更换了拟合的滚动窗口期、公司绩效的衡量指标、Boosting 模型的参数、其他机器学习算法后,上述结果仍然稳健。

与现有文献相比,本文的贡献如下:第一,首次应用机器学习方法来研究中国的公司治理问题,评估了公司高管特征整体对于公司绩效的预测能力,本文的研究丰富了“高管特征与公司绩效的关联”这一领域的文献。第二,采用前沿的Boosting 回归树的方法,规避了传统线性模型的缺陷,更好地分析了变量之间的非线性和交互关

研究结果

研究意义

② 从概念方面,解释性模型基于理论中变量的因果关系,而预测性模型则基于可测变量之间的关联;从优化方法的角度,解释性建模的目标是最小化模型偏差以准确地刻画潜在的理论模型,而预测性建模的目标是同时最小化模型偏差和样本方差。模型偏差和样本方差的最小化之间存在一定的权衡关系,因此预测准确性的提高需要牺牲一定的理论解释性。

③ 比如 Adams 等^[9]发现管理层的特征与组织变量的相互作用对公司表现有显著影响。

系.第三,探究了不同高管特征对于预测公司绩效的重要程度,并分析了相对重要的高管特征对公司绩效的预测机制,这一系列结论对选聘公司高管具有非常重要的启示意义.

1 文献综述

1.1 高管特征与公司业绩

在现代公司制度下,高管是公司的核心决策者,因此“高管特征与公司业绩的关联”是经济、金融和管理学中的经典研究问题.早期的实证研究发现,高管的个人特征能够很大程度上解释公司的资本结构、投资决策和组织架构^[2,10].后续的一系列研究从高管特征的不同角度出发,验证了多维度的高管特征对公司业绩的影响.这些高管特征的维度可以被归纳为以下五类:先天特质、人生经历、个性特征、能力水平以及管理风格.同时,公司组织结构也可能影响高管作用.

先天特质方面,已有文献从高管的性别、年龄、容貌等角度研究. Kaplan 和 Sorensen^[11]的研究发现男、女性高管在管理风格上没有显著差异.高管年龄的影响主要体现在年龄较大的高管的管理风格更加保守^[2,12].

高管的人生经历一方面是高管个人特征的反映,另一方面也可能会通过影响他们的风险偏好、专业技能、对行业发展的预期来影响公司决策的制定.这部分研究的切入点主要包括幼年生活经历^[13,14]、家庭生活^[15,16]、职业经历^[2,12,13,17]和教育经历^[2,12]等.

高管的性格特点也会影响公司决策和公司绩效.比如过度自信的高管可能会高估自己的能力和公司的收益,从而制定更激进的投资和财务决策,特别是公司的研发创新投入,进而影响公司绩效^[18-20].此外,高管是否乐于寻求刺激^[21]和他们的风险态度^[22-24]也会影响他们的管理风格.

近年来,针对高管个人能力的研究也非常丰富,受到研究者关注的高管个人能力主要包括基本能力、沟通交流能力和执行能力等.个人能力本身是难以度量的,因此越来越多的研究通过非结构化数据和机器学习技术来构建高管的多维能力指标,并分析这些指标与公司绩效之间的关联^[11,25,26].

与上述维度相比,高管的管理风格是内生性更强的指标.一方面,管理风格受到高管个人经历、性格、专业水平等特质的影响;另一方面,管理风格本身难以观测和度量.比较典型的研究是 Bandiera 等^[27]利用高管的日程记录和非监督机器学习算法,刻画了“领导型”和“管理型”两种管理风格,并发现领导型高管所在的公司的绩效更佳.

此外,公司的组织结构也是影响高管作用的重要因素.例如,管理层的薪酬制度和高管持股情况通过影响高管工作的努力程度和公司治理水平来影响公司绩效^[24,28,29].创始人高管以及家族企业中的高管对公司绩效的影响也有所不同,比如创始人高管会有更多的研发投入、资本支出和兼并收购活动^[30,31]. Khanna 等^[32]发现高管在企业内部的裙带关系会影响上市公司违规犯罪的概率.

关于中国上市公司高管与公司业绩之间关联的研究也层出不穷,但限于数据的可得性,国内的研究大多立足于一些可观测的高管特征指标,并且更多地考虑了中国独特的公司体制和经济环境.具体而言,中国上市公司高管特征的分析大多围绕高管的职业经历展开,比如有政府任职经历的高管会更多地参与非生产性活动,从而降低资源配置效率,损害公司业绩^[33],但是民营公司高管的政府背景对公司价值没有显著影响^[34].高管的学术经历能帮助降低公司的债务融资成本^[35],但人大代表或政协委员的经历会提高发债成本^[36].程新生和赵畅发现权威专业董事能够促进企业创新活跃度^[37].此外,徐莉萍等^[38]从个人家庭生活经历的角度入手,发现高管离婚后公司盈利能力和市值水平都有所下降.风险投资和创始人持股会对公司治理效果产生影响^[39].此外,CEO 与董事之间的关联关系也受到学者的关注,比如 CEO 对董事会的影响力的提高使得 CEO 的违规倾向增加^[40],CEO 与董事之间存在“老乡”关系会提高公司违规的可能性和公司的财务风险^[41,42].而不同类型的董事(咨询董事或监督董事)与 CEO 的互动关系也有所差异^[43].

从上述研究可知,以往关于高管特征的研究主要围绕某一类个人特质,探讨其对公司决策及业绩的影响,缺少全面比较公司高管特征相对

重要性的研究. 此外, 我国特有的制度和文化环境也使得国内的高管研究难以直接借鉴西方的研究结论, 需要更加综合地考虑高管特征对公司业绩的影响. 这为本文的研究提供了难得的契机. 在已有文献的基础上, 本文结合了各角度的高管特征来探讨哪些特征与公司业绩有更为显著的关联.

1.2 机器学习与公司金融研究

机器学习作为一种新兴的统计工具, 越来越多地受到研究者的关注. 目前, 已有研究从两个角度尝试将机器学习应用到公司金融的研究领域: 无监督学习 (unsupervised learning) 和有监督学习 (supervised learning).

在无监督学习中, 训练样本并不对应标记信息, 因此它的目标是通过训练总结、刻画出多维数据的内在关联, 进一步将高维甚至非结构化的变量转化为低维的、可解释性强的变量. 这一方法能够从复杂的数据集中归纳、生成本身难以观测的变量, 从而帮助解决公司金融研究中常面临的关键变量难以精确度量的问题. 典型的研究包括 Bandiera 等^[27] 利用无监督学习中的潜在狄利克雷配置方法来分析 CEO 日常活动记录的数据, 刻画出了“领导型”和“管理型”两种 CEO 的行为特征, 并进一步发现“领导型”CEO 与好的公司业绩更相关. Li 等^[44] 采用单词植入模型, 构建了五个公司文化价值的指标, 并研究了其对公司行为决策的影响. 总体而言, 无监督学习有助于基

于大数据来构建难以直接观测和衡量的指标, 大大拓展了公司金融中可研究的领域.

有监督学习主要解决从特征变量 (x) 到结果变量 (y) 的预测问题. 与传统的线性回归模型相比, 机器学习模型能够应对变量之间更加复杂的非线性和交互关系, 在样本外达到更好的拟合效果^[45]. 近年来, 有监督学习已经在资产定价领域备受关注^[46-48], 但在公司金融领域的应用较少. 只有 Gow 等^[49] 基于 CEO 个性特征来预测公司样本外的特征, 包括融资决策、投资决策以及经营表现. 简言之, 有监督学习提高了传统方法的预测精度, 在公司金融领域的应用仍有广大的探索空间.

2 数据来源和变量说明

2.1 数据来源

本文选取的样本为 2008 年 ~ 2016 年所有的中国 A 股上市公司, 研究对象为上市公司的最高管理层——CEO 和董事长. 数据来源为国泰安经济金融研究数据库 (CSMAR) 和锐思数据库 (RESSET). 其中 CEO 和董事长的个人特征数据来自 CSMAR 从 2008 年起开始数据收录的“上市公司人物特征”子数据库. 剔除了 CEO 特征与董事长特征存在缺失的样本点, 同时保证了 CEO 和董事长个人特征数据的完整, 最终保留样本 18 170 个, 平均占各年度全部上市公司的 85.06%. 表 1 报告了样本分布情况.

表 1 样本分布
Table 1 Sample distribution

年份	样本量	沪深上市公司数量	样本占上市公司总数比 (%)	保证 CEO 数据完整 的公司数量	保证董事长数据完整 的公司数量
2008	1 208	1 603	75.36	1 252	1 548
2009	1 406	1 752	80.25	1 439	1 717
2010	1 751	2 107	83.10	1 791	2 087
2011	2 092	2 341	89.36	2 128	2 327
2012	2 348	2 470	95.06	2 398	2 463
2013	1 895	2 515	75.35	1 919	2 497
2014	2 271	2 632	86.28	2 326	2 598
2015	2 477	2 823	87.74	2 541	2 781
2016	2 722	3 118	87.30	2 788	3 073
总计	18 170	21 361	85.06	18 582	21 091

2.2 变量定义

本文从盈利能力和价值这两个角度对目标变量——公司绩效进行衡量. 公司的盈利能力对应公司的财务绩效, 公司的价值对应公司的市场绩效. 具体而言, 本文采用行业中值调整后的资产收益率($adEBITDAper$)作为财务绩效指标, 采用行业中值调整后的托宾 Q 值($adTobinQ$)作为市场绩效指标. 其中资产收益率定义为息税折旧摊销前利润除以期末总资产; 托宾 Q 值的定义为公司总市值除以总资产. 考虑到公司业绩受行业因素和宏观经济因素影响, 在实际分析过程中, 本文基于行业中值对这两个公司绩效指标进行了调整, 即实际绩效指标减去公司所属行业所有公司当年该绩效指标的中位值.

CEO 和董事长的多维度个人特征是本文的核心变量. 本文结合文献和中国 A 股上市公司数据可得性两个方面来选取相关变量. 在文献综述部分, 将高管特征的维度归纳为先天特质、人生经历、个性特征、能力水平、管理风格, 以及公司的组织结构. 其中, 先天特质包括性别、年龄和容貌等高管与生俱来的种种特征, 而在这些特征当中, 能够被客观测度并可得变量包括性别和年龄; 人生经历和能力水平可以从高管的职业经历以及教育经历来切入, 本文从生产经营、市场管理、财务法律、金融领域的职业经验来衡量高管在各个领域的专业水平, 同时海外经历以及研究经历也是本文关注的指标. 此外, 考虑到国有企业在中国经济中发挥的影响, 本文认为高管的政府工作经历也可能是预测公司绩效的重要指标. 组织结构方面, 采用高管持股和兼职情况来对公司绩效进行预测. 而由于个性和管理风格方面的特征难以观测和准确度量, 暂时未将其纳入考虑, 但这会是下一步的重要研究方向.

具体而言, 本文选取的高管特征包括年龄、性别、话语权、年末持股比例、公司外兼职、职能经验、海外经验、学术经验、金融工作背景和政府工作背景. 其中, 话语权由 CEO 和董事长是否两权分离来衡量; 职能经验涵盖三个方面, 第一是生产、技术和设计岗位经验, 第二是市场、人力、管

理岗位经验, 第三是财务、法律职能岗位经验, 这三方面职能经验描述了高管主要的职场经验和能力; 海外经验反映该 CEO 或董事长是否有海外工作或求学经历; 学术经验反映该 CEO 或董事长是否有学术研究工作经历; 金融背景反映该 CEO 或董事长是否曾任职于金融机构^④; 政府工作背景反映该 CEO 或董事长是否曾任职于政府部门, 衡量其政治关联度.

在基准模型中, 公司层面的变量包括: 公司规模、公司寿命、杠杆率、重资产率和股权性质. 其中, 公司规模由公司总资产规模衡量; 杠杆率由上市公司总负债与总资产的比率衡量; 重资产率由上市公司固定资产净额与总资产的比率衡量; 股权性质指的是国有股份占比. 上述所有变量的具体定义, 均在表 2 中列明.

从探究因果关系的角度出发, 上述变量之间存在明显的内生性问题, 而对变量之间内生性的处理方式也是本文与以往大部分高管相关研究相比的重要区别之一. 内生性指的是在解释性模型中, 解释变量与残差项之间的相关性, 通常由遗漏解释变量、反向因果和测量误差导致^[4]. 解释性建模旨在识别模型和证明因果关系, 因此内生性的存在会使得参数估计有偏. 然而, 与证明因果关系的解释性建模不同的是, 本文的预测性建模并不关注因果关系, 而是相关关系, 所以变量之间内生性的存在并不会影响对预测效果的评估^[4].

2.3 描述性统计

表 2 分别列出了被预测变量、CEO 特征、董事长特征和公司层面其他变量的描述性统计, 包括样本量、均值、标准差和中位数. 为了防止极值带来的统计偏差, 本文对连续型变量进行了 1% 和 99% 的 Winsorize 处理.

从高管的个人特征变量的描述性统计结果(面板 B 和面板 C)中, 可以发现: 首先, 近四分之一(24.96%)样本中 CEO 和董事长由同一人担任; 第二, 男性在公司“一把手”职位中占绝对优势, 女性 CEO 仅占 6.02%, 而女性董事长仅占 4.53%; 第三, 我国上市公司 CEO 在其他公

^④ 此处统计中的金融机构包括监管部门、政策性银行、商业银行、保险公司、证券公司、基金管理公司、证券登记结算公司、期货公司、投资银行、信托公司、投资管理公司、交易所等.

司兼职的比例高于董事长, 超过一半 (52.92%) CEO 在其他公司兼职, 董事长在其他公司兼职的比例仅为 26.51%; 第四, 我国上市公司 CEO 和董事长在职能背景方面具有类似的特征, 平均 25% 至 30% 的 CEO 或董事长具有生产、技术、设计等职能经验, 约 18% 的 CEO 或董事长具有财

务、法律等专业职能经验, 超过 90% 的 CEO 或董事长具有市场、战略、人力等管理职能经验; 第五, CEO 或董事长具有海外工作、求学经历的比例较低, 仅约 5%; 最后, 我国上市公司的董事长比 CEO 具备更丰富的政府工作、学术研究和金融工作经历。

表 2 主要变量的定义和描述性统计

Table 2 Definitions of variables and descriptive statistics

变量名	定义	样本量	均值	标准差	中位数
Panel A 公司绩效					
<i>adEBITDAper</i>	经过行业中值调整的 <i>EBITDA</i> 占总资产的比例	18 170	0.015 7	0.119 1	0.000 3
<i>adTobinQ</i>	经过行业中值调整的托宾 <i>Q</i> 值	17 300	0.342 0	1.518 3	0.006 8
Panel B CEO 个人特征					
<i>Duality</i>	CEO 和董事长是否两职分离, 1 为是	18 170	0.249 6	0.432 8	0.000 0
<i>Gender. c</i>	是否为女性, 1 为是	18 170	0.060 2	0.237 8	0.000 0
<i>Age. c</i>	年龄的对数	18 170	3.871 0	0.135 5	3.871 2
<i>Share. c</i>	年末持股比重	18 170	0.048 7	0.115 1	0.000 0
<i>Parttime. c</i>	是否在其他公司兼职, 1 为是	18 170	0.529 2	0.499 2	1.000 0
<i>ProFun. c</i>	是否有生产、技术、设计职能经验, 1 为是	18 170	0.294 1	0.455 6	0.000 0
<i>MgtFun. c</i>	是否有市场、战略、人力管理职能经验, 1 为是	18 170	0.996 6	0.057 8	1.000 0
<i>SkiFun. c</i>	是否有财务、法律职能经验, 1 为是	18 170	0.186 1	0.389 2	0.000 0
<i>Oversea. c</i>	是否有海外工作、求学经历, 1 为是	18 170	0.054 4	0.226 8	0.000 0
<i>GovBack. c</i>	是否有政府工作经历, 1 为是	18 170	0.186 8	0.389 8	0.000 0
<i>Academic. c</i>	是否有学术研究经历, 1 为是	18 170	0.164 6	0.370 8	0.000 0
<i>FinBack. c</i>	是否有金融行业工作经历, 1 为是	18 170	0.057 9	0.233 6	0.000 0
Panel C 董事长个人特征					
<i>Duality</i>	CEO 和董事长是否两职分离, 1 为是	18 170	0.249 6	0.432 8	0.000 0
<i>Gender. b</i>	是否为女性, 1 为是	18 170	0.045 3	0.208 0	0.000 0
<i>Age. b</i>	年龄的对数	18 170	3.942 4	0.137 4	3.951 2
<i>Share. b</i>	年末持股比重	18 170	0.077 5	0.141 2	0.000 0
<i>Parttime. b</i>	是否在其他公司兼职, 1 为是	18 170	0.265 1	0.441 4	0.000 0
<i>ProFun. b</i>	是否有生产、技术、设计职能经验, 1 为是	18 170	0.248 3	0.432 1	0.000 0
<i>MgtFun. b</i>	是否有市场、战略、人力管理职能经验, 1 为是	18 170	0.925 4	0.262 7	1.000 0
<i>SkiFun. b</i>	是否有财务、法律职能经验, 1 为是	18 170	0.172 0	0.377 4	0.000 0
<i>Oversea. b</i>	是否有海外工作、求学经历, 1 为是	18 170	0.053 7	0.225 4	0.000 0
<i>GovBack. b</i>	是否有政府工作经历, 1 为是	18 170	0.350 0	0.477 0	0.000 0
<i>Academic. b</i>	是否有学术研究经历, 1 为是	18 170	0.243 3	0.429 1	0.000 0
<i>FinBack. b</i>	是否有金融行业工作经历, 1 为是	18 170	0.086 3	0.280 8	0.000 0
Panel D 公司层面变量					
<i>Ln Asset</i>	总资产的对数	18 170	21.868 1	1.395 2	21.690 7
<i>State. share</i>	国有股份占比	18 170	0.063 5	0.155 9	0.000 0
<i>Leverage</i>	资产负债率	18 170	0.435 5	0.221 8	0.428 6
<i>PPE</i>	固定资产占总资产的比例	18 170	0.225 5	0.171 6	0.190 3
<i>Esty</i>	公司寿命的对数	18 170	2.719 2	0.381 1	2.772 6

3 研究方法和模型构建

3.1 Boosting 回归树

本文采用的机器学习方法为渐进梯度回归树 (gradient boosting regression trees, GBRT)^⑤, 简称“Boosting 回归树”^[51,52]. Boosting 回归树有效地结合了机器学习中的决策树和集成思想, 通过加权多个基回归树来提高拟合效果, 在样本内和样本外的拟合精确度都非常高^[53].

机器学习方法的目标是找到一个回归函数 $\hat{f}(x)$ 来最小化损失函数 $\psi(y, f)$ ^⑥ 的期望值

$$\hat{f}(x) = \arg \min_{f(x)} E_{y|x} [\psi(y, f(x)) | x] \quad (1)$$

而 Boosting 的基本思想是: 从初始训练集中得到一个基回归树, 然后在当前预测误差的基础上训练新的基回归树, 迭代多次, 最后加权结合多个基学习器得到最后的回归函数 $\hat{f}(x)$. 其中, 梯度提升回归树 (gradient boost regression tree, GBRT) 是常用的 Boosting 算法之一, 它的关键点是通过梯度提升来得到每一步的基回归树, 使得模型在每次迭代后都向损失函数负梯度的方向移动, 达到损失函数随着迭代次数增加而逐渐减小的效果. 借鉴 Friedman^[52], 本文采用的具体算法步骤如下: 首先, 设定初始的回归函数

$$\hat{f}(x) = \arg \min_{\rho} \sum_{i=1}^N \psi(y_i, \rho) \quad (2)$$

即使得误差平方和最小的常数. 计算损失函数的负梯度

$$z_i = - \left. \frac{\partial \psi(y_i, f(x_i))}{\partial f(x_i)} \right|_{f(x_i) = \hat{f}(x_i)} \quad (3)$$

利用损失函数负梯度在当前模型的值 (z_i) 作为残差的近似值, 基于 x_i 拟合出新的回归树模型

$$g(x) = E(z | x) \quad (4)$$

选择使得误差最小的梯度下降幅度

$$\rho = \arg \min_{\rho} \sum_{i=1}^N \psi(y_i, \hat{f}(x_i) + \rho g(x_i)) \quad (5)$$

进而得到经过本轮迭代后新的预测函数

$$\hat{f}(x) \leftarrow \hat{f}(x) + \rho g(x) \quad (6)$$

然后循环以上迭代过程 ($t = 1, \dots, T$).

Boosting 回归树的估计过程主要涉及三个参数: 回归树的交互深度 (interaction depth)、收缩参数 (shrinkage parameter) 和回归树的数量. 在本文的模型中, 参数的调整优化 (tuning parameters) 仅建立在样本内信息之上, 而最优参数的搜索范围则是根据经验法则和运算成本来确定的.

交互深度指的是每个基回归树有几次分叉. 根据经验法则和算力限制设置了最大交互深度, 即最优交互深度的搜索范围. 较多的分叉意味着基回归树能考虑到更复杂的交互性, 但更容易过度拟合. 同时, 过高的交互深度也会带来较大的计算成本. Hastie 等^[54]认为在多数问题中最优交互深度都较低, 因此本文的研究将交互深度的上限设置为 6.

收缩参数指的是被加入模型中的新的基学习器的权重. 根据以往基于 GBRT 模型的研究经验^[50,54], 收缩参数一般设置为 0.01 或 0.001, 且在算力允许的情况下取值尽可能小. 越小的收缩参数越可能带来更优的预测效果, 但本文并没有尝试小于 0.001 的收缩参数, 主要原因是更小的收缩参数需要更高的运算性能, 会极大地提高计算成本 (包括存储成本和 CPU 运算时间), 且对预测的边际提升效果很小^[54], 容易过度拟合^[52]. 因此, 本文取收缩参数值为 0.001.

最优的收缩参数和最优的回归树数量是相互影响的. 基回归树的数目与收缩参数之间需要一定的平衡, 小的收缩参数往往需要更多的基学习器. 本文的模型中基回归树的数目上限设置为 10 000.^⑦在每一次滚动的训练样本中, 采用 5 折交叉验证 (cross-validation) 的方法来确定最优的基回归树的数目. 上述模型参数的优化过程完全没有用到样本外的信息.

作为一种集成的树形回归算法, Boosting 虽然可以显著提升预测准确度, 但模型的可解释性

⑤ 本文使用的 Boosting 算法是 Ridgeway^[50]使用的 R 语言工具包“gbm”, 参数设置如下: distribution = “gaussian”, n. trees = 10 000, interaction. depth = 6, shrinkage = 0.001. 在稳健性检验部分验证了参数调整对预测效果的影响.

⑥ 本文选用的损失函数是平方误差 (squared error), 这是连续性结果变量的分析中最常用且在数学上更易于处理的损失函数^[49,50].

⑦ 本文分析部分所涉及的拟合过程取得最佳优化效果时的回归树的数量通常处于 3 000 至 5 000 之间.

却不尽如人意. 为了解决这一问题, 采用了相对重要性 (relative importance) 和部分依赖图 (partial dependence plot) 来进一步挖掘模型背后的经济学内涵. 本文的目的并非仅仅是预测, 更是试图回答哪些变量对预测效果影响较大以及预测模式如何. 这些问题的回答一方面能够有助于验证以往文献中“因果关系”的可靠性以及预测模型的可靠性, 比如本文通过部分依赖图关系图印证了高管持股比例与公司业绩之间存在着非线性的复杂关系; 另一方面能够推动对变量之间关系的理解, 从而提出新的理论^[4]. 上述两点意义对于正在形成期的理论而言尤为重要^[55,56]. 如果仅仅关注预测效果, 而忽视预测模型的可解释性, 那么整个模型就变成了一个“黑箱模型”, 内在预测机制被掩盖, 无法带来更多启发.

一个变量的相对重要性指的是一个变量在模型拟合过程中, 相对于模型中其他变量的重要程度. 根据 Friedman^[52], 将变量重要性的定义为给定模型中其他部分不变, 在模型中加入该变量带来的平方误差的下降幅度. 换言之, 相对重要性即在模型中其他部分不变的前提下, 在模型中去掉该变量或用一随机数变量将该变量替代后, 模型拟合能力的恶化程度相对于在其他变量上做相同操作后模型拟合能力恶化程度的大小. 如果一个变量对应的恶化程度相对较大, 则其在模型拟合过程中的相对重要性较大. 本文通过计算相对重要性来回答高管的高维特征之中, 哪些特征对于公司绩效的预测效果贡献更大的问题. 为了便于比较, 模型中所有变量的相对重要性指标之和被标准化为 1.

本文希望基于 Boosting 模型的估计挖掘出一些解释性意义, 因此进一步引入部分依赖图来分析单个变量对于预测结果的影响模式. 某一特征变量对于被预测变量的部分依赖关系是在控制其他特征为观测值的情况下, 衡量这一特征变量值的变动对模型拟合效果的影响. 这一指标能够在一定程度上解决 relative importance 无法反映正负关系的问题. 利用部分依赖图, 可以考察某一高管特征对公司绩效的预测模式

$$\hat{f}_{x_s}(x_s) = E_{x_c} [\hat{f}(x_s, x_c)] = \int \hat{f}(x_s, x_c) dP(x_c) \quad (7)$$

模型(7)中 x_s 为某一高管特征变量, $\hat{f}_{x_s}(x_s)$ 刻画在 x_s 的每个取值下, 被预测变量对应的估计值, x_c 为模型中除 x_s 之外高管的其他特征变量

$$P(x_c) = \int p(X_{all}) dx_s \quad (8)$$

式(8)中 $\int p(X_{all})$ 是模型中所有特征变量的联合概率密度. 部分依赖关系可以由下述方法近似估计^[52]. 在 x_s 的特定取值下, 对于 Boosting 回归树生成的每一棵树 b , 可以计算

$$\hat{f}_{x_s}(x_s) = \frac{1}{n} \sum_{i=1}^n \hat{f}(x_s, x_c^{(i)}) \quad (9)$$

式中 $x_c^{(i)}$ 为模型中 x_c 在训练集中的第 i 组观察值 (共 n 组). 上述的计算过程把 n 组观察值中的目标变量取为特定取值 x_s , 并控制其他变量为数据集中的观测值. 在特定取值 x_s 下, 对 Boosting 回归树生成的 B 棵树进行上述公式的计算, 取均值后得到 x_s 对应的被预测变量的值 y_s 的近似估计

$$\hat{y}_s(x_s) = \frac{1}{B} \sum_{b=1}^B \hat{f}_{x_s}(x_s)_b \quad (10)$$

计算出 x_s 的每一个不同取值对应的 \hat{y}_s 后, 就可以绘制出 x_s 对于被预测变量的部分依赖图: 横坐标为给定的特征变量值 (x_s); 在给定特征变量值下, 控制其他特征为数据集中各自的观测值, 纵坐标为公司业绩预测值的平均值 $\hat{y}_s(x_s)$.

本文选择 Boosting 回归树作为主要研究方法有以下两点原因. 1) Boosting 回归树有较强的拟合能力. 根据 Hastie 等^[54]的论述, 在众多统计学习方法中, 基于“树”的学习器具有易于构建、可解读、对变量的严格单调变换保持不变、不受极端值影响等优点. 另外, 回归树本身能够进行变量筛选, 因此即使加入不相关的变量, 也不会对结果造成影响. 但是, 回归树是一种弱学习器, 单独使用会得到不精确、波动大的结果. 因此, Boosting 回归树通过对集成许多基回归树, 能在尽量保持其固有优势的情况下, 显著提升其准确性. 2) Boosting 回归树有较好的可解释性. 大多数有监督的机器学习方法的唯一目的是预测, 因此往往会为了提高模型的复杂度和预测准确性而牺牲模型的可解释性^[57], 导致了绝大部分机器学习方法都是“黑箱模型”, 极大地限制了机器学习在

公司金融领域的应用. 而 Boosting 回归树是机器学习中为数不多的解释性较强的模型^⑧, 不仅可

以提高模型的预测性, 而且能通过计算变量的相对重要性和绘制部分依赖图来促进对理论的探索.

表 3 模型构建

Table 3 Model construction

被预测变量	模型	预测特征变量
公司绩效	基准模型	$\ln Asset, Leverage, State. Share, PPE, Esty$
	CEO 模型	$Duality, Gender. c, Age. c, Share. c, Parttime. c, ProFun. c, MgtFun. c, SkiFun. c, Oversea. c, GovBack. c, Academic. c, FinBack. c, \ln Asset, Leverage, State. Share, PPE, Esty$
	董事长模型	$Duality, Gender. b, Age. b, Share. b, Parttime. b, ProFun. b, MgtFun. b, SkiFun. b, Oversea. b, GovBack. b, Academic. b, FinBack. b, \ln Asset, Leverage, State. Share, PPE, Esty$

注: 公司绩效指标包括财务绩效 $adEBITDAper$ 和市场绩效 $adTobinQ$. 各变量具体含义请见表 2.

3.2 模型构建

本文对中国上市公司高管的个人特征进行综合考量. 首先构建了只包含公司基本特征但不含公司高管人物特征的基准模型 (以下简称“基准模型”). 基准模型仅由五个描述公司特征的控制变量组成, 包括公司规模 ($\ln Asset$)、杠杆率 ($Leverage$)、国有股份占比 ($State. Share$)、重资产比例 (PPE) 和公司寿命 ($Esty$). 在基准模型的基础上, 分别加入 CEO 特征族和董事长特征族, 进一步构建了两个包含高管人物特征的模型 (以下简称“CEO 模型”和“董事长模型”). 三个模型如表 3 所示.

基于上述模型, 本文的研究设计主要按照以下几个步骤进行:

首先, 按照一年滚动窗口期的拟合结构, 用 Boosting 回归树分别对三个模型进行拟合, 在样本内数据中通过参数调整来优化拟合能力, 并将优化后的拟合模型在样本外数据中测试, 评估拟合效果. 一年滚动窗口期是指以单年数据为训练集, 次年数据为测试集, 如此不断滚动拟合 T 次 ($T=8$). 具体而言, 首先使用 2008 年数据对模型进行训练和交叉验证, 并用模型预测 2009 年的公司绩效; 接着使用 2009 年数据进行训练和交叉验证, 并在 2010 年数据上进行预测; 依次向后续年份滚动推进. 由训练集的拟合结果可计算得样本内拟合优度 R_{is}^2 ; 由测试集的预测结果可计算得模型的样本外拟合优度 R_{oos}^2 和样本外均方误差 MSE_{oos} . 以 R_{is}^2 、 R_{oos}^2 和 MSE_{oos} 为主要指标,

评估 CEO 模型和董事长模型在预测公司绩效的能力上与基准模型相比分别有何提升. 以 R_{is}^2 为例, 鉴于不同年份数据之间存在差异, 计算整体 R_{is}^2 时取每一年拟合程度的均值, R_{is}^2 计算方法为

$$R_{is}^2 = 1 - \frac{\sum_{t=1}^T SSE_t}{\sum_{t=1}^T SST_t} \quad (11)$$

式中 SSE_t 为第 t 期样本内的预测误差平方和, SST_t 为第 t 期的总残差平方和. 类似地, R_{oos}^2 通过样本外的预测误差来构造. 样本外均方误差 MSE_{oos} 指的是样本外拟合值与真实值之间差值的平方的均值.

第二, 考察高管各项特征对于预测公司绩效的相对重要性. 利用 Boosting 回归树的相对重要性指标, 计算三个模型中各变量的对于模型拟合的相对重要性, 取 T 次拟合结果中各变量的相对重要性的均值作为该变量综合的相对重要性. 在 CEO 模型和董事长模型中, 将各人物特征的相对重要性相加可以得到 CEO 或董事长人物特征的的综合的相对重要性, 从而可以进一步对比 CEO 或董事长特征族在模型中的相对重要程度, 从中总结出对公司绩效预测能力较强的高管特征.

第三, 对于相对重要性较高的高管特征, 利用 Boosting 回归树生成部分依赖图, 进一步分析其对于公司绩效的预测模式和边际影响.

最后, 对本文主要的分析结论进行稳健性检

⑧ 随机森林 (random forest) 模型和 XGBoost 也是可解释性较强的模型, 本文在稳健性分析中加以讨论.

验,包括改变滚动时间窗口、改变公司绩效衡量方式、调整 Boosting 模型的估计参数、更换机器学习方法和更换高管特征变量。

4 实证检验与结果分析

4.1 高管特征能够预测公司业绩吗？

本部分利用 Boosting 回归树的方法和中国 A 股上市公司的样本,分别实证检验了 CEO 和董事长的个人特征是否能够较大幅度上提高模型对公司业绩的预测效果.表 4 列出了“Boosting-CEO 模型”和“Boosting-董事长模型”的拟合效果,并加入了仅包括公司层面指标的“Boosting-基准模型”作为对照,通过分析二者相对于“Boosting-基准模型”的拟合效果分别提升的程度,判定 CEO

和董事长的个人特征整体对公司业绩预测的贡献程度.同时,将“OLS-基准模型”、“OLS-CEO 模型”和“OLS-董事长模型”^⑨作为对照组,检验 Boosting 回归树的机器学习方法是否能显著提高模型的预测能力.在表 4 的面板 A 和面板 B 中,分别以财务绩效和市场绩效作为公司业绩的代理变量.

首先,表 4 中 Boosting 回归树的拟合效果相对于 OLS 模型有大幅提高.结果表明,Boosting 回归树在训练集拟合能力和样本外拟合能力两方面皆优于 OLS.比如在列(2)的对比中,Boosting 回归树在 R^2_{IS} 和 R^2_{OOS} 指标上比 OLS 分别提高了 12.23% (22.69% - 10.46%) 和 4.07% (13.43% - 9.36%).这说明选择 Boosting 回归树来改进传统线性模型的效果是十分显著的.

表 4 主要拟合结果
Table 4 Main fitting results

Panel A 财务绩效						
		基准模型	CEO 模型	提升效果	董事长模型	提升效果
	模型	(1)	(2)	(2) - (1)	(3)	(3) - (1)
$R^2_{IS}(\%)$	Boosting	19.94	22.69	2.75	21.09	1.15
	OLS	9.43	10.46	1.03	10.33	0.90
$R^2_{OOS}(\%)$	Boosting	12.35	13.43	1.08	13.26	0.91
	OLS	9.17	9.36	0.19	9.54	0.37
MSE_{OOS}	Boosting	0.012 4	0.012 3	-0.000 1	0.012 3	-0.000 1
	OLS	0.012 9	0.012 8	-0.000 1	0.012 8	-0.000 1
Panel B 市场绩效						
		基准模型	CEO 模型	提升效果	董事长模型	提升效果
	模型	(1)	(2)	(2) - (1)	(3)	(3) - (1)
$R^2_{IS}(\%)$	Boosting	65.87	67.97	2.10	67.81	1.94
	OLS	48.86	50.17	1.31	49.91	1.05
$R^2_{OOS}(\%)$	Boosting	37.72	38.62	0.90	38.78	1.06
	OLS	27.88	28.43	0.55	28.44	0.56
MSE_{OOS}	Boosting	1.78	1.75	-0.03	1.75	-0.03
	OLS	2.05	2.03	-0.02	2.03	-0.02

第二,中国 A 股上市公司 CEO 和董事长的个人特征对于预测公司绩效的作用相对于公司自身的特征指标(包括公司规模、杠杆率、国有股份占比、重资产比例和公司寿命)比较有限.综合各个指标的结果,发现最大的提升效果为 CEO 对公司财务绩效的 R^2_{IS} (2.75%), 远远低于公司层面

特征对于公司业绩的预测力度 (19.94%). 这一结果可能的原因是: 与美国等成熟国家市场相比,中国的产权不明晰以及缺乏对投资者的保护,因而中国公司面临着更为严重的代理问题.这些代理问题主要体现在管理层与股东之间、大股东与小股东之间的利益背离,进而导致了管理

⑨ 与 Boosting 模型类似, OLS 模型也是采用一年期滚动窗口回归,并分别计算了其在样本内的拟合效果和样本外的泛化能力.

层和控股股东更多地追求个人利益而忽视公司整体业绩^[58]，所以**高管特征整体而言对公司绩效的预测能力比较有限**。

第三，**CEO 和董事长在预测财务绩效和市场绩效方面的相对能力略有差异**。从面板 A 的列 (1) 和列 (3) 的结果来看，相比于“Boosting-基准模型”，“Boosting-CEO 模型”预测财务绩效的 R^2_{00s} 提升了 1.08%；而类似的“Boosting-董事长模型”预测财务绩效的 R^2_{00s} 提高了 0.91%。因此，CEO 对于财务绩效的预测能力大于董事长。而从面板 B 的 R^2_{00s} 结果中，却发现了相反的结果：CEO 对于市场绩效的预测能力弱于董事长。这一结果可能的原因在于我国公司主要用会计业绩来考核 CEO 而较少考虑公司的市场绩效^[59]，所以 CEO 会更加致力于提升公司的财务绩效；而董事长在非国有上市公司中往往直接持有公司股票并

控股，在国有企业中更是由政府部门直接任命或委派，是控股国有资本的直接代理人^[60]，因此董事长相比于 CEO 更多地代表公司股东的利益，对市场绩效更为重视。

4.2 高管个人特征的相对重要性分析

表 5 中报告了 CEO 模型和董事长模型中各变量对于预测公司绩效的相对重要性，其中面板 A 和面板 B 分别展示了各变量对于预测财务绩效和市场绩效的结果。每一个面板的左侧为 CEO 模型，右侧为董事长模型。为了便于区分，在属于人物特征的变量后标记了“*”。从而发现：**在众多的高管特征中，持股比例和年龄对预测公司绩效的相对重要性较高**（在面板 A 和面板 B 中都大于 1%），而其他诸如政府背景所代表的政府资源、金融工作背景所代表的金融资源和不同职能的工作经验等特征的相对重要性较低。

表 5 各变量相对重要性排序

Table 5 The ranking of relative variable importance

Panel A 财务绩效

排序	CEO 模型		董事长模型	
	变量	相对重要性 (%)	变量	相对重要性 (%)
1	资产负债率	31.77	资产负债率	32.25
2	公司规模	21.27	公司规模	21.30
3	重资产占比	18.52	重资产占比	18.73
4	年龄*	6.45	年龄*	6.39
5	国有控股比例	5.51	公司寿命	5.66
6	公司寿命	5.50	国有控股比例	5.59
7	持股比例*	4.91	持股比例*	4.92
8	影响力 (兼职) *	1.05	管理职能*	0.81
9	专业职能*	1.00	金融工作背景*	0.77
10	学术经验*	0.83	政府背景*	0.75
11	政府背景*	0.76	专业职能*	0.70
12	金融工作背景*	0.68	CEO 董事长兼任*	0.47
13	性别*	0.49	学术经验*	0.42
14	CEO 董事长兼任*	0.49	性别*	0.35
15	生产职能*	0.41	生产职能*	0.34
16	海外经历*	0.35	影响力 (兼职) *	0.34
17	管理职能*	0.00	海外经历*	0.22
	CEO 特征族相对重要性总计	17.43	董事长特征族相对重要性总计	16.47

表 5 (续)
Table 5 (Continue)

Panel B 市场绩效

排序	CEO 模型		董事长模型	
	变量	相对重要性 (%)	变量	相对重要性 (%)
1	公司规模	50.06	公司规模	50.19
2	资产负债率	25.57	资产负债率	25.76
3	重资产占比	8.92	重资产占比	9.13
4	持股比例*	5.10	持股比例*	4.94
5	公司寿命	3.06	公司寿命	3.08
6	年龄*	2.96	年龄*	3.07
7	国有控股比例	1.42	国有控股比例	1.47
8	学术经验*	0.76	学术经验*	0.57
9	影响力 (兼职) *	0.46	政府背景*	0.41
10	专业职能*	0.35	生产职能*	0.37
11	海外经历*	0.28	专业职能*	0.32
12	生产职能*	0.27	CEO 董事长兼任*	0.19
13	政府背景*	0.25	影响力 (兼职) *	0.18
14	CEO 董事长兼任*	0.24	性别*	0.10
15	性别*	0.18	管理职能*	0.08
16	金融工作背景*	0.13	海外经历*	0.07
17	管理职能*	0.00	金融工作背景*	0.07
	CEO 特征族相对重要性总计	10.97	董事长特征族相对重要性总计	10.37

注: 1. 依照相对重要性高低进行排序, 人物特征后标记*。

2. 下划线代表相对重要性指标在面板 A 和面板 B 中的均大于 1% 的人物特征变量。

持股比例反映了高管对公司决策的话语权和公司绩效对公司高管的激励水平。高管年龄反映了高管对于生产经营和管理公司的内部资源的积累。相比而言, CEO 和董事长的其他人物特征反映的是其对个人自身技能和外部资源的积累。例如, 生产职能、专业职能和管理职能的工作经验和学术经验可以反映个人自备技能的积累, 政府工作背景、金融工作背景和海外经历等可以反映外部资源积累。上述分析结果表明高管的持股比例和年龄能够更大程度上预测公司绩效。其原因主要是在中国 A 股上市公司中, 高管对公司内部资源的积累和把控能力对于公司绩效的影响大于其个人自身技能和对外部资源的积累把控能力带来的影响。

4.3 重要高管特征对于公司绩效的预测模式

在变量重要性的分析中, 本文发现持股比例和年龄是高管特征中两个最为重要的预测公司绩效的因素。那么这两个特征和公司绩效的关联究竟是怎样的呢? 为了回答这一问题, 本文中采用部分依赖图来考察持股比例与年龄对于公司绩效

的预测模式。图 1 至图 4 展示了 CEO 持股比例和年龄对于公司绩效的预测模式^⑩, 横轴为持股比例或年龄, 纵轴为公司绩效。

图 1 和图 2 分别展示了中国 A 股上市公司的 CEO 持股比例与公司财务绩效和市场绩效的部分依赖图。根据表 2 描述性统计的结果, 中国 A 股上市公司的 CEO 和董事长的持股比例平均值很低, 中位数皆为 0。进一步观察发现, 18 170 个观察值中, 仅有 53 个 CEO 的持股比例高于 60%, 仅有 76 个董事长的持股比例高于 60%, 高比例占股的现象极少。因此本文主要分析持股比例在 0% 到 60% 区间的部分依赖图, 高管的持股比例对公司绩效的预测效果具有明显的非线性特点。整体来看, 随着高管持股比例的提升, 公司绩效呈上升趋势。公司绩效受到 CEO 持股比例影响的程度在不同持股比例下是不同的。部分依赖图中的曲线初始阶段较为平缓, 后半段陡升。这说明当持股比例处于中等以下水平区间时 (0% 至约 40%), 持股比例增加对提升公司绩效的作用较为微弱; 而当持股

⑩ 董事长的持股比例和年龄的分析结果与 CEO 的类似。限于篇幅, 本文在正文中仅报告了 CEO 相关的部分依赖图。

比例处于较高水平时(约 40% 以上), 增加持股比例对于提升公司绩效的作用明显。

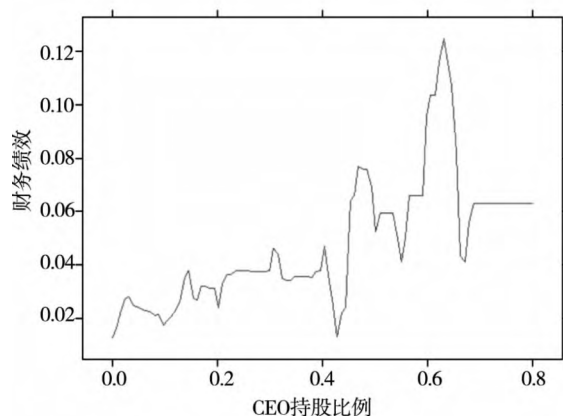


图 1 CEO 持股比例与财务绩效

Fig. 1 CEO ownership and financial performance

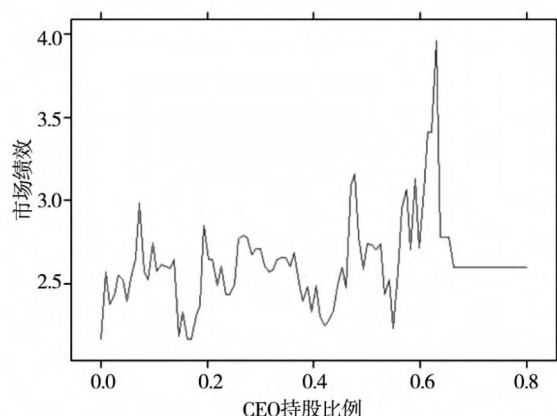


图 2 CEO 持股比例与市场绩效

Fig. 2 CEO ownership and market performance

以往文献中针对高管持股的激励效果存在争论, 主流观点有以下三种. 首先, 管理层所有权的增加可以使得管理者与股东的利益趋同, 缓解代理问题^[61]. 其次, 当高管持股水平超过一定阈值后, 他们的权力难以受到控制, 就能更多地谋取自身的利益而牺牲股东利益, 即“壕沟防御效应 (entrenchment effect)”^[62]. 最后, 高管持股与公司绩效之间也可能没有明显关联^[63]. 基于这三类效应, 目前大量的实证结果验证了高管持股与公司业绩之间的非线性关联^①.

基于上述理论, 本文认为图 1 和图 2 的结果

是因为当高管持股比例较高时, 利益趋同效应更加明显. 与此同时, 随着持股比例的提升, 其对公司绩效的作用会在局部出现“下探”现象, 即在局部区间对公司绩效有负面影响. 理论上, 这种“下探”对应着“壕沟防御效应”, 即提升高管持股比例可能使高管对公司的控制能力增强, 受到的约束变弱, 进而可以追求更大范围的个人利益, 提高代理成本. 该非线性关系也与 Kim 和 Lu^[24] 的发现一致: 当高管持股比例高于一定水平时, 经理人会过于风险厌恶, 从而放弃一些对公司有利的但高风险的投资行为 (例如研发投入), 从而导致企业价值下降. 总体而言, 本文发现高管持股对公司业绩的作用效果是非线性的, 且在不同持股比例区间可能呈现出利益趋同效应、壕沟防御效应或没有影响, 这对已有理论是互相补充印证的.

图 3 和图 4 分别展示了中国 A 股上市公司的 CEO 年龄与公司财务绩效和市场绩效的部分依赖图. 高管的年龄对于公司绩效也具有明显的非线性预测作用, 且对财务绩效和市场绩效的预测模式差异较大. 从财务绩效来看, 年龄与绩效的关系近似与“U”型结构, 特别年轻 CEO 与特别年长的 CEO 管理的公司财务绩效更佳. 而市场绩效却会随着 CEO 年龄增长而逐步提升, 同时在年龄超过一定范围后, 市场业绩会出现随年龄增长而大幅下降的趋势. 这说明年龄与公司绩效之间的关联较为复杂.

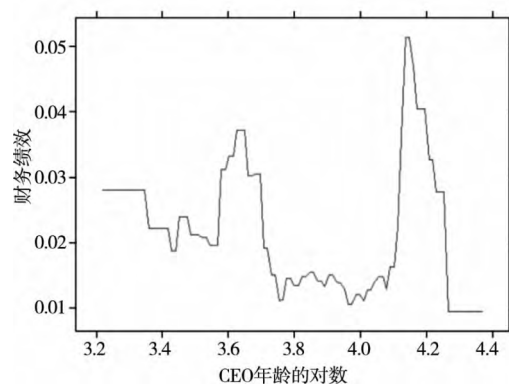


图 3 CEO 年龄与财务绩效

Fig. 3 CEO age and financial performance

① 对于财务绩效, 刘鲁彬^[64] 的实证研究发现, 当高管持股低于约 23% 或高于约 61% 时, 其与扣除了非经常性损益后的净资产回报率呈显著正相关, 否则呈显著负相关. 对于市场绩效, Morck 等^[65] 的分段回归分析表明, 高管持股比例低于 5% 或高于 25% 时, 其与托宾 Q 值呈显著正相关, 否则呈显著负相关. 类似地, 韩亮亮等^[66] 利用小规模的中国市场样本发现当高管持股比例低于 8% 或高于 25% 时, 其与托宾 Q 值呈正相关, 否则呈负相关关系.

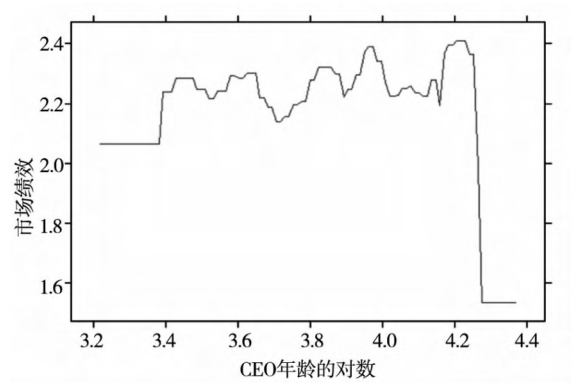


图 4 CEO 年龄与市场绩效

Fig. 4 CEO age and market performance

从理论上讲, CEO 年龄可能会有三种机制来预测公司决策和公司业绩. 第一, 年龄的变化给高管带来的个人心理和生理的变化将会间接影响公司的投融资行为. 比如年轻的 CEO 容易过度自信, 因此更容易从事并购以及高风险项目, 而年老的 CEO 管理风格更为保守^[67, 68]. 其次, CEO 年龄是其在公司中的地位与权力的体现, 年龄大的 CEO 在公司拥有更强的控制力, 会利用自身权力谋取私利, 不利于公司业绩提升. 最后, 年龄大意味着更高的社会资本和经验, 因此年龄大的 CEO 为企业贡献更多的人力资

本, 帮助企业提升业绩. 结合图 3 和图 4 的结果, 本文认为年龄与公司业绩的关系体现了上述三种预测机制, 且不同作用机制的相对强弱与其他变量存在较强的交互关系, 比如企业的国有性质会使得年龄带来的管理权力效应更为显著. 因此, 年龄与公司业绩的关联较为复杂, 会在未来研究中关注.

5 稳健性检验

5.1 更换滚动时间窗口

在基本结果中, 选取 1 年为窗口期进行滚动拟合. 为了避免窗口期长短选择对结果的影响, 在表 6 中报告了以 3 年为窗口期的结果, 结果与表 4 一致, 即中国上市公司 CEO 和董事长个人特征对于预测公司绩效的作用较为有限. 以 CEO 模型为例, 预测财务绩效和市场绩效相比于基准模型的 R^2_{IS} 的净提升分别为 2.92% 和 2.88%. 基于上述前提进行模型拟合后, 持股比例和年龄依然是高管对公司绩效预测程度较大的个人特征, 通过了稳健性检验.^⑫

表 6 更换滚动窗口期 (三年)

Table 6 Alternative rolling windows (three years)

Panel A 财务绩效					
模型	基准模型	CEO 模型		董事长模型	
	(1)	(2)	(2) - (1)	(3)	(3) - (1)
$R^2_{IS}(\%)$	23.41	26.33	2.92	26.24	2.83
$R^2_{OOS}(\%)$	13.10	14.97	1.87	15.06	1.96
MSE_{OOS}	0.012 1	0.011 9	-0.000 2	0.011 9	-0.000 2
Panel B 市场绩效					
模型	基准模型	CEO 模型		董事长模型	
	(4)	(5)	(5) - (4)	(6)	(6) - (4)
$R^2_{IS}(\%)$	55.61	58.49	2.88	58.33	2.72
$R^2_{OOS}(\%)$	29.70	31.62	1.92	31.50	1.80
MSE_{OOS}	1.90	1.85	-0.05	1.85	-0.05

表 7 更换公司绩效的衡量方式

Table 7 Alternative measures of corporate performance

Panel A 财务绩效					
模型	基准模型	CEO 模型		董事长模型	
	(1)	(2)	(2) - (1)	(3)	(3) - (1)
$R^2_{IS}(\%)$	8.61	10.16	1.55	10.00	1.39
$R^2_{OOS}(\%)$	2.73	3.02	0.29	3.11	0.38
MSE_{OOS}	0.000 04	0.000 04	0.000 0	0.000 04	0.000 0

⑫ 限于篇幅, 没有在文中报告稳健性分析中的变量相对重要性的实证结果.

表 7 (续)

Table 7 (Continue)

Panel B 市场绩效

模型	基准模型	CEO 模型		董事长模型	
	(4)	(5)	(5) - (4)	(6)	(6) - (4)
R_{IS}^2 (%)	31.15	33.11	1.96	33.70	2.55
R_{OOS}^2 (%)	20.07	21.87	1.80	22.14	2.07
MSE_{OOS}	0.001 3	0.001 3	0.000 0	0.001 3	0.000 0

5.2 更换公司绩效的衡量指标

在基本分析中,分别采用行业中值调整后的资产收益率(息税折旧摊销前利润除以期末总资产)和托宾 Q 值作为公司的财务和市场绩效评价指标.为了避免可能的变量衡量误差带来的结果偏差,更换了一组公司绩效衡量指标来分析.财务指标更换为行业中值调整后的资产收益率(净利润除以期末总资产),市场指标更换为行业中值调整后的个股年度超额回报率.

表 7 报告了更换公司绩效后的预测效果.整体来看,中国上市公司 CEO 和董事长个人特征对于预测公司绩效的作用依然十分有限.例如,对于预测财务绩效,CEO 模型和董事长模型的 R_{IS}^2 相比于基准模型的 R_{IS}^2 的净提升分别为 1.55% 和 1.39%.净利润这一指标容易受到公司管理层自发的盈余管理行为的影响,从而一定程度上偏离公司的真实经营绩效,因而表 7 中 CEO 和董事长对公司绩效的预测能力与表 4 相比略大.总体而言,“中国上市公司的 CEO 和董事长对于公司绩效的影响力较低”这一发现是稳健的.同时,持股比例和年龄依然是 CEO 和董事长共有的、对预测公司绩效具有较高相对重要性的人物特征.

5.3 Boosting 模型的参数调整

Boosting 模型的估计过程中设定了多个参数,包括回归树的交互深度、收缩参数和回归树的数量.其中,模型将回归树数目的上限设为 10 000,而模型的回归树数目是通过交叉验证最优选择的.本文拟合过程取得最佳优化效果时的回归树的数量通常处于 3 000 至 5 000 之间,因此上限设置是合理的.

本文还检验了回归树的交互深度和收缩参数的选择是否合理.在绝大多数的 GBM 模型中,回归树的交互深度都在 2 至 9 之间,而收缩系数在 0.001 和 0.1 之间^[54].基于参数检验的基本思想,构建了一系列可能的参数组合,遍历这些参数下的估计结果,从而判断在基本分析中所使用的参数是否合理.如果预测结果对参数的选择十分敏感的话,那么选择的参数则是不合理.

具体而言,本文将交互深度设为{2, 4, 6, 8},收缩参数设为{0.001, 0.01, 0.1},共构成 12(=4×3) 个参数组合,并分别在每个参数下进行估计.表 8 以财务绩效的预测模型为例,展示了参数变动对预测效果的影响.结果发现在这些参数组合下,CEO 和董事长对公司绩效预测效果的提升程度与表 4 的结果相近.

表 8 更换参数组合

Table 8 Alternative parameters

参数组合		CEO 模型			董事长模型		
交互深度	收缩参数	R_{IS}^2 (%)	R_{OOS}^2 (%)	MSE_{OOS}	R_{IS}^2 (%)	R_{OOS}^2 (%)	MSE_{OOS}
2	0.001	18.90	12.59	0.012 4	17.90	12.41	0.012 4
2	0.01	19.14	12.62	0.012 4	18.85	12.53	0.012 4
2	0.1	18.36	11.46	0.012 6	18.38	12.21	0.012 4
4	0.001	21.08	13.16	0.012 3	19.58	12.88	0.012 4
4	0.01	20.99	13.00	0.012 3	20.51	13.01	0.012 3
4	0.1	19.66	12.32	0.012 4	19.55	12.29	0.012 4
6	0.001	22.69	13.43	0.012 3	21.09	13.26	0.012 3
6	0.01	21.68	13.26	0.012 3	20.43	13.03	0.012 3
6	0.1	20.87	12.59	0.012 4	20.95	12.61	0.012 4
8	0.001	23.61	13.70	0.012 2	22.88	13.71	0.012 2
8	0.01	23.86	13.61	0.012 2	22.83	13.47	0.012 3
8	0.1	21.42	12.86	0.012 4	20.32	12.33	0.012 4

5.4 更换机器学习方法

5.4.1 随机森林 (Random Forest)

Random Forest 是机器学习算法中另一种常用的回归树模型. 文献表明, Boosting 的预测能力强于 Random Forest, 尤其是在指标维度低于 4 000 维时^[69,70]. 因为 Random Forest 在解决回归问题时, 无法做出超越训练集数据范围的预测, 在噪音较大的问题上容易出现过度拟合^[45]. 所以本文没有

将 Random Forest 作为主要的研究方法, 但采用这一方法作为稳健性检验.

表 9 报告了 Random Forest 模型拟合的基本结果. 发现, 相对于 Boosting 模型, Random Forest 模型的结果中高管特征对公司业绩的预测性有所提高, 但总体而言仍处于较低水平. 比如, 高管个人特征的加入对公司财务和市场绩效的预测误差分别降低了 0.001 0 和 0.13.

表 9 更换机器学习算法 (随机森林)

Table 9 Alternative machine learning algorithms (random forest)

Panel A 财务绩效

	基准模型	CEO 模型		董事长模型	
	(1)	(2)	(2) - (1)	(3)	(3) - (1)
$R^2_{IS}(\%)$	63.51	80.83	17.32	80.82	17.31
$R^2_{OOS}(\%)$	15.20	22.14	6.94	22.60	7.40
MSE_{OOS}	0.012 0	0.011 0	-0.001 0	0.011 0	-0.001 0

Panel B 市场绩效

	基准模型	CEO 模型		董事长模型	
	(6)	(7)	(7) - (6)	(8)	(8) - (6)
$R^2_{IS}(\%)$	82.06	91.09	9.03	91.12	9.06
$R^2_{OOS}(\%)$	37.02	41.99	4.97	42.04	5.02
MSE_{OOS}	1.79	1.66	-0.13	1.66	-0.13

5.4.2 XGBoost

XGBoost (Extreme Gradient Boosting) 是除 GBM 之外另一种优化 GBRT 问题的算法, 它与 GBRT 的主要区别在于在优化过程中不仅考虑了损失函数的一阶导数信息, 而且对损失函数进行了二阶泰勒展开, 同时用到了一阶和二阶导数信息^[71]. 但由于参数数量的增多, XGBoost 所需的

运算时间更长. 在表 10 中报告了 XGBoost 的估计结果, 发现高管模型相对于基准模型的提升程度与表 4 中 GBM 的结果相似, 多数指标的提升幅度仍处于较低水平. 比如, CEO 个人特征的加入对财务绩效的样本外预测误差 (MSE_{OOS}) 的降低幅度仅为 0.000 2. 综上, 本文的基本结果是稳健的.

表 10 更换机器学习方法 (XGBoost)

Table 10 Alternative machine learning algorithms (XGBoost)

Panel A 财务绩效

	基准模型	CEO 模型		董事长模型	
	(1)	(2)	(2) - (1)	(3)	(3) - (1)
$R^2_{IS}(\%)$	22.13	23.58	1.45	23.38	1.25
$R^2_{OOS}(\%)$	11.44	12.55	1.11	12.53	1.09
MSE_{OOS}	0.012 6	0.012 4	-0.000 2	0.012 4	-0.000 2

Panel B 市场绩效

	基准模型	CEO 模型		董事长模型	
	(6)	(7)	(7) - (6)	(8)	(8) - (6)
$R^2_{IS}(\%)$	65.75	68.43	2.68	66.73	0.98
$R^2_{OOS}(\%)$	37.11	37.88	0.77	37.89	0.78
MSE_{OOS}	1.80	1.78	-0.02	1.78	-0.02

5.5 更换高管特征变量

在主要结果的 CEO 模型及董事长模型中,

高管个人特征大部分为虚拟变量. 而回归树的方法可能会导致这些变量相对于连续型变量的

预测能力较低. 因此, 在原有一系列指标的基础上, 将高管个人特征的虚拟变量转化为高管团队特征的连续变量. 例如, 在 CEO 模型中, 以高管团队作为一个整体, 将“是否有学术研究经历”转换为拥有学术研究经历的高管占高管总人数的比重. 同理, 本文转换了其他虚拟变量, 包括是否在其他公司兼职、是否有生产、技术、设计职能经验, 是否有市场、战略、人力管理职能经验, 是否有财务、法律职能经验, 是否有海

外工作、求学经历、是否有政府工作经历和是否有金融行业工作经历. 董事长的个人特征也转化为董事会成员中拥有相应特征的成员占比. 模型的拟合结果在表 11 中汇报. 发现, 以特征指标的连续变量为预测指标的模型拟合效果有了一定提升, 比如在对财务绩效的预测中, Boosting-CEO 模型相对基准模型的样本外提升效果从 1.08 变为 1.47, 但仍然处于较低的水平, 这说明基本结论是稳健的.

表 11 高管特征均值的拟合结果

Table 11 Fitting results on the mean value of executive features

Panel A 财务绩效

	模型	基准模型 (1)	CEO 模型 (2)	提升效果 (2) - (1)	董事长模型 (3)	提升效果 (3) - (1)
$R^2_{IS}(\%)$	Boosting	19.94	22.53	2.59	25.34	5.40
	OLS	9.43	10.88	1.45	10.47	1.04
$R^2_{OOS}(\%)$	Boosting	12.35	13.82	1.47	14.45	2.10
	OLS	9.17	9.92	0.75	9.68	0.51
MSE_{OOS}	Boosting	0.012 4	0.012 2	-0.000 2	0.012 1	-0.000 3
	OLS	0.012 9	0.012 8	-0.000 1	0.012 8	-0.000 1

Panel B 市场绩效

	模型	基准模型 (4)	CEO 模型 (5)	提升效果 (5) - (4)	董事长模型 (6)	提升效果 (6) - (4)
$R^2_{IS}(\%)$	Boosting	65.87	69.53	3.66	69.42	3.55
	OLS	48.86	50.99	2.13	50.89	2.03
$R^2_{OOS}(\%)$	Boosting	37.72	38.88	1.16	38.77	1.05
	OLS	27.88	28.92	1.14	28.96	1.08
MSE_{OOS}	Boosting	1.78	1.75	-0.03	1.75	-0.03
	OLS	2.05	2.02	-0.03	2.02	-0.03

6 结束语

以往高管特征研究大多仅研究单一特征与公司业绩之间的因果关联, 缺乏全面的高管特征分析, 特别是其对公司绩效的预测能力的分析. 本文首次采用机器学习中的 Boosting 回归树算法全面考察了多维度的高管特征与公司业绩之间的关联. 具体而言, 本文以我国 2008 年~2016 年上市公司为样本, 总结了高管的高维特征变量, 研究了高管的个人特征是否能较大程度地预测公司业绩, 并进一步分析了对公司业绩预测程度较大的高管个人特征及其预测模式. 研究发现: (1) 整体而言, 公司 CEO 和董事长的特征对公司业绩的预测力较弱; (2) 在众多高管个人特征之中,

高管持股比例和年龄对公司业绩的预测程度较高; (3) 高管持股比例和年龄与公司业绩之间的关联呈现出非线性的特点, 与以往的理论吻合. 本研究不仅从一个更为全面的视角推进了中国高管特征与公司业绩关联的研究, 也对公司遴选高管和激励机制设计等提供了有益启发.

本文首次应用机器学习方法来研究中国的公司治理的问题, 评估了公司高管特征对于公司绩效的整体预测能力, 丰富了“公司高管特征与公司绩效之间关系”这一领域的文献. 此外, 本文采用前沿的 Boosting 回归的方法规避了传统线性模型的缺陷, 更加适用于分析变量之间的非线性和交互关系. 最后, 本文探究了不同公司高管特征对于预测公司绩效的重要程度, 并分析了相对重要的变量对公司绩效的作用效果, 这一系列结论

对于选聘公司高管具有实践意义。

本文开启了中国高管特征研究的一个新的维度: 高维高管特征对公司绩效的预测能力. 它对于深入理解中国高管在企业经营管理中所扮演的角色具有重要意义. 基于机器学习方法和预测性模型, 未来仍有广阔的研究空间值得探索. 首先, 机器学习方法能够利用非结构性数据来构建无法直接观测的变量 (比如高管的性格特点和行为

等), 从而挖掘在理论上尚未证实的重要高管特征, 丰富高管个人特征对公司绩效的影响以及公司治理方面的研究. 第二, 除了公司绩效这一直接衡量公司表现的变量之外, 本文用类似的模型可以进一步探究中国高管特征对公司经营决策的预测能力, 从而丰富高管对公司的影响机制方面的研究. 总体而言, 本文开拓了新的研究视角, 为后续研究奠定了基础.

参考文献:

- [1] Hambrick D C, Mason P A. Upper echelons: The organization as a reflection of its top managers [J]. *Academy of Management Review*, 1984, 9(2): 193–206.
- [2] Bertrand M, Schoar A. Managing with style: The effect of managers on firm policies [J]. *Quarterly Journal of Economics*, 2003, 118(4): 1169–1208.
- [3] Jensen D D, Cohen P R. Multiple comparisons in induction algorithms [J]. *Machine Learning*, 2000, 38(3): 309–338.
- [4] Shmueli G. To explain or to predict? [J]. *Statistical Science*, 2010, 25(3): 289–310.
- [5] Dawes R M. The robust beauty of improper linear models in decision making [J]. *American Psychologist*, 1979, 34(7): 571–582.
- [6] Forster M, Sober E. How to tell when simpler, more unified, or less ad hoc theories will provide more accurate predictions [J]. *The British Journal for the Philosophy of Science*, 1994, 45(1): 1–35.
- [7] 吴昊旻, 墨沈微, 孟庆玺. 公司战略可以解释高管与员工的薪酬差距吗? [J]. *管理科学学报*, 2018, 21(9): 105–117.
Wu Haomin, Mo Shenwei, Meng Qingxi. Is the corporate strategy responsible for the CEO-employee pay gap? [J]. *Journal of Management Sciences in China*, 2018, 21(9): 105–117. (in Chinese)
- [8] 王 曾, 符国群, 黄丹阳, 等. 国有企业 CEO 政治晋升与在职消费关系研究 [J]. *管理世界*, 2014, 5: 157–171.
Wang Zeng, Fu Guoqun, Huang Danyang, et al. The relation between the political promotion for CEOs of state-owned firms and perquisite consumption [J]. *Management World*, 2014, 5: 157–171. (in Chinese)
- [9] Adams R, Almeida H, Ferreira D. Powerful CEOs and their impact on corporate performance [J]. *Review of Financial Studies*, 2005, 18, 1403–432.
- [10] Li M, Lu Y, Phillips G M. CEOs and the product market: When are powerful CEOs beneficial? [J]. *Journal of Financial and Quantitative Analysis*, 2019, 54(6): 2295–2326.
- [11] Kaplan S N, Sorensen M. Are CEOs Different? Characteristics of Top Managers [R]. Chicago: NBER Working Paper no. 23832, 2017.
- [12] Barker V L, Mueller G C. CEO characteristics and firm R&D spending [J]. *Management Science*, 2002, 48(6): 782–801.
- [13] Malmendier U, Tate G, Yan J. Overconfidence and early-life experiences: The effect of managerial traits on corporate financial policies [J]. *The Journal of Finance*, 2011, 66(5): 1687–1733.
- [14] Bernile G, Bhagwat V, Rau P R. What doesn't kill you will only make you more risk-loving: Early-life disasters and CEO behavior [J]. *The Journal of Finance*, 2017, 72(1): 167–206.
- [15] Roussanov N, Savor P. Marriage and managers' attitudes to risk [J]. *Management Science*, 2014, 60(10): 2496–2508.
- [16] Cronqvist H, Yu F. Shaped by their daughters: Executives, female socialization, and corporate social responsibility [J]. *Journal of Financial Economics*, 2017, 126(3): 543–562.
- [17] Schoar A, Zuo L. Shaped by booms and busts: How the economy impacts CEO careers and management styles [J]. *Review of Financial Studies*, 2017, 30(5): 1425–1456.
- [18] Malmendier U, Tate G. CEO overconfidence and corporate investment [J]. *Journal of Finance*, 2005, 60(6): 2661–2700.
- [19] Hirshleifer D, Low A, Teoh S H. Are overconfident CEOs better innovators [J]. *The Journal of Finance*, 2012, 67(4):

- 1457 – 1498.
- [20] Malmendier U, Tate G. Behavioral CEOs: The role of managerial overconfidence [J]. *Journal of Economic Perspectives*, 2015, 29(4): 37 – 60.
- [21] Sunder J, Sunder S V, Zhang J. Pilot CEOs and corporate innovation [J]. *Journal of Financial Economics*, 2017, 123(1): 209 – 224.
- [22] Chava S, Purnanandam A. CEOs versus CFOs: Incentives and corporate policies [J]. *Journal of Financial Economics*, 2010, 97(2): 263 – 278.
- [23] Cain M D, McKeon S B. CEO personal risk-taking and corporate policies [J]. *Journal of Financial and Quantitative Analysis*, 2016, 51(1): 139 – 164.
- [24] Kim E H, Lu Y. CEO ownership, external governance, and risk-taking [J]. *Journal of Financial Economics*, 2011, 102(2): 272 – 292.
- [25] Kaplan S N, Klebanov M M, Sorensen M. Which CEO characteristics and abilities matter [J]. *The Journal of Finance*, 2012, 67(3): 973 – 1007.
- [26] Falato A, Li D, Milbourn T. Which skills matter in the market for CEOs? Evidence from pay for CEO credentials [J]. *Management Science*, 2015, 61(12): 2845 – 2869.
- [27] Bandiera O, Prat A, Hansen S, et al. CEO behavior and firm performance [J]. *Journal of Political Economy*, 2020, 128(4): 1325 – 1369.
- [28] Coles J, Daniel N D, Naveen L. Managerial incentives and risk-taking [J]. *Journal of Financial Economics*, 2006, 79(2): 431 – 468.
- [29] Fahlenbrach R, Stulz R M. Bank CEO incentives and the credit crisis [J]. *Journal of Financial Economics*, 2011, 99(1): 11 – 26.
- [30] Fahlenbrach R. Founder-CEOs, investment decisions, and stock market performance [J]. *Journal of Financial and Quantitative Analysis*, 2009, 44(02): 439 – 466.
- [31] Mehrotra V, Morck R, Shim J, et al. Adoptive expectations: Rising sons in Japanese family firms [J]. *Journal of Financial Economics*, 2013, 108(3): 840 – 854.
- [32] Khanna V, Kim E H, Lu Y. CEO connectedness and corporate fraud [J]. *The Journal of Finance*, 2015, 70(3): 1203 – 1252.
- [33] 逯 东, 林 高, 黄 莉, 等. “官员型”高管、公司业绩和非生产性支出——基于国有上市公司的经验证据 [J]. *金融研究*, 2012, 6: 139 – 153.
Lu Dong, Lin Gao, Huang Li, et al. Administrative experienced executives, the performance of the company, and nonproductive expenditures: Based on the empirical evidence of state-owned listed firms [J]. *Journal of Financial Research*, 2012, 6: 139 – 153. (in Chinese)
- [34] 吴文锋, 吴冲锋, 刘晓薇. 中国民营上市公司高管的政府背景与公司价值 [J]. *经济研究*, 2008, 7: 130 – 141.
Wu Wenfeng, Wu Chongfeng, Liu Xiaowei. Political connection and market valuation: Evidence from China individual-controlled listed firms [J]. *Economic Research Journal*, 2008, 7: 130 – 141. (in Chinese)
- [35] 周楷唐, 麻志明, 吴联生. 高管学术经历与公司债务融资成本 [J]. *经济研究*, 2017, 7: 169 – 183.
Zhou Kaitang, Ma Zhiming, Wu Liansheng. Managerial academic experience and cost of debt [J]. *Economic Research Journal*, 2017, 7: 169 – 183. (in Chinese)
- [36] 林晚发, 钟辉勇, 李青原. 高管任职经历的得与失? ——来自债券市场的经验证据 [J]. *金融研究*, 2018, 456(6): 171 – 188.
Lin Wanfa, Zhong Huiyong, Li Qingyuan. The gains and losses of senior management experience? Evidence from the bond market [J]. *Journal of Financial Research*, 2018, 456(6): 171 – 188. (in Chinese)
- [37] 程新生, 赵 旸. 权威董事专业性、高管激励与创新活跃度研究 [J]. *管理科学学报*, 2019, 22(3): 40 – 52.
Cheng Xinsheng, Zhao Yang. Authoritative professional directors, executive incentives and innovation activity [J]. *Journal of Management Sciences in China*, 2019, 22(3): 40 – 52. (in Chinese)
- [38] 徐莉萍, 赖丹丹, 辛 宇. 不可承受之重: 公司高管婚变的经济后果研究 [J]. *管理世界*, 2015, (5): 117 – 133.
Xu Liping, Lai Dandan, Xin Yu. Unbearable weight: A study of the economic consequences of corporate executive divorces [J]. *Management World*, 2015, (5): 117 – 133. (in Chinese)
- [39] 陈 闯, 张 岩, 吴晓晖. 风险投资、创始人与高管薪酬——多边代理视角 [J]. *管理科学学报*, 2017, 20(6): 78 – 88.

- Chen Chuang, Zhang Yan, Wu Xiaohui. Venture capital, founder and top management team compensation: A view from multiple agency theory [J]. Journal of Management Sciences in China, 2017, 20(6): 78–88. (in Chinese)
- [40] 陆瑶, 李茶. CEO对董事会的影响力与上市公司违规犯罪[J]. 金融研究, 2016, (1): 176–191.
Lu Yao, Li Cha. CEO influence over the board and corporate fraud [J]. Financial Research, 2016, (1): 176–191. (in Chinese)
- [41] 陆瑶, 胡江燕. CEO与董事间的老乡关系对我国上市公司风险水平的影响[J]. 管理世界, 2014, 3: 131–138.
Lu Yao, Hu Jiangyan. The impact of hometown connectedness between the CEO and board of directors on corporate risk level [J]. Management World, 2014, 3: 131–138. (in Chinese)
- [42] 陆瑶, 胡江燕. CEO与董事间老乡关系对公司违规行为的影响研究[J]. 南开管理评论, 2016, 2: 52–62.
Lu Yao, Hu Jiangyan. The impact of hometown connectedness between the CEO and board of directors on corporate fraud [J]. Nankai Business Review, 2016, 2: 52–62. (in Chinese)
- [43] 龚辉锋, 茅宁. 咨询董事、监督董事与董事会治理有效性[J]. 管理科学学报, 2014, 17(2): 81–94.
Gong Huifeng, Mao Ning. Advisory directors, monitoring directors and board governance effectiveness [J]. Journal of Management Sciences in China, 2014, 17(2): 81–94. (in Chinese)
- [44] Li K, Mai F, Shen R, et al. Measuring Corporate Culture Using Machine Learning [R]. Available at SSRN 3256608, 2018.
- [45] Mullainathan S, Spiess J. Machine learning: An applied econometric approach [J]. Journal of Economic Perspectives, 2017, 31(2): 87–106.
- [46] Ban G, Karoui N E, Lim A E B. Machine learning and portfolio optimization [J]. Management Science, 2018, 64(3): 1136–1154.
- [47] Gu S, Kelly B, Xiu D. Empirical Asset Pricing Via Machine Learning [R]. Chicago: NBER Working Paper no. 25398, 2019.
- [48] Feng G, Giglio S, Xiu D. Taming the Factor zoo: A Test of New Factors [R]. Chicago: NBER Working Paper no. 25481, 2019.
- [49] Gow I D, Kaplan S N, Larcker D F, et al. CEO Personality and Firm Policies [R]. Chicago: NBER Working Paper no. 22435, 2016.
- [50] Ridgeway G. Generalized boosted models: A guide to the GBM package [J]. Update, 2019, 1(1): 1–39.
- [51] Freund Y, Schapire R E. A decision-theoretic generalization of on-line learning and an application to Boosting [J]. Journal of Computer and System Sciences, 1997, 55(1): 119–139.
- [52] Friedman J H. Greedy function approximation: A gradient Boosting machine [J]. Annals of Statistics, 2001, 29(5): 1189–1232.
- [53] Schapire R E, Freund Y. Boosting: Foundations and Algorithms [R]. MA: MIT Press, 2012.
- [54] Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning [R]. NY: Springer New York, 2009.
- [55] Doshi-Velez F, Kim B. Towards a rigorous science of interpretable machine learning [EB/OL]. arXiv, Preprint at <https://arxiv.org/abs/1702.08608>, 2017.
- [56] Molnar C. Interpretable Machine Learning: A Guide for Making Black Box Models Explainable [M]. Victoria: Leanpub Publishing, 2019.
- [57] James G, Witten D, Hastie T, et al. An Introduction to Statistical Learning with Applications in R [R]. NY: Springer Texts in Statistics (Spring-Verlag, New York), 2013.
- [58] Kato T, Long C. Executive turnover and firm performance in China [J]. American Economic Review, 2006, 96: 363–367.
- [59] 唐松, 孙铮. 政治关联、高管薪酬与企业未来经营绩效[J]. 管理世界, 2014, 5: 93–105.
Tang Song, Sun Zheng. The political connections, the CEO's salary, and firm's future management performances [J]. Management World, 2014, 5: 93–105. (in Chinese)
- [60] 黄继承, 盛明泉. 高管背景特征具有信息含量吗? [J]. 管理世界, 2013, 9: 144–153.
Huang Jicheng, Sheng Mingquan. Do the background characteristics of executives have information contents? [J]. Management World, 2013, 9: 144–153. (in Chinese)
- [61] Jensen M C, Murphy K J. Performance pay and top management incentive [J]. Journal of Political Economy, 1990, 98: 225–264.
- [62] Demsetz H. The structure of ownership and the theory of the firm [J]. Journal of Law and Economics, 1983, 26(2): 375

– 390.

- [63] 于东智. 股权结构、治理效率与公司绩效[J]. 中国工业经济, 2001, 5: 54 – 62.
Yu Dongzhi. Ownership structure, governance efficiency, and corporate performance [J]. China Industrial Economy, 2001, 5: 54 – 62. (in Chinese)
- [64] 刘鲁彬. 我国上市公司管理层持股比例与公司绩效关系实证研究[R]. 南昌: 南昌大学, 2012.
Liu Lubin. The Empirical Research on the Relationship between Managerial Ownership Structure and Firm Performance of Chinese Listed Companies [R]. Nanchang: Nanchang University, 2012. (in Chinese)
- [65] Morck R, Shleifer A, Vishny R W. Management ownership and market valuation: An empirical analysis [J]. Journal of Financial Economics, 2010, 20(88): 293 – 315.
- [66] 韩亮亮, 李 凯, 宋 力. 高管持股与企业价值——基于利益趋同效应与壕沟防守效应的经验研究[J]. 南开管理评论, 2006, 9: 35 – 41.
Han Liangliang, Li Kai, Song Li. Managerial ownership and firm value: Empirical evidence based on the convergence of interest effects and the entrenchment effects [J]. Nankai Business Review, 2006, 9: 35 – 41. (in Chinese)
- [67] Bertrand M, Mullainathan S. Enjoying the quiet life? Corporate governance and managerial preferences [J]. Journal of Political Economy, 2003, 111(5): 1043 – 1075.
- [68] Yim S. The acquisitiveness of youth: CEO age and acquisition behavior [J]. Journal of Financial Economics, 2013, 108 (1): 250 – 273.
- [69] Caruana R, Niculescu-Mizil A. An Empirical Comparison of Supervised Learning Algorithms [C] // Proceedings of the 23rd International Conference on Machine Learning, ACM, 2006: 161 – 168.
- [70] Caruana R, Karampatziakis N, Yessenalina A. An Empirical Evaluation of Supervised Learning in High Dimensions [C] // Proceedings of the 25th International Conference on Machine Learning, ACM, 2008: 96 – 103.
- [71] Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System [C] // Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining, ACM, 2016: 785 – 794.

Managerial individual characteristics and corporate performance: Evidence from a machine learning approach

LU Yao, ZHANG Ye-qing, LI Bo, ZHAO Hao-yu

School of Economics and Management, Tsinghua University, Beijing 100084, China

Abstract: In the existing literature on corporate governance, most of the research related to managerial characteristics has two main limitations. First, most of the papers focus on the relationship between one managerial individual characteristic and corporate performance but lack a comprehensive understanding of the potential non-linear relationship and interactions among some of the important independent variables. Second, existing research tests casual inference but ignores the predictive performance of the model. In this paper, we first examine if managerial individual characteristics can predict corporate performance by using a machine learning approach: Boosting regression trees. Using a sample of listed firms in the Chinese A-share market from 2008 to 2016, we study whether these individual characteristics could predict corporate performance. The evidence shows that: 1) The individual characteristics of Chinese executives including CEOs and chairmen could predict corporate performance only to a limited degree. 2) Among multiple individual characteristics, managerial ownership and executive age are the two most important predictors of corporate performance. 3) The relations between predictors and corporate performance are non-linear, consistent with the prior literature. This paper initiates a new, more thorough perspective in Chinese executive research using machine learning methods and has important implications for selecting executives and designing incentive mechanisms.

Key words: machine learning; Boosting regression trees; corporate governance; corporate performance