

# 董事会特征如何影响公司绩效？： 来自机器学习的证据

陆瑶 彭章 余思凝  
(清华大学经济管理学院)

## How Do Board Characteristics Affect Firm Performance? : Evidence from Machine Learning

Yao Lu Zhang Peng Sining Yu  
(School of Economics and Management, Tsinghua University)

### 第一作者（通讯作者）：

作者姓名：陆瑶，工作单位：清华大学经管学院金融系副教授。研究方向：公司金融，公司治理，企业投资融资，企业并购与重组，人力资本与金融 (Labor and Finance)，资本市场改革与发展。电子邮箱：[luyao@sem.tsinghua.edu.cn](mailto:luyao@sem.tsinghua.edu.cn)，联系电话：8610-62797399/13691288499，通讯地址：北京市海淀区清华大学经管学院伟伦楼 347。

### 第二作者：

作者姓名：彭章，工作单位：清华大学经管学院金融系博士研究生。研究方向：公司金融、公司治理、劳动与金融、机器学习在公司金融的应用、金融科技。电子邮箱：[pengzh.15@sem.tsinghua.edu.cn](mailto:pengzh.15@sem.tsinghua.edu.cn)，联系电话：18811328006，通讯地址：北京市海淀区清华大学紫荆 17 号楼 824A 室。

### 第三作者：

作者姓名：余思凝，工作单位：清华大学经管学院金融系硕士研究生，邮编：100084，电子邮箱：[ysn19@mails.tsinghua.edu.cn](mailto:ysn19@mails.tsinghua.edu.cn)，联系电话：13521015801，通讯地址：北京市海淀区清华大学紫荆公寓 608B。

### \*所受基金资助：

陆瑶感谢国家自然科学基金优秀青年科学基金资助项目（71722001）、清华大学自主科研计划课题项目（2015THZWWY09）的支持

# 董事会特征如何影响公司绩效？：

## 来自机器学习的证据

**摘要** 以往董事会文献大多关注单一董事会特征如何解释公司绩效，缺乏同时考虑多维董事会特征对公司绩效的预测性研究。本文运用机器学习方法探究 46 个董事会特征对公司绩效的预测能力，发现：1. 董事会特征可以预测公司绩效；2. 董事平均薪酬、董事会持股比例、董事平均任期、学历异质性的预测能力较强，董事会结构特征预测能力较弱；3. 公司绩效与董事平均薪酬基本呈正相关，而与董事会持股比例、董事平均任期、学历异质性呈非线性关系。该结果说明，我国公司中董事会特征与公司绩效的关系紧密且复杂，运用机器学习方法可以帮助我们认识这些复杂关系、提升对公司绩效预测的准确性。

**关键词** 董事会特征 公司绩效 机器学习

## How Do Board Characteristics Affect Firm Performance? : Evidence from Machine Learning

**Abstract** Previous literature about the board of directors focuses on how a specific board characteristic explains firm performance, little research studies how multiple board characteristics predict firm performance. We investigate the relation between 46 board characteristics and firm performance using machine learning approaches, we find: 1. board characteristics can predict firm performance; 2. Directors' average compensation, board ownership, directors' average tenure, and education diversity have strong predict power, but board structure characteristics have weak predict power; 3. Directors' average compensation and firm performance are positively related, the relations between directors' ownership, average tenure, education diversity, and firm performance are non-linear. Our findings suggest that the relations between board characteristics and firm performance are close and complicated. Machine learning approaches can help us identify those complicated relations and improve the accuracy of predicting firm performance.

**Key Words** Board Characteristics; Firm Performance; Machine Learning.

**JEL Classification** C53; C81; G34

## 一、引言

随着 2020 年 3 月 1 日新《证券法》的实施，我国证券市场迈入了崭新的阶段。新《证券法》显著增大了公司违法违规的处罚力度、提升了投资者保护水平、强化了信息披露要求。更加市场化、规范化、透明化的环境对我国公司的治理水平提出了更高的要求。董事会是公司治理的核心，是公司权力的最高行使者（邓峰，2011）。如何建立一个高效的、最大程度保护投资者利益的董事会不仅仅是一个重要的学术问题，更是关乎公司和投资者切身利益的实际问题，还是监管部门需要探索的政策问题。

虽然国内外学者从不同角度对董事会做了大量研究，但董事会特征与公司绩效的关系仍然存在很大争议（Adams, 2017）。而且，董事会作为一个整体，有多维度的特征，但以往文献往往都只针对单个特征的影响，缺乏对董事会特征总体上的分析。那么，董事会特征能否影响公司绩效？具体是哪些特征更为重要？这些重要特征对公司绩效的作用模式是怎样的？我们试图运用机器学习的方法，通过考察我国上市公司 46 个董事会特征对公司绩效的预测能力，在一定程度上回答这些问题。

机器学习，又称统计学习，是人工智能的代表性技术，目前被各个领域广泛运用。机器学习的目的在于自动地寻找数据中复杂的结构和模式，从而帮助预测（Hastie et al., 2009; 周志华, 2016; 李斌等, 2019）。机器学习方法与传统公司金融运用的计量方法在研究思路上存在很大差异：以往的公司金融文献大多运用解释性模型（explanatory model），侧重于解释变量间的因果关系（即因果推断），从理论假设出发，运用统计方法（如线性回归、结构方程模型等）检验假设；而机器学习的模型是预测性模型（predictive model），它侧重于模型的预测能力<sup>1</sup>，不对变量间的关系进行事先的假设，因此它能够发现解释性模型不曾假设的、更加复杂的关系，从而帮助我们发现新的关系、补充和发展现有理论。具体来说，相较于解释性模型中最常用的线性最小二乘回归（OLS）<sup>2</sup>，目前的机器学习方法有以下两点优势：1. 一些机器学习算法（如决策树、深度神经网络等）能够识别非线性的、交互的关系；2. 机器学习的参数正则化、模型筛选等技术可以帮助我们在众多的、相关度高的变量中筛选出重要变量。

研究董事会特征与公司绩效时，机器学习方法的优点可以得到充分利用。首先，董事会特征与公司绩效的关系可能是非线性的、交互的，如董事会规模与公司绩效可能呈倒 U 型关系（于东智，2003），且两者之间的关系可能受到公司规模等其他因素的影响，传统的解释性模型难以准确捕捉非线性的、交互的关系，而机器学习方法则可以自动识别这些关系；其次，董事会特征众多且存在高度相关性，由于多重共线性问题，传统的 OLS 回归无法判断高度相关的特征中哪一个才是重要特征，而机器学习算法中可以帮助我们选出重要特征。因此，本文使用 LASSO, 随机森林和渐进梯度提升树三种机器学习方法探究了董事会特征与公司绩效的关系。

本文收集了 2008—2017 年我国 A 股上市公司的 8 个公司基本特征、46 个董事会特征作为自变量，收集了公司财务绩效指标总资产回报率（ROA）和市场绩效指标托宾 Q 值作为因变量，首先运用 OLS<sup>3</sup>, LASSO, 随机森林和渐进梯度回归树（简称 Boosting 回归树）方法，构建只包含 8 个基本公司特征的模型，我们称之为“基准模型”。其次，同样运用 OLS, LASSO, 随机森林和渐进梯度回归树方法，构建包含 8 个公司基本特征和 46 个董事会特征的模型，我们称之为“董事会模型”。我们先比较基准模型和董事会模型的预测能力，预测能力以样本外  $R^2$  来衡量，样本外  $R^2$  越

---

<sup>1</sup> 预测能力能够反映理论解释现象的能力（Shmueli, 2010），如果一个从理论出发的模型有很差的预测能力，那说明这个模型所依托的理论并不能够很好地解释现象，这个理论仍然需要补充和修正。因此，考察模型的预测能力也非常重要，解释性模型与预测性模型是相辅相成的。

<sup>2</sup> 严格意义上来说，传统最小二乘回归（OLS）也是机器学习方法的一种，但目前的机器学习算法的表现远远超出 OLS 回归，因此在本文中，如果没有特殊说明，我们提到“机器学习”时，是不包括传统的 OLS 算法的。

<sup>3</sup> 我们运用最小二乘方法（OLS）作为比较基准。

高说明模型预测能力越好。我们发现董事会模型的样本外  $R^2$  均显著高于基准模型，说明董事会特征变量可以有效预测公司绩效，因此董事会特征对公司绩效有重要影响。我们比较 OLS 和三种机器学习方法，发现三种机器学习方法的预测能力均好于 OLS，其中 LASSO 方法的预测能力提升并不高，但是随机森林和渐进梯度回归树的提升非常大。接下来，对于董事会模型，我们考察了随机森林和渐进梯度提升树方法得出的变量重要性排序，发现董事平均薪酬、董事会持股比例、董事平均任期指标和董事学历异质性在每个排序中都名列前茅，说明这些指标的预测能力更强、是更为重要的董事会特征。而董事会规模、独立董事比例等董事会结构指标的排名较后，预测能力较弱。最后，对于董事会模型，我们考察了随机森林和渐进梯度提升树方法下，董事平均薪酬、董事会持股比例、董事平均任期和董事学历异质性的部分依赖图，部分依赖图可以反映这些变量对公司绩效的作用模式。我们发现：董事平均薪酬越高，公司绩效越好；董事会持股比例与公司绩效呈非线性关系，董事会持股比例在 0 至一个非常小的值（约为 2% 至 5%）之间，董事会持股比例提升能够带来非常大的公司绩效提升。但是超过这个阈值后，董事会持股比例对公司绩效的影响不大；董事平均任期与公司绩效呈非线性关系，公司董事平均任期在约 2 至 4 年的区间内，平均任期越长，公司绩效越差，但是过了 4 年之后，平均任期越长，公司绩效越高；董事学历异质性则与公司绩效呈 U 型关系。

在学术上，本文有以下两方面贡献：首先，本文同时研究了 46 个董事会特征，并运用机器学习方法比较了各个特征对公司绩效的预测能力、探究了一些重要特征与公司绩效的关系。本文在一定程度上弥补了之前董事会文献的不足：以往的董事会文献大多只研究单个董事会特征变量，没有同时考虑和比较多个特征的作用，也没有考虑这些特征可能存在的非线性的、交互的作用，而机器学习方法则可以解决这些问题。第二，我们使用机器学习的方法研究了董事会特征对公司绩效的预测能力，丰富了机器学习方法在公司金融方面运用的文献，目前机器学习方法在金融领域更多地被用在构造变量、预测股票收益等，而运用机器学习方法研究公司绩效影响因素的文献较少，本文则在一定程度上补充了这部分空白。

本文的发现也有以下现实意义：从公司股东角度，本文的发现可以作为选举董事的参考，本文对 46 个董事会特征进行了重要性排序，并分析了一些重要特征与公司绩效的关系，因此，股东可以根据这些结果，有针对性地选择董事、组成高效的董事会；从外部投资者和分析师的角度，本文的发现有助于投资者和分析师进行公司绩效的预测，本文发现随机森林和渐进梯度提升树这类方法的预测能力远远好于线性模型，且加入董事会特征能够更好地预测公司绩效，因此，投资者和分析师在进行公司未来业绩预测时，可以考虑采取加入董事会特征的机器学习方法；从政策制定者角度，本文的发现也有助于相关部门制定董事会相关政策。董事会是公司治理的关键，但什么样的董事会是有利于股东的仍众说纷纭，而我们发现很多董事会特征与公司特征都存在非线性关系，这些发现可以作为政策制定的依据。

接下来本文的安排如下：第二部分简单介绍了董事会特征相关文献和目前机器学习方法在金融中的应用；第三部分介绍本文使用的三种机器学习方法和研究设计；第四部分介绍样本、变量与数据来源；第五部分展示结果；第六部分进行稳健性检验；最后进行总结。

## 二、相关文献

### （一）董事会特征文献

作为公司最主要的决策机构和股东利益的代表，董事会在公司日常经营、管理和战略决策中起到重大作用，关于董事会特征对公司影响的研究也一直是学术界关注的焦点。董事会作为一个团队，其特征可以分为四类：董事会结构、董事会行为、董事会激励特征和董事个人特征。

董事会结构，如董事会规模、独立性、董事长是否兼任总经理等指标，一直是公司治理文献中非常重要的研究对象，然而董事会结构特征如何影响公司治理和公司绩效都存在很大争议。由于大规模的董事会存在高的沟通和协调成本，导致其决策效率更低，因此董事会规模一般被认为与公司绩效、公司价值呈负相关（如 Jensen, 1993; Yermack, 1996; Hermalin and Weisbach, 2003）。但董事会规模也不是越小越好，越大、越多元化的公司需要的董事会规模越大，而且较大的董事会也可以防止公司做出极端决策、减少业绩的方差（Cheng, 2008），一些实证研究也发现了董事会规模与公司绩效之间存在正相关（如 Dalton et al., 1999）。于东智（2003）、于东智和池国华（2003）运用我国上市公司数据，发现董事会规模与公司绩效之间存在倒 U 型关系。董事会独立性与公司绩效之间的关系更加复杂，从理论上来说，一方面，董事会独立性越高，其监督作用越强，公司治理、公司绩效越好，另一方面，独立董事对公司不够了解，引入独立董事也可能增加董事会沟通决策的成本。大量实证研究则难以给出统一的结论（Adams, 2017）<sup>4</sup>。董事长是否兼任总经理一般被看作是管理层防御（entrenchment）的表现，但实证研究并没有发现该现象（Jayaraman et al., 2015）。关于我国公司的研究也存在争议，于东海（2003）发现董事长与总经理兼任并不是影响公司绩效的重要因素；雷海民等（2012）发现董事长与总经理两职分离有利于提高公司运营效率；李晓（2015）则发现董事长与总经理兼任与绩效呈负向关系。

董事会行为也是公司治理文献中的经典研究问题之一。会议是董事会最主要的行为特征，是衡量董事会监督职能与有效性的重要指标（Lipton and Lorsch, 1992）。Vafeas（1999）发现董事会会议次数与公司价值呈负相关，但是会议之后企业业绩有所改善。于东智（2001）发现我国上市公司董事会会议次数与公司价值指标正相关，非正常会议次数与以前年度绩效负相关、与当年和以后年度正相关。牛建波和李胜楠（2007）也发现董事会会议次数增加有利于公司价值提升，这些结果都在一定程度上说明董事会会议可以起到提升公司绩效、增加公司价值的作用。

董事会激励特征主要包括董事薪酬和董事持股两个方面。董事薪酬方面，Brick et al.（2006）发现超额的董事薪酬与公司表现负相关，Adams and Ferreira（2008a）则发现给予董事会议报酬显著提高了董事会议出席率。杨青等（2009）发现董事薪酬对公司业绩没有显著影响，郑志刚等（2017）则发现高的独立董事薪酬能够显著改善公司绩效，朱滔（2015）则将董事薪酬分为董事监督薪酬与超额薪酬，发现董事监督薪酬与公司未来业绩显著正相关，而超额薪酬则与未来业绩显著负相关。董事持股方面，Maug（1997）发现给予董事股票、期权激励可以激励董事最大化股东利益，Chen et al.（2008）研究美国共同基金发现，董事的持股比例越高，其监督动力更强。Bhagat and Bolton（2013）也发现董事会持股与公司绩效正相关。

董事个人特征是近年来董事会相关研究的热点问题之一。董事的性别（Adams and Ferreira, 2008b; Liu et al., 2014; Kim and Starks, 2016）、年龄（Jiang et al., 2016）、任期（如 Huang and Hilary, 2018）、背景（如王跃堂等，2006；魏刚等，2007）、技能（Adams et al., 2018）等可能影响到董事在公司治理中的作用，然而在众多的特征中，究竟是哪些特征起作用仍无定论。此外，作为一个团队，董事会成员特征的多样性和异质性也被广泛的研究，但是也存在很大争议。一方面，多样性高的董事会会在能力上更加互补，导致更好的绩效（如 Erhardt et al., 2003），但另一方面，多样性高的董事会可能更难达成一致意见，导致公司绩效下降（Arrow, 1951; Shehata et al., 2017）。

虽然目前已经有很多关于董事会特征的研究，但是这些研究存在四个问题：第一，目前研究的结论并不一致；第二，大多研究并没有考虑董事会特征非线性的、交互的作用；第三，目前的大多数研究都只针对单个的董事会特征，很少考虑多个特征的共同作用和交互作用；第四，目前的研究都基于解释性模型，很少考虑董事会特征对公司绩效的预测能力。因此，更深入地考察董

---

<sup>4</sup> Dahya et al.（2008）、Black and Kim（2012）等研究发现董事会独立性与公司绩效呈正相关；而 Yermack（1996）、Kumar and Sivaramakrishnan（2008）等研究发现董事会独立性越高，公司绩效越差；Bhagat and Black（2002）、Hermalin and Weisbach（2003）等研究则发现董事会独立性对公司绩效没有影响。Liu et al.（2015）、王跃堂等（2006）运用我国的数据发现董事会独立性与绩效呈正相关。

事会特征与公司绩效的关系是非常必要的。本文从上述 4 个方面构造了 46 个董事会特征变量，运用机器学习模型考察了这些变量对公司绩效的预测能力。一方面，本文同时考察并比较众多的、不同方面的董事会特征的作用；另一方面，机器学习模型也可以让我们考虑变量之间的非线性的、交互的关系，这在一定程度上弥补了现有文献的不足。

## （二）机器学习方法在金融研究中的应用

一些学者运用机器学习方法来量化文字、图片等信息。上市公司的定期报表等信息披露、分析师的报告等等都提供了大量的文字的、非会计信息，机器学习中的文本分析方法，如词汇分类词典法、朴素贝叶斯分类器、LDA (Latent Dirichlet Allocation) 模型等，为我们利用这些信息提供了良好的工具 (沈艳等, 2019)。这些方法被大量用于研究公司报表 (如 Li, 2010; Jegadeesh and Wu, 2013)、分析师报告 (Huang et al., 2014) 和新闻媒体报告 (如 Tetlock, 2007; Tetlock et al., 2008; Heston and Sinha, 2015)。例如，Li (2010) 运用朴素贝叶斯算法分析美国上市公司年报和季报文件中的管理层讨论和分析部分，发现该部分语气积极的公司未来业绩更好；Bandiera et al. (2017) 运用 LDA 模型分析了 CEO 的日常活动调查记录将 CEO 分为“领导者”和“管理者”，并发现雇佣“领导者”CEO 的公司生产率更高。近年来，我国学者也针对中国上市公司做了大量相关研究。谢德仁和林乐 (2015)、林乐和谢德仁 (2017) 运用分词技术和词频统计分析了上市公司管理层说明会的文本，发现管理层语调可以预测公司业绩，且管理层正面语调对分析师评级的水平及变动都有正面影响、而负面语调则没有影响。王克敏等 (2018) 也运用分词和文本分析技术分析了我国上市公司年报，发现管理者会操纵文本复杂性来获取超额报酬、提高公司估值。马黎琚等 (2019) 运用支持向量机模型 (SVM) 对 2009-2015 的分析师报告进行了文本分类和情感分析，发现分析师报告中的前瞻性语句的情感积极程度与报告发布后的投资者反映显著正相关。除了文字之外，机器学习中的深度学习方法还可以处理图片、语音等其他信息，沈艺峰等 (2017) 运用该方法研究了高管的长相和语音对 IPO 市场的影响。

一些学者则看到机器学习在变量筛选、预测方面的强大能力，并将运用在金融领域。在资产定价方面，Gu et al. (2019) 运用随机森林、神经网络等方法，考察了 94 个股票层面特征、他们与 8 个时间序列变量的交互项、74 个行业虚拟变量 (一共超过 900 个变量) 对股票收益的预测能力，发现树模型 (如随机森林和提升树) 和神经网络模型的预测能力好于线性模型，这些模型也在众多特征中选出了预测能力最强的特征。Feng et al. (2019) 运用双选择的 LASSO 方法，对近 30 年来文献发现的 150 个因子进行评价和筛选。李斌等 (2019) 运用 12 种机器学习方法，发现我国 A 股市场上运用机器学习的策略能够获得比传统线性算法和所有单因子算法更好的投资绩效，并发现交易摩擦因子在 A 股市场上有较强的预测能力。在公司金融方面，于焕杰和杜子芳 (2017) 运用北京市 12 万家公司数据建立了随机森林模型，用于对各类市场主体的违法违规风险进行识别。Erel et al. (2019) 运用机器学习方法考察了美国董事个人特征对其表现的影响，发现男性、有更多的过去或现在的董事职位、更少的资格证、过大的社会网络的董事表现更差。Gow et al. (2016) 则将多种机器学习方法结合，首先运用文本分析方法分析了 CEO 的电话会议，构造 CEO 个人特征的衡量方式，再用这些特征构造机器学习模型预测公司决策和绩效。

大数据、机器学习技术正在被广泛运用于金融领域，但目前研究处于起始阶段，关于公司最重要的指标——绩效——的研究仍然很少。本文则试图运用机器学习方法探究董事会特征与公司绩效的关系，通过新的工具发现新的关系，增加我们对于董事会的公司治理作用的理解。

## 三、方法与模型

### （一）机器学习方法

本文使用了 LASSO 回归，随机森林（Random Forest）和渐进梯度回归树（Boosting 回归树）这三个机器学习方法对董事会特征与公司绩效的关系进行分析。选择 LASSO 模型是因为该模型是一个典型的正则化线性模型，它常用于筛选变量、避免过度拟合。选择随机森林和渐进梯度回归树模型是因为这两个模型属于目前最流行的集成算法，大量实践显示它们具有很好的预测能力。

### 1. LASSO（Least Absolute Shrinkage and Selection Operator）回归

LASSO 回归模型是一个线性模型，与 OLS 回归非常类似，其区别在于，OLS 以残差平方和作为损失函数，选取能够最小化残差平方和的参数作为估计值，这样会导致在样本特征很多、样本数量相对较少时出现过拟合问题。而 LASSO 回归则对损失函数添加了惩罚项，其损失函数如式（1）所示，其中  $\lambda$  为正则化参数， $\lambda > 0$ ， $i$  为观测个体， $\beta$  为线性模型系数。若  $\lambda$  取 0，该模型则退化成 OLS 模型；若将惩罚项  $\|\beta\|_1$  换为  $\|\beta\|_2^2$ ，该模型则变为岭回归模型。虽然岭回归也可以降低过拟合风险，但是 LASSO 回归的优势在于， $\beta$  的估计值中会出现更多的零向量，有助于我们在众多的、互相存在高度相关性的指标中选出重要的指标。

$$\min_{\beta} \sum (y_i - \beta^T x_i)^2 + \lambda \|\beta\|_1 \quad (1)$$

在 LASSO 回归的估计中，我们需要选择正则化参数  $\lambda$ 。我们按照通常做法，选择 5 折交叉验证后最优  $\lambda$  值进行估计，在交叉验证过程中，我们只运用训练集样本。我们运用 R 软件估计 OLS 和 LASSO 模型。

### 2. 随机森林

随机森林和梯度提升树都属于目前非常流行的集成学习方法，由于其预测性能出色，他们被大量运用于实践当中。相较于 OLS 回归和 LASSO 回归此类线性回归模型，随机森林和梯度提升树都是基于决策树的模型，他们能够处理变量之间的交互作用。然而，单独的决策树模型容易发生拟合问题，随机森林则可以很大程度上减少过拟合。

随机森林的基本思想为：从全体样本中运用随机抽取样本  $b$ ，再从特征中随机抽取  $m$  个，运用样本  $b$  和  $m$  个特征生成回归树  $T_b$ ，进行  $B$  次之后（Bootstrap 方法），对每棵树的结果取均值即为随机森林的预测值， $\hat{f}_{rf}^B(x) = \frac{1}{B} \sum_{b=1}^B T_b(x)$ 。

我们使用 R 语言的 randomForest 工具包进行随机森林模型的拟合。在随机森林方法中，主要参数有树的数量（ $B$ ）和特征个数（ $m$ ）。一般来说，树的个数越大，预测性能越好，因此我们在能力范围内选择了最大的树的数量，2000 棵<sup>5</sup>。特征个数则没有特定的规则，因此我们在主要结果中选取 R 中 randomForest 函数的设定值，即总特征个数的三分之一<sup>6</sup>，并用一系列其他取值进行稳健性检验。

### 3. 渐进梯度回归树（Gradient Boosting Regression Tree, 简称 Boosting 回归树）

渐进梯度回归树（GBRT），简称 Boosting 回归树，其基本思想是：首先，从初始训练集中得到一个回归树作为基回归树，得到当前的预测误差，然后在当前预测误差的基础上，沿着损失函数梯度下降的方向训练新的基回归树，迭代多次，最后加权结合多个基回归树得到最后的回归函数。具体的算法步骤如下：首先，设定初始的回归函数  $f_0(x) = \arg \min_c \sum_{i=1}^N L(y_i, c)$ ， $L(\cdot, \cdot)$  为损失函数，我们这里选取误差平方和作为损失函数， $f_0(x)$  即误差平方和最小的常数。然后循环以下迭代过程（ $b = 1, \dots, B$ ）：

- (1) 计算损失函数的负梯度： $z_{i,b} = -\frac{\partial L(y_i, f(x))}{\partial x} \Big|_{f(x)=f_{b-1}(x)}$ 。
- (2) 利用  $z_{i,b}$  作为残差的近似值，对  $z_{i,b}$  拟合一个新回归树  $T_b(x) = E(z|x)$ 。
- (3) 选择使得误差最小的梯度下降幅度  $\rho = \arg \min_{\rho} \sum_{i=1}^N L(y_i, \hat{f}_{b-1}(x) + \rho T_b(x))$ 。

<sup>5</sup> 一般而言，几百棵树的预测性能已经足够好了，在我们的样本中，500 棵树的结果与 2000 棵的差异并不大，但是为了结果的可靠性，我们尽可能往大选。然而树的数量过大会导致计算量过大，权衡之后我们选取 2000 棵树。

<sup>6</sup> 对于基准模型来说， $m$  为 2；对于加入董事会特征的模型来说， $m$  为 18。

(4) 经过本轮迭代后新的预测函数： $f_b(x) = f_{b-1}(x) + \nu\rho T_b(x)$ ， $\nu$ 为收缩参数。

迭代后的最终模型 $f_B(x)$ 即为我们的预测模型。我们运用 R 软件的 `gbm` 工具包进行该模型的拟合。Boosting 回归树模型的拟合过程主要有三个参数：回归树的交互深度，收缩参数 ( $\nu$ ) 和回归树的数量 ( $B$ )。交互深度指的是每个基回归树有几次分叉，更多的分叉树意味着模型考虑了变量之间更加复杂的交互性，然而，大的交互深度会大大提高模型复杂度和计算量，Hastie *et al.* (2009) 也发现在多数问题中，最优交互深度都较低，因此我们将交互深度设为 5<sup>7</sup>，此外，我们还运用 2, 4 和 6 作为稳健性检验。收缩参数即加入模型中的新的回归树的权重，也被看做是 boosting 模型的学习速率，该值一般设为 0.01 或者 0.001 (Hastie *et al.*, 2009)，越小的收缩参数几乎一定会带来更优的预测效果，但是计算成本会大大提升，因此我们将其设为 0.001，并用 0.1 和 0.01 作为稳健性检验。我们将回归树的数量最大值设置为 10000，并在每一次滚动训练样本时用 5 折交叉验证的方法确定最优的回归树数量。在交叉验证过程中，我们只运用训练集样本。

## (二) 研究设计

### 1. 基准模型和董事会特征模型

为了探究董事会特征对公司绩效的影响，我们建立了不含有董事会特征、只有公司基本特征的模型（即基准模型）和含有董事会特征和公司基本特征的模型（即董事会特征模型），分别运用 OLS 及上述三种机器学习方法建立模型。

基准模型中只包含八个目前文献中最常用的控制变量：公司规模 (*Size*)、杠杆率 (*Lev*)、固定资产比例 (*PPE*)、无形资产比例 (*Intang*)、公司年龄的对数 (*ln\_age*)、公司是否国企 (*SOE*)、第一大股东持股比例 (*Own1*) 和第十大股东持股比例 (*Own10*)。董事会特征模型除了这 8 个变量之外，还包含 46 个董事会特征变量（具体变量见第四部分）。若董事会特征对公司绩效有显著影响，那么考虑董事会特征应该能够更加精准地预测公司绩效，因此我们可以对比董事会模型和基准模型的预测能力：若董事会模型的预测能力较基准模型有大幅提升，则说明董事会特征对董事绩效有显著影响；若预测能力相当或甚至不如基准模型，则说明董事会特征的影响不大。

### 2. 模型表现评价方法

我们运用一年为窗口期进行滚动拟合，即每一年的样本仅作为训练集建立模型，运用下一年度的数据作为测试集。即运用  $t$  年度的观测值建立模型，再用  $t+1$  年的观测值对模型进行评价。为了评估机器学习模型的表现，我们也进行了 OLS 回归，以 OLS 回归的性能作为基准。这里，我们主要考察董事会特征对公司绩效的作用的三个方面：预测能力、董事会特征变量的重要性和作用模式。

在预测能力方面，我们借鉴 Gu *et al.* (2019)，运用样本外  $R^2$  ( $R^2_{\text{OOS}}$ ) 作为衡量标准进行判断。 $R^2_{\text{OOS}}$  的计算方式如式 (2) 所示：

$$R^2_{\text{OOS}} = 1 - \frac{\sum_{t=2008}^{t=2016} (y_{i,t+1} - \hat{y}_{i,t+1})^2}{\sum_{t=2008}^{t=2016} (y_{i,t+1} - \bar{y}_{i,t+1})^2} \quad (2)$$

这里  $t$  表示年度， $y_{i,t+1}$  表示  $t+1$  年（即测试集）的公司绩效， $\hat{y}_{i,t+1}$  为运用  $t$  年度数据训练出的模型对  $t+1$  年的公司绩效的预测值， $\bar{y}_{i,t+1}$  为  $t+1$  年公司绩效的均值。 $R^2_{\text{OOS}}$  越大，说明模型预测越精准。此外，为了判断模型对样本内数据的拟合程度，我们运用传统的样本内  $R^2$  ( $R^2_{\text{IS}}$ ) 进行评价。 $R^2_{\text{IS}}$  的构造与  $R^2_{\text{OOS}}$  类似， $R^2_{\text{IS}}$  越大，说明模型对样本内数据拟合程度越好。但是需要注意的是， $R^2_{\text{IS}}$  并不是越大越好，高的  $R^2_{\text{IS}}$  可能是模型过度拟合的结果，因此， $R^2_{\text{IS}}$  的结果只能作为参考，不能完全以  $R^2_{\text{IS}}$  对模型进行评价。

<sup>7</sup> 在随机森林的 `randomForest` 工具包中，我们不能直接设置交互深度，但是可以通过设置模型的最大节点数和最小节点数来控制。一般来说，在数据量少或者特征量少 的情况下，可以不限 制最大节点数，决策树会长到最大。因为我们的模型相比其他随机森林模型在特征量上并不多，因此我们并不限制该值。



在董事会特征变量的重要性（即预测能力强弱）方面，我们主要考察了随机森林和 boosting 回归树模型中各变量的重要性。由于基准模型中不含有董事会特征变量，我们在这里只考虑董事会模型。OLS 和 LASSO 这类线性模型可以直接通过系数的大小反映其重要性，而随机森林和 boosting 回归树并不是线性模型，变量的重要性无法直接观测，但是，各个机器学习方法也提供了对应的衡量变量重要性的方法。我们用随机森林中常用的均方误差增加幅度（increase in MSE）来衡量随机森林模型中变量的重要性，其基本思想是：分别对模型中每一个变量随机赋值，其他变量不变，计算该操作导致的均方误差上升的幅度，如果变量越重要，均方误差上升的幅度越大。我们用 boosting 回归树中常用的相对重要性指标（Friedman, 2001）来衡量 boosting 回归树中变量的重要性，其基本思想是：对于 Boosting 回归树的每一个决策树，计算需要评估的变量作为节点时带来的均方误差的减小幅度之和，运用所有决策树中均方误差减小值之和的均值作为其重要性的衡量，如果变量越重要，那么它应该可以带来更多的均方误差下降。变量的重要性越高，说明其对公司绩效的预测能力越强，也意味着该变量对公司绩效的影响更大。

在作用模式方面，我们考察了随机森林和 boosting 回归树的部分依赖图（Partial Dependence Plot）。部分依赖图可以展示特征变量对拟合模型预测结果的边际效应，通过部分依赖图，我们可以直观地看到特征变量对公司绩效影响的模式。部分依赖图的基本思想为：设我们需要考察的特征变量向量为  $X_S$ ，其余变量为  $X_C$ ， $X_S$  的部分依赖函数为  $f_S(X_S) = E_{X_C} f(X_S, X_C)$ ，该函数的估计值则为  $\bar{f}_S(X_S) = \frac{1}{N} \sum_{i=1}^N f(X_S, x_{ic})$ ，其中  $N$  为训练集观测值个数， $\{x_{1c}, x_{2c}, \dots, x_{Nc}\}$  为  $X_C$  在训练集中的观测值，根据  $\bar{f}_S(X_S)$  画出的图即为部分依赖图。同样，由于基准模型中不含有董事会特征变量，我们在这里只考虑董事会模型。

## 四、数据与变量

### （一）数据来源与样本

表 1 样本分布

年份	样本个数	比例 (%)
2008	1,197	6.06
2009	1,274	6.45
2010	1,389	7.04
2011	1,817	9.21
2012	2,144	10.86
2013	2,224	11.27
2014	2,148	10.88
2015	2,311	11.71
2016	2,558	12.96
2017	2,676	13.56
总计	19,738	100.00

本文数据来源于国泰安（CSMAR）数据库。由于一部分董事会特征数据起始于 2008 年，且金融受到政策管制更多，其杠杆率、董事背景与其他公司差异较大，故我们使用 2008 至 2017 年我国 A 股非金融业上市公司为样本进行分析。我们运用 stata 软件进行数据的预处理，我们剔除了公司资产、固定资产小于 0、杠杆率小于 0 或大于 1、含有缺失值的样本，并对公司层面的连续变量在 1% 和 99% 处进行缩尾处理。最终的样本包括 2960 家公司的 19738 个观测值。样本分布如表 1 所示。

### （二）变量

在公司绩效方面，我们选用了总资产负债率（ROA）和托宾 Q 值（TobinQ）分别衡量公司的财务绩效和市场绩效。

在公司基本特征方面，我们运用了八个目前文献中最常用的控制变量：公司规模（Size）、杠杆率（Lev）、固定资产比例（PPE）、无形资产比例（Intang）、公司年龄的对数（ln\_age）、公司是否国企（SOE）、第一大股东持股比例（Own1）和第十大股东持股比例（Own10）。

表 2 变量定义及描述性统计

变量名	变量定义和计算方式	样本量	均值	标准差	中位数
ROA	总资产收益率，公司息税前利润/总资产	19738	0.079	0.057	0.074
TobinQ	托宾 Q 值，（公司市值+债务面值）/总资产	19738	2.654	1.915	2.033
Size	总资产，取对数	19738	22.014	1.267	21.857
Lev	杠杆率，总负债/总资产	19738	0.445	0.211	0.442
PPE	固定资产比例，固定资产/总资产	19738	0.230	0.170	0.196
Intang	无形资产比例，无形资产/总资产	19738	0.047	0.053	0.034
ln_age	公司年龄，公司年龄加 1 后取对数	19738	2.747	0.375	2.833
SOE	国企虚拟变量，若公司为国企取 1，否则为 0	19738	0.426	0.495	0.000
Own1	公司第一大股东持股比例	19738	0.351	0.150	0.331
Own10	公司前十大股东持股比例	19738	0.569	0.155	0.577
Co_CEO	董事长与总经理是否为同一人，是为 1，否则为 0	19738	0.239	0.427	0.000
BdScale	董事会规模，即董事会人数的自然对数	19738	2.151	0.198	2.197
IndPct	独立董事比例	19738	0.371	0.052	0.333
ln_meeting	董事会会议次数，加 1 取对数	19738	2.319	0.338	2.303
ln_mean_comp	董事平均薪酬（元），取对数	19738	12.030	0.754	12.057
BoardOwn	董事会持股比例	19738	0.105	0.180	0.000
nopay_ratio	未领薪酬的董事占比	19738	0.201	0.183	0.167
founder	董事会中创始人占比	19738	0.065	0.105	0.000
prod_ratio	有生产职业背景的董事的比例	19738	0.058	0.097	0.000
rd_ratio	有研发职业背景的董事的比例	19738	0.160	0.155	0.111
design_ratio	有设计职业背景的董事的比例	19738	0.018	0.050	0.000
hr_ratio	有人力资源管理职业背景的董事的比例	19738	0.021	0.047	0.000
manage_ratio	有管理职业背景的董事的比例	19738	0.779	0.144	0.778
market_ratio	有营销职业背景的董事的比例	19738	0.130	0.127	0.111
fin_ratio	有金融职业背景的董事的比例	19738	0.192	0.154	0.167
acc_ratio	有财务职业背景的董事的比例	19738	0.238	0.118	0.222
law_ratio	有法律职业背景的董事的比例	19738	0.076	0.074	0.091
femaleratio	董事会中女性占比	19738	0.132	0.115	0.111
foreignratio	董事会中外籍成员占比 <sup>8</sup>	19738	0.009	0.036	0.000
gatratio	董事会中港、澳、台籍成员占比	19738	0.004	0.021	0.000
heternation1	国籍异质性，计算公式为 $-\sum_c p_c \ln(p_c)$ ，其中 $p_c$ 为国籍 $c$ 的董事占比	19738	0.042	0.132	0.000

<sup>8</sup> 董事国籍信息从董事简历中获得。

<i>finback_ratio</i>	曾在金融机构（或相关监管机构）任职的董事的比例	19738	0.121	0.123	0.111
<i>oversea1_ratio</i>	有海外工作背景的董事的比例	19738	0.047	0.081	0.000
<i>oversea2_ratio</i>	有海外学习背景的董事的比例	19738	0.041	0.067	0.000
<i>dazhuan_ratio</i>	大专学历的董事的比例	19738	0.073	0.108	0.000
<i>bach_ratio</i>	本科学历的董事的比例	19738	0.244	0.184	0.222
<i>master_ratio</i>	硕士学历的董事的比例	19738	0.235	0.193	0.222
<i>phd_ratio</i>	博士学历的董事的比例	19738	0.102	0.206	0.000
<i>other_ratio</i>	其他学历（如荣誉博士等）董事的比例	19738	0.027	0.066	0.000
<i>heteredu</i>	学历异质性，计算公式为 $-\sum_e p_e \ln(p_e)$ ，其中 $p_e$ 为学历 $e$ 的董事占比	19738	2.880	1.176	2.975
<i>party_ratio</i>	党员董事占比 <sup>9</sup>	19738	0.093	0.223	0.000
<i>fgo_ratio</i>	曾在政府机构任职的董事占比	19738	0.238	0.167	0.222
<i>fgo8_ratio</i>	曾在政府机构任职，级别在副处级及以上的董事占比	19738	0.175	0.147	0.143
<i>academic_ratio</i>	有学术背景的董事占比	19738	0.308	0.168	0.300
<i>mtmt_ratio</i>	在公司管理层任职的董事占比	19738	0.243	0.123	0.222
<i>ln_codir_mean</i>	平均兼任其他公司董事的公司总数，加1取对数	19738	0.627	0.442	0.560
<i>ln_age_mean</i>	董事平均年龄，加1取对数	19738	3.944	0.068	3.945
<i>age_var</i>	董事年龄的变异系数，及董事年龄的标准差/均值	19738	0.155	0.048	0.153
<i>age_skew</i>	董事年龄的偏度	19738	0.249	0.660	0.249
<i>age_kurt</i>	董事年龄的峰度	19738	2.635	0.923	2.430
<i>ln_age_gap</i>	董事年龄差异，即该公司董事年龄的最大值-最小值，加1取对数	19738	3.200	0.336	3.258
<i>ln_tenure_mean</i>	董事平均任期（月份数），加1取对数	19738	3.449	0.710	3.597
<i>tenure_var</i>	董事任期的变异系数，及董事任期的标准差/均值	19738	0.543	0.379	0.549
<i>tenure_skew</i>	董事任期的偏度	19738	-0.024	1.147	0.000
<i>tenure_kurt</i>	董事任期的峰度	19738	2.574	2.241	1.942
<i>ln_tenure_gap</i>	董事任期差异，即该公司董事任期的最大值-最小值，加1取对数	19738	3.360	1.668	4.043

在董事会特征方面，我们试图尽可能多地包含董事会特征，因此，我们根据国泰安数据库上市公司人物特征数据库和治理结构数据库中信息，构造了46个董事会特征变量。在董事会结构特征方面，我们使用了董事长是否兼任总经理（*Co\_CEO*）、董事会规模（*BdScale*）、独立董事比例（*IndPct*）这三个常用的董事会结构特征；在董事会行为方面，我们使用了董事会会议次数（*ln\_meeting*）这个指标；在董事会激励特征方面，我们运用了董事平均薪酬（*ln\_mean\_comp*）、董事会持股比例（*BoardOwn*）、未领薪酬的董事占比（*nopay\_ratio*）这三个指标；在董事个人特征方面，我们从年龄、任期、性别、学历、职业经历、国籍、政治关联等方面入手，构造了39个指标。变量的具体定义、描述性统计如表2所示。

<sup>9</sup> 董事是否党员信息从董事简历中获得。

图 1 为董事会特征变量的相关系数图。可以看到，某些董事会特征之间存在强烈的相关性，因此使用 OLS 回归可能存在严重的共线性问题，导致我们无法判断究竟是哪些董事会特征起到重要作用。

[这里插入图 1]

## 五、结果与分析

### (一) 董事会特征对公司绩效的预测能力

表 3 展示了分别运用 OLS, LASSO, 随机森林和 Boosting 回归树方法建立的基准模型和董事会模型的样本内  $R^2$  ( $R^2_{is}$ ) 和样本外  $R^2$  ( $R^2_{oos}$ )。

比较董事会模型与基准模型，可以看出，无论采取哪种方法，样本内和样本外  $R^2$  均有大幅度提升。用 ROA 衡量公司绩效时， $R^2_{is}$  提升了 0.043—0.133，与基准模型的  $R^2_{is}$  值 0.159 相比，提升幅度在 27.04%—83.65% 之间； $R^2_{oos}$  提升了 0.037—0.052，基准模型的  $R^2_{oos}$  值 0.135 相比，提升幅度在 27.41%—38.52% 之间。用 TobinQ 衡量公司绩效时， $R^2_{is}$  提升了 0.023—0.067，与基准模型的  $R^2_{is}$  值 0.421 相比，提升幅度在 5.46%—15.91% 之间； $R^2_{oos}$  提升了 0.021—0.024，基准模型的  $R^2_{oos}$  值 0.147 相比，提升幅度在 14.29%—16.33% 之间。样本内  $R^2$  ( $R^2_{is}$ ) 的结果说明，相比基准模型，董事会模型能够更好地拟合样本内数据。从样本外  $R^2$  ( $R^2_{oos}$ ) 的结果可以看出，相比基准模型，董事会模型对公司绩效的预测能力 ( $R^2_{oos}$ ) 有了大幅提升，因此，董事会特征与公司绩效之间存在密切联系。

表 3 基准模型与董事会模型的样本内和样本外  $R^2$

表 3.1 用 ROA 衡量公司绩效		OLS	LASSO	随机森林	Boosting
基准模型的 $R^2_{is}$	(1)	0.159	0.158	0.825	0.307
董事会模型的 $R^2_{is}$	(2)	0.222	0.210	0.868	0.440
基准模型的 $R^2_{oos}$	(3)	0.135	0.135	0.242	0.173
董事会模型的 $R^2_{oos}$	(4)	0.176	0.178	0.279	0.225
$\Delta R^2_{is}$	(2)-(1)	0.063	0.052	0.043	0.133
$\Delta R^2_{oos}$	(4)-(3)	0.041	0.043	0.037	0.052

表 3.2 用 TobinQ 衡量公司绩效		OLS	LASSO	随机森林	Boosting
基准模型的 $R^2_{is}$	(1)	0.421	0.421	0.903	0.659
董事会模型的 $R^2_{is}$	(2)	0.466	0.457	0.926	0.726
基准模型的 $R^2_{oos}$	(3)	0.147	0.147	0.329	0.297
董事会模型的 $R^2_{oos}$	(4)	0.169	0.171	0.353	0.318
$\Delta R^2_{is}$	(2)-(1)	0.045	0.036	0.023	0.067
$\Delta R^2_{oos}$	(4)-(3)	0.022	0.024	0.024	0.021

接下来我们比较 OLS 和三种机器学习方法。由于样本内  $R^2$  ( $R^2_{is}$ ) 高可能是过度拟合的结果，因此我们在这里主要比较样本外  $R^2$  ( $R^2_{oos}$ )。可以看到，相比 OLS, LASSO 方法对基准模型的  $R^2_{oos}$  没有提升<sup>10</sup>，而 LASSO 对董事会模型的  $R^2_{oos}$  略有提升。当公司绩效的衡量方式为 ROA 时，董事会模型的  $R^2_{oos}$  提升了 0.002，相比 OLS 结果提升了 1.14%；当公司绩效衡量方式为 TobinQ 时，

<sup>10</sup> 公司基本特征都是重要特征，LASSO 模型将这些特征全部保留，因此 LASSO 与 OLS 的估计是一样的。

董事会模型的  $R^2_{\text{oos}}$  提升了 0.002，相比 OLS 结果提升了 1.18%。这说明 LASSO 方法的预测性能较 OLS 稍有提升，但是幅度不大。相比 OLS，随机森林方法对基准模型和董事会模型的预测能力都有大幅提升，公司绩效的衡量方式为 ROA 时，基准模型的  $R^2_{\text{oos}}$  提升了 0.107，提升幅度为 79.26%，董事会模型的  $R^2_{\text{oos}}$  提升了 0.103，提升幅度为 58.52%；当公司绩效衡量方式为 TobinQ 时，基准模型的  $R^2_{\text{oos}}$  提升了 0.182，提升幅度为 123.81%，董事会模型的  $R^2_{\text{oos}}$  提升了 0.184，相比 OLS 结果提升了 108.88%。相比 OLS，Boosting 回归树对基准模型和董事会模型的预测能力也都有较大提升，公司绩效的衡量方式为 ROA 时，基准模型的  $R^2_{\text{oos}}$  提升了 0.038，提升幅度为 28.15%，董事会模型的  $R^2_{\text{oos}}$  提升了 0.049，提升幅度为 27.84%；当公司绩效衡量方式为 TobinQ 时，基准模型的  $R^2_{\text{oos}}$  提升了 0.150，提升幅度为 102.04%，董事会模型的  $R^2_{\text{oos}}$  提升了 0.149，相比 OLS 结果提升了 88.17%。综上所述，机器学习方法，特别是随机森林和 Boosting 回归树，大大提升了公司基本特征变量和董事会特征变量对公司绩效的预测能力。而 OLS 和 LASSO 模型这类线性模型的预测效果较差则在一定程度上说明董事会特征与公司绩效之间可能不是线性关系，而是存在交互关系的。

## （二）变量重要性

我们发现随机森林和 Boosting 回归树都可以大幅提升公司基本特征及董事会特征对公司绩效的预测能力，那么，这些变量中哪些能起到关键的预测作用呢？我们试图运用两个模型中各个变量的重要性来回答这个问题。

表 4 展示了重要性在前 20 名的变量。可以看到，6 个公司基本特征，即公司规模（Size）、杠杆率（Lev）、固定资产比例（PPE）、无形资产比例（Intang）、第一大股东持股比例（Own1）和第十大股东持股比例（Own10），在四个模型中均排在前列，说明这些公司基本特征确实能够对公司绩效产生很大的影响。

在董事会激励特征方面，无论是用随机森林还是 Boosting 回归树，无论用 ROA 还是 TobinQ 作为公司绩效的衡量标准，董事平均薪酬（ln\_mean\_comp）和董事会持股比例（BoardOwn）的重要性都在前 10 名之内，未领薪酬的董事占比（nopay\_ratio）的重要性排名也都在 14—32 名之间（20 名之后的排序见附录），处于中等水平。说明在众多特征中，董事会激励特征对公司绩效的预测能力最强。

在董事个人特征方面，只有董事平均任期（ln\_tenure\_mean）和学历异质性（heteredu）的重要性在四个模型中均排在前列。董事平均任期（ln\_tenure\_mean）的重要性排名在 8 至 19 名之间，而且其他董事任期相关特征（ln\_tenure\_gap, tenure\_var, tenure\_skew 和 tenure\_kurt）的排名也很高（在 9 至 30 名之间），说明董事的任期有很强的预测能力，是影响公司绩效的重要董事会特征。学历异质性（heteredu）的排名在 12 至 18 名之间，说明董事会学历的异质性对公司绩效的影响较大。而且，值得注意的是，其他董事会学历相关变量的重要性并不高，说明董事会学历的多样性比董事学历高低的影响更大。

值得注意的是，在以 TobinQ 作为公司绩效时，有学术背景的董事占比（academic\_ratio）和有研发职业背景的董事的比例（rd\_ratio）的排名均在前 20 名，而以 ROA 为公司绩效时该变量排名不高，这可能是因为有学术、研发背景的董事更加重视公司的成长和长期表现，而对短期财务表现并不敏感，故他们对企业 TobinQ 影响更大。在以 ROA 作为公司绩效的模型中，在公司管理层任职的董事占比（mtmt\_ratio）的排名在均前 20 名，而在以 TobinQ 为公司绩效时该变量排名不高，这可能是因为公司管理层任职的董事可能更加注重短期的财务表现。

董事会结构特征变量的重要性都不高。其中，董事会规模（BdScale）的重要性排名在 21 至 42 名之间，独立董事比例的重要性排名在 30 至 44 名之间，董事长是否兼任总经理（Co\_CEO）排名则在 34 至 54 名之间，这些结果说明董事会结构对公司绩效的影响并不大，他们的作用没有我们认

为的重要<sup>11</sup>。董事会行为变量，即董事会会议次数 (*ln\_meeting*) 的重要性排名在 15 至 22 名，处在中上水平，说明董事会会议次数对公司绩效也有一定影响。

综上所述，我们可以得出：董事会激励特征，特别是董事平均薪酬和董事会持股比例，都对公司绩效有很强的预测能力，是影响公司绩效的重要因素；董事个人特征中，董事平均任期和董事学历异质性有较强的预测能力，也对公司绩效有重要影响；董事会结构特征的预测能力较低，其影响较弱；董事会行为特征的预测能力处于中上水平，对公司绩效有一定影响。

表 4 变量重要性列表 (前 20 名)

变量排序	ROA 随机森林	ROA Boosting	TobinQ 随机森林	TobinQ Boosting
1	<b>Lev</b>	<b>Lev</b>	<b>Size</b>	<b>Size</b>
2	<i>ln_mean_comp</i>	<i>ln_mean_comp</i>	<b>Lev</b>	<b>Lev</b>
3	<b>Size</b>	<b>Size</b>	<b>Own10</b>	<b>PPE</b>
4	<b>Own10</b>	<b>PPE</b>	<i>BoardOwn</i>	<b>Own10</b>
5	<b>PPE</b>	<b>Own10</b>	<b>PPE</b>	<i>BoardOwn</i>
6	<i>BoardOwn</i>	<b>Intang</b>	<b>SOE</b>	<i>ln_mean_comp</i>
7	<b>Own1</b>	<b>Own1</b>	<i>founder</i>	<b>Intang</b>
8	<i>founder</i>	<i>ln_tenure_mean</i>	<i>ln_mean_comp</i>	<i>academic_ratio</i>
9	<i>ln_tenure_gap</i>	<i>ln_age_mean</i>	<b>Intang</b>	<i>ln_tenure_mean</i>
10	<b>Intang</b>	<i>BoardOwn</i>	<b>Own1</b>	<i>ln_age_mean</i>
11	<i>tenure_var</i>	<i>ln_tenure_gap</i>	<i>ln_tenure_gap</i>	<i>age_skew</i>
12	<i>tenure_skew</i>	<i>age_var</i>	<i>heteredu</i>	<i>age_kurt</i>
13	<i>heteredu</i>	<i>age_kurt</i>	<i>academic_ratio</i>	<i>age_var</i>
14	<i>nopay_ratio</i>	<i>age_skew</i>	<i>tenure_skew</i>	<b>Own1</b>
15	<b>ln_age</b>	<i>heteredu</i>	<i>ln_meeting</i>	<i>ln_codir_mean</i>
16	<i>ln_tenure_mean</i>	<i>tenure_var</i>	<i>tenure_var</i>	<i>fin_ratio</i>
17	<b>SOE</b>	<i>ln_codir_mean</i>	<i>ln_age_mean</i>	<i>ln_meeting</i>
18	<i>tenure_kurt</i>	<i>tenure_skew</i>	<b>ln_age</b>	<i>heteredu</i>
19	<i>mtmt_ratio</i>	<i>manage_ratio</i>	<i>ln_tenure_mean</i>	<i>rd_ratio</i>
20	<i>heternation1</i>	<i>mtmt_ratio</i>	<i>rd_ratio</i>	<i>acc_ratio</i>

注：加粗的变量为公司基本特征，其余为董事会特征。

### (三) 重要董事会特征对公司绩效的作用模式

在上节中，我们发现董事平均薪酬 (*ln\_mean\_comp*)、董事会持股比例 (*BoardOwn*)、董事平均任期 (*ln\_tenure\_mean*) 和董事学历异质性 (*heteredu*) 的重要性排名在所有模型中都处于前列，是影响公司绩效的重要董事会特征。那么，这些特征与公司绩效存在怎样的关系呢？接下来，我们利用部分依赖图考察这些变量对公司绩效的作用模式。

#### 1. 董事平均薪酬

图 2 是董事平均薪酬 (*ln\_mean\_comp*) 的部分依赖图。该图中 x 轴为 *ln\_mean\_comp*，处于上方的两张图随机森林模型的部分依赖图，处于下方的两张图 **Boosting** 回归树的部分依赖图。左侧图中的公司绩效衡量方式为 *ROA*，右侧为 *TobinQ*。可以看到，董事平均薪酬与公司绩效基本上呈正相关。这可能由三个原因导致：第一，高的董事薪酬激励了董事提升公司绩效；第二，高薪酬的董事本身更有能力，他们能够更好地提升公司绩效；第三，公司可能会根据其绩效来决定董事薪酬。为了排除第三个原因可能导致的影响，我们在第六部分中剔除了该变量进行稳健性检验。

<sup>11</sup> 由于独立董事比例 (*IndPct*) 的重要性排名较差，故未加入其他独立董事相关指标 (如独立董事平均任期等)。

[这里插入图 2]

## 2. 董事会持股

图 3 是董事会持股比例 (*BoardOwn*) 的部分依赖图。该图中 x 轴为 *BoardOwn*，处于上方的两张图为随机森林模型的部分依赖图，处于下方的两张图为 **Boosting** 回归树的部分依赖图。左侧图中的公司绩效衡量方式为 *ROA*，右侧为 *TobinQ*。可以看到，董事会持股比例在 0 至一个非常小的值（约为 2% 至 5%）之间，董事会持股比例提升能够带来非常大的公司绩效提升。但是超过这个阈值后，董事会持股比例对公司绩效的影响不大，甚至当董事会持股比例在 35% 至 40% 附近时，董事会持股比例越高，公司 *ROA* 会有明显降低，这可能因为董事会持股比例处于这个阶段时，大股东掏空现象比较严重。

[这里插入图 3]

## 3. 董事平均任期

图 4 是董事平均任期 (*ln\_tenure\_mean*) 的部分依赖图。该图中 x 轴为 *ln\_tenure\_mean*，处于上方的两张图为随机森林模型的部分依赖图，处于下方的两张图为 **Boosting** 回归树的部分依赖图。左侧图中的公司绩效衡量方式为 *ROA*，右侧为 *TobinQ*。可以看到，在所有图中，*ln\_tenure\_mean* 在 3.2 至 3.9 附近的区间内，公司绩效有明显的下降，当 *ln\_tenure\_mean* 过了 4 之后，公司绩效则明显回升，说明公司董事平均任期在 2 至 4 年的区间内，平均任期越长，公司绩效越差，但是过了 4 年之后，平均任期越长，公司绩效越高。该结果与 Huang and Hilary (2018) 运用美国上市公司进行检验的结果相左，Huang and Hilary (2018) 发现董事平均任期与公司绩效之间呈倒 U 型关系，他们认为新董事的“干中学” (learning-by-doing) 可以提升公司绩效，但过长的董事任期使其地位根深蒂固，反而导致公司治理下降、公司绩效降低。

[这里插入图 4]

我们的结果可能是由于我国的公司治理政策和董事会换届所致。《公司法》中规定，每一届董事的任期不能超过 3 年，《中外合资企业法》规定中外合资企业董事任期为 4 年。郑志刚等 (2016) 发现，独立董事明确提出反对意见会导致其未来连任概率降低。因此，在临近董事会换届时，独立董事可能会为了增大连任的概率更倾向于提赞同意见，而不是从公司角度出发提出意见，导致公司绩效下降。而连任之后，由于独立董事最长任期不能超过六年，他们则会从公司利益角度提出意见，从而改善公司治理、提升公司绩效。而美国没有对董事任期的规定和限制，所以运用美国数据的结果与我国数据的结果有较大差异。

## 4. 学历异质性

图 5 是董事学历异质性的部分依赖图。该图中 x 轴为 *heteredu*，处于上方的两张图为随机森林模型的部分依赖图，处于下方的两张图为 **Boosting** 回归树的部分依赖图。左侧图中的公司绩效衡量方式为 *ROA*，右侧为 *TobinQ*。可以看到，董事会学历异质性与公司绩效呈 U 型关系<sup>12</sup>，即随着学历异质性程度增加，公司绩效先降低后升高。这可能是因为当董事学历程度类似时，董事会成员的沟通更通畅，董事会效率更高，但更多样的学历背景的董事也可以能力互补，因此高的学历异质性也会提升公司绩效。

[这里插入图 5]

# 六、稳健性检验

## (一) 更换训练集滚动时间窗口

---

<sup>12</sup> 用 **Boosting** 回归树预测 *TobinQ* 的部分依赖图 (右下) 中，*heteredu* 与 *TobinQ* 在 *heteredu* 小于 3 的区域内呈负相关，但由于负相关程度较低，图中表现并不明显。

在上文中我们运用一年为窗口期进行滚动拟合。为了保证结果的稳健性，我们换用3年为窗口期再次建立模型，即用  $t$  至  $t+2$  年的数据作为训练集建立模型，用  $t+3$  年的数据作为测试集评价模型效果。表5展示了运用3年为窗口期滚动拟合的结果，可以看到，无论运用何种模型，董事会模型的样本内和样本外  $R^2$  均有大幅提升，而且随机森林与 Boosting 回归树的解释与预测能力也都大大强于 OLS 和 LASSO 模型。

表5 三年期滚动窗口拟合结果

表 5.1 用 ROA 衡量公司绩效		OLS	LASSO	随机森林	Boosting
基准模型的 $R^2_{is}$	(1)	0.147	0.147	0.849	0.301
董事会模型的 $R^2_{is}$	(2)	0.205	0.203	0.888	0.378
基准模型的 $R^2_{oos}$	(3)	0.100	0.101	0.225	0.149
董事会模型的 $R^2_{oos}$	(4)	0.154	0.154	0.270	0.200
$\Delta R^2_{is}$	(2)-(1)	0.058	0.056	0.039	0.077
$\Delta R^2_{oos}$	(4)-(3)	0.054	0.053	0.045	0.051

表 5.2 用 TobinQ 衡量公司绩效		OLS	LASSO	随机森林	Boosting
基准模型的 $R^2_{is}$	(1)	0.323	0.323	0.892	0.565
董事会模型的 $R^2_{is}$	(2)	0.371	0.370	0.922	0.632
基准模型的 $R^2_{oos}$	(3)	0.117	0.117	0.305	0.257
董事会模型的 $R^2_{oos}$	(4)	0.137	0.139	0.331	0.290
$\Delta R^2_{is}$	(2)-(1)	0.048	0.047	0.030	0.067
$\Delta R^2_{oos}$	(4)-(3)	0.020	0.022	0.026	0.033

## (二) 改变模型参数

随机森林模型和 Boosting 回归树模型都需要设定一些参数，这些参数对模型构建也有一定影响，为了保证结果的稳健性，我们运用不同参数再次建立模型。

表6 更换随机森林模型的参数

		$m = 7$	$m = 12$	$m = 22$	$m = 27$
ROA	$R^2_{is}$	0.854	0.864	0.869	0.870
	$R^2_{oos}$	0.264	0.276	0.279	0.278
TobinQ	$R^2_{is}$	0.907	0.920	0.927	0.929
	$R^2_{oos}$	0.314	0.342	0.356	0.356

### 1. 随机森林

影响随机森林模型的重要参数为特征个数  $m$ ，在正文中我们运用变量个数的三分之一作为特征个数，在这里我们选取了4个特征个数：7，12，22和27<sup>13</sup>，进行稳健性检验，表6展示了该结果。可以看到，这些结果与主要结果非常接近，说明无论选取哪个特征个数，随机森林的拟合结果均好于 OLS 和 LASSO。

### 2. Boosting 回归树

<sup>13</sup> 选取7是因为这是变量个数的平方根。分类问题一般选取变量个数的平方根作为特征个数。选取27是因为这是变量个数的二分之一。为了更加稳健，我们在7至18，18至27之间再选取其中间值（12和22）作为特征个数。



影响 Boosting 回归树的重要参数为回归树的交互深度和收缩参数 ( $\nu$ )。在正文中, 我们将交互深度设为 5、收缩参数设为 0.001, 这里, 我们选取 2, 4, 6 这三个交互深度和 0.1, 0.01, 0.001 三个收缩参数, 运用这 9 组参数重新建立模型, 结果如表 7 所示。可以看到, 无论采用哪个参数组合, Boosting 回归树的拟合结果均好于 OLS 和 LASSO。

表 7 更换 Boosting 回归树模型的参数

	收缩参数 交互深度	$\nu = 0.001$			$\nu = 0.01$			$\nu = 0.1$		
		2	4	6	2	4	6	2	4	6
<i>ROA</i>	$R^2_{is}$	0.333	0.397	0.431	0.339	0.407	0.429	0.330	0.356	0.360
	$R^2_{oos}$	0.207	0.221	0.228	0.206	0.220	0.226	0.199	0.197	0.201
<i>TobinQ</i>	$R^2_{is}$	0.666	0.720	0.747	0.690	0.730	0.746	0.670	0.708	0.713
	$R^2_{oos}$	0.302	0.318	0.325	0.307	0.317	0.323	0.299	0.309	0.305

### (三) 去除董事平均薪酬变量

由于公司可能根据其绩效来决定董事薪酬, 导致董事薪酬与公司绩效呈高度相关, 从而导致我们高估了董事会特征的影响。为了避免问题可能造成的影响, 我们剔除董事平均薪酬变量, 重新建立了 OLS, LASSO, 随机森林和 Boosting 回归树模型。结果如表 8 所示, 可以看到, 当公司绩效衡量指标为 *TobinQ* 时, 该结果与主要结果非常接近; 当公司绩效衡量指标为 *ROA* 时, 董事会模型的样本内  $R^2$  和样本外  $R^2$  均有下降, 但是相比基准模型, 董事会模型的  $R_{is}$  和  $R_{oos}$  仍有显著提高。这个结果也在一定程度上意味着董事薪酬更多地取决于财务绩效而非市场绩效。

表 8 除去董事平均薪酬的结果

表 8.1 用 <i>ROA</i> 衡量公司绩效		OLS	LASSO	随机森林	Boosting
基准模型的 $R^2_{is}$	(1)	0.159	0.158	0.825	0.307
董事会模型的 $R^2_{is}$	(2)	0.193	0.180	0.862	0.383
基准模型的 $R^2_{oos}$	(3)	0.135	0.135	0.242	0.173
董事会模型的 $R^2_{oos}$	(4)	0.145	0.149	0.263	0.197
$\Delta R^2_{is}$	(2)-(1)	0.034	0.022	0.037	0.076
$\Delta R^2_{oos}$	(4)-(3)	0.010	0.014	0.021	0.024
表 8.2 用 <i>TobinQ</i> 衡量公司绩效		OLS	LASSO	随机森林	Boosting
基准模型的 $R^2_{is}$	(1)	0.421	0.421	0.903	0.659
董事会模型的 $R^2_{is}$	(2)	0.465	0.457	0.925	0.720
基准模型的 $R^2_{oos}$	(3)	0.147	0.147	0.329	0.297
董事会模型的 $R^2_{oos}$	(4)	0.168	0.171	0.352	0.318
$\Delta R^2_{is}$	(2)-(1)	0.044	0.036	0.022	0.061
$\Delta R^2_{oos}$	(4)-(3)	0.021	0.024	0.023	0.021

## 七、结 论

为了探究我国董事会特征如何影响公司绩效，本文运用 LASSO，随机森林和渐进梯度回归树三种机器学习方法，研究了 46 个董事会特征对公司绩效的预测能力。对比加入董事会特征的模型与只包含公司基本特征的模型，可以发现董事会特征对公司绩效有较强的预测能力。对比 OLS 模型，我们发现三个机器学习算法都能够提高董事会特征模型对公司绩效的预测能力，而且随机森林与渐进梯度提升树模型的预测能力都远远高于 OLS 和 LASSO 模型。接下来，我们运用随机森林和渐进梯度提升树的变量重要性排名考察了这些董事会特征对公司绩效影响的大小，发现董事平均薪酬、董事会持股、董事平均任期和董事学历异质性在各个排序中都名列前茅，说明这些指标是影响公司绩效的重要董事会特征。我们还运用部分依赖图考察了董事平均薪酬、董事会持股、董事平均任期和董事学历异质性对公司绩效的作用关系，发现董事平均薪酬越高，公司绩效越好，但董事会持股、董事平均任期、董事学历异质性与公司绩效呈非线性关系。最后，我们更换了训练集时间窗口、改变模型参数、去除董事平均薪酬指标，发现董事会特征仍然可以解释和预测公司绩效。

本文的结果说明，我国公司中董事会特征与公司绩效有紧密且复杂的联系。在众多特征中，董事平均薪酬、董事会持股比例、董事平均任期和学历异质性对公司绩效的影响较大，而董事会规模、独立董事比例、董事长与总经理是否兼任这些董事会结构指标的影响并没有我们想象中大。此外，很多董事会特征公司绩效之间的关系是非线性的。这些发现都有助于加深我们对董事会在公司治理中作用的理解。

本文也在一定程度上弥补了传统的董事会研究的不足：以往文献从各个方面研究了董事会特征与公司治理、公司绩效的关系，但是这些文献大多都只考虑单个董事会特征，而且目前广泛使用的线性回归方法也很难识别董事会特征中的非线性的、交互的关系，而本文则运用机器学习方法同时探究了 46 个董事会特征与公司绩效之间的关系，在一定程度上解决了这两个问题。本文也丰富了目前机器学习在公司金融中的应用的相关研究。本文的研究方法和发现都有一定的现实意义：第一，我们的发现有助于公司股东选择合适的董事、组成高效的董事会；第二，我们发现运用机器学习方法可以帮助投资者和分析师提升对公司绩效预测的准确性；第三，我们的发现也可以作为相关部门制定董事会政策的参考。

最后，本文仍然存在一些限制与不足。一方面，由于数据所限，我们无法穷尽所有的董事会特征；另一方面，机器学习方法远不只我们所使用的三种，但由于计算能力所限，我们也无法穷尽所有的方法。之后的研究可以尝试运用文本分析方法考察更多的董事会特征、并运用其他的机器学习方法（如神经网络等）考察董事会特征与公司绩效的关系。

## 参考文献

- [1] 邓峰, 2011, 董事会制度的起源、演进与中国的学习, 《中国社会科学》, 第1期, 第165-177+224页。
- [2] 雷海民、梁巧转、李家军, 2012, 公司治理影响公司运营效率吗——基于中国上市公司的非参数检验, 《中国工业经济》, 第9期, 第109-121页。
- [3] 李斌、邵新月、李玥阳, 2019, 机器学习驱动的基本面量化投资研究, 《中国工业经济》, 第8期, 第61-78页。
- [4] 李晓, 2015, 两职兼任、独立董事与公司绩效的关联性——基于中国证券市场的经验检验, 《财会月刊》, 第30期, 第51-56页。
- [5] 林乐、谢德仁, 2007, 分析师荐股更新利用管理层语调吗? ——基于业绩说明会的文本分析, 《管理世界》, 2007年第11期, 第125-145+188页。
- [6] 马黎璐、伊志宏、张澈, 2019, 廉价交谈还是言之有据——分析师报告文本的信息含量研究, 《管理世界》, 第7期, 第182-200页。
- [7] 牛建波、李胜楠, 2007, 控股股东两权偏离、董事会行为与企业价值: 基于中国民营企业面板数据的比较研究, 《南开管理评论》, 第2期, 第31-37页。
- [8] 沈艳、陈赟、黄卓, 2019, 文本大数据分析在经济学和金融学中的应用: 一个文献综述, 《经济学(季刊)》, 第4期, 第1153-1186页。
- [9] 沈艺峰、王夫乐、黄娟娟、纪荣嵘, 2017, 高管之“人”的先天特征在IPO市场中起作用吗? 《管理世界》, 第9期, 第141-154+188页。
- [10] 王克敏、王华杰、李栋栋、戴杏云, 2008, 年报文本信息复杂性与管理者自利——来自中国上市公司的证据, 《管理世界》, 第12期, 第120-132+194页。
- [11] 王跃堂、赵子夜、魏晓雁, 2006, 董事会的独立性是否影响公司绩效? 《经济研究》, 第5期, 第62-73页。
- [12] 魏刚、肖泽忠、N. Travlos、邹宏, 2007, 独立董事背景与公司经营绩效, 《经济研究》, 第3期, 第92-105+156页。
- [13] 谢德仁、林乐, 2015, 管理层语调能预示公司未来业绩吗? ——基于我国上市公司年度业绩说明会的文本分析, 《会计研究》, 第2期, 第20-27+93页。
- [14] 杨青、高铭、B. B. Yurtoglu, 2009, 董事薪酬、CEO薪酬与公司业绩: 合谋还是共同激励? 《金融研究》, 第6期, 第111-127页。
- [15] 于东智, 2003, 董事会、公司治理与绩效——对中国上市公司的经验分析, 《中国社会科学》, 第3期, 第29-41+205-206页。
- [16] 于东智, 2001, 董事会行为、治理效率与公司绩效——基于中国上市公司的实证分析, 《管理世界》, 第2期, 第200-202页。
- [17] 于东智、池国华, 2004, 董事会规模、稳定性与公司绩效: 理论与经验分析, 《经济研究》, 第4期, 第70-79页。
- [18] 于焕杰、杜子芳, 2017, 基于随机森林的企业监管方法研究, 《管理世界》, 第9期, 第180-181页。
- [19] 郑志刚、李俊强、黄继承、胡波, 2016, 独立董事否定意见发表与换届未连任, 《金融研究》, 第12期, 第159-174页。
- [20] 周志华, 2016, 《机器学习》, 北京: 清华大学出版社。
- [21] 朱滔, 2015, 董事薪酬、CEO薪酬与公司未来业绩: 监督还是合谋? 《会计研究》, 第8期, 第159-174页。
- [22] Adams, R. B., and D. Ferreira, 2008a, Do directors perform for pay?, *Journal of Accounting Economics*, 46(1), 154-171.
- [23] Adams, R. B., and D. Ferreira, 2008b, Women in the boardroom and their impact on governance and performance, *Journal of Financial Economics*, 94(2), 291-309.
- [24] Adams, R. B., 2017, Boards, and the directors who sit on them, *The Handbook of the Economics of Corporate Governance*, 1, 291-382.
- [25] Adams, R. B., A. C. Akyol, and P. Verwijmeren, 2018, Director skill sets, *Journal of Financial Economics*, 130(3), 641-662.
- [26] Arrow, K., 1951, *Social Choice and Individual Values*. New York: John Wiley and Sons, Inc..
- [27] Bandiera, O., S. Hansen, A. Prat, and R. Sadun, 2017, CEO behavior and firm performance, NBER Working paper.
- [28] Bhagat, S., and B. Black, 2002, The non-correlation between board independence and long-term firm performance, *Journal of Corporation Law*, 27(2), 231-273.
- [29] Bhagat, S., and B. Bolton, 2013, Director ownership, governance, and performance, *Journal of Financial and Quantitative Analysis*, 48(1), 105-135.

- [30] Black, B., and W. Kim, 2012, The effect of board structure on firm value: A multiple identification strategies approach using Korean data, *Journal of Financial Economics*, 104(1), 203-226.
- [31] Brick I. E., O. Palmon, and J. K. Wald, 2006, CEO compensation, director compensation, and firm performance: Evidence of cronyism? *Journal of Corporate Finance*, 12(3), 403-423.
- [32] Chen Q., I. Goldstein, and W. Jiang, 2008, Directors' ownership in the US mutual fund industry, *Journal of Finance*, 2008, 63(6), 2629-2677.
- [33] Cheng, S., 2008, Board size and the variability of corporate performance, *Journal of Financial Economics*, 87(1), 157-176.
- [34] Dahya, J., O. Dimitrov, and J. J. McConnell, 2008, Dominant shareholders, corporate boards, and corporate value: A cross-country analysis, *Journal of Financial Economics*, 87(1), 73-100.
- [35] Dalton, D. R., C. M. Daily, J. L. Johnson, and A. E. Ellstrand, 1999, Number of directors and financial performance: A meta-analysis, *Academy of Management Journal*, 42(6), 674-686.
- [36] Erel, I., L. H. Stern, C. Tan, and M. S. Weisbach, 2019, Selecting directors using machine learning, NBER Working paper, 2019.
- [37] Erhardt, N. L., J. D. Werbel, and C. B. Shrader, 2003, Board of director diversity and firm financial performance, *Corporate Governance: An International Review*, 11(2), 102-111.
- [38] Feng, G., S. Giglio, and D. Xu, 2019, Taming the factor zoo: A test of new factors, NBER Working paper.
- [39] Friedman, J. H., 2001, Greedy function approximation: A gradient boosting machine, *The Annals of Statistics*, 29(5), 1189-1232.
- [40] Gow, I. D., S. N. Kaplan, D. F. Larcker, and A. A. Zakolyukina, 2016, CEO personality and firm policies, NBER Working paper.
- [41] Gu, S., B. Kelly, and D. Xiu, 2019, Empirical asset pricing via machine learning, NBER Working paper.
- [42] Hastie, T., R. Tibshirani, and J. Friedman, 2009, *The Elements of Statistical Learning*, New York: Springer.
- [43] Hermalin, B. E., and M. S. Weisbach, 2003, Board of directors as an endogenously determined institution: A survey of the economic literature, *Federal Reserve Bank of New York Economic Policy Review*, 9(1), 7-26.
- [44] Heston, S. L., and N. R. Sinha, 2017, News vs. sentiment: Predicting stock returns from news stories, *Financial Analysts Journal*, 73(3), 67-83.
- [45] Huang, A. H., A. Y. Zang, and R. Zheng, 2014, Evidence on the information content of text in analyst reports, *The Accounting Review*, 89(6), 2151-2180.
- [46] Huang, S., and G. Hilary, 2018, Zombie board: Board tenure and firm performance, *Journal of Accounting Research*, 56(4), 1285-1329.
- [47] Jayaraman, N., V. K. Nanda, and H. E. Ryan, 2015, Does combining the CEO and chair roles cause poor firm performance?, SSRN Working Paper.
- [48] Jegadeesh, N., and D. Wu, 2013, Word power: A new approach for content analysis, *Journal of Financial Economics*, 110(3), 712-729.
- [49] Jensen, M. C., 1993, The modern industrial revolution, exit, and the failure of internal control systems, *Journal of Finance*, 48(3), 831-880.
- [50] Jiang, W., H. Wan, and S. Zhao, 2016, Reputation concerns of independent directors: Evidence from individual director voting, *Review of Financial Studies*, 29(3), 655-696.
- [51] Kim, D., and L. T. Starks, 2016, Gender diversity on corporate boards: Do women contribute unique skills? *American Economic Review*, 106(5), 267-271.
- [52] Kumar, P., and K. Sivaramakrishnan, 2008, Who monitors the monitor? The effect of board independence on executive compensation and firm value, *Review of Financial Studies*, 21(3), 1371-1401.
- [53] Li, F., 2010, The information content of forward-looking statements in corporate filings-A naïve Bayesian machine learning approach, *Journal of Accounting Research*, 48(5), 1049-1102.
- [54] Lipton, M., and J. W. Lorsch, 1992, A modest proposal for improved corporate governance, *The Business Lawyer*, 48, 59-77.
- [55] Liu, Y., M. K. Miletkov, Z. Wei, and T. Yang, 2015, Board independence and firm performance in China, *Journal of Corporate Finance*, 30, 223-244.
- [56] Liu, Y., Z. Wei, and F. Xie, 2014, Do women directors improve firm performance in China? *Journal of Corporate Finance*, 28, 169-184.
- [57] Maug, E., 1997, Board of directors and capital structure: Alternative forms of corporate restructuring, *Journal of Corporate Finance*, 3, 113-139.
- [58] Shehata, N., A. Salhin, and M. El-Helaly, 2017, Board diversity and firm performance: Evidence from SMEs, *Applied Economics*, 49(46-48), 4817-4832.
- [59] Shmueli, G., 2010, To explain or to predict? *Statistical Science*, 25(3), 289-310.
- [60] Tetlock, P. C., 2007, Giving content to investor sentiment: The role of media in the stock media, *Journal of Finance*, 62(3), 1139-1168.

- [61] Tetlock, P. C., M. Saar-Tsechansky, and S. Macskassy, 2008, More than words: Quantifying language to measure fundamentals, *Journal of Finance*, 63(3), 1437-1467.
- [62] Vafeas, N., 1999, Board meeting frequency and firm performance, *Journal of Financial Economics*, 53(1), 113-142.
- [63] Yermack, D., 1996, Higher market valuation of companies with a small board of directors, *Journal of Financial Economics*, 40(2), 185-211.

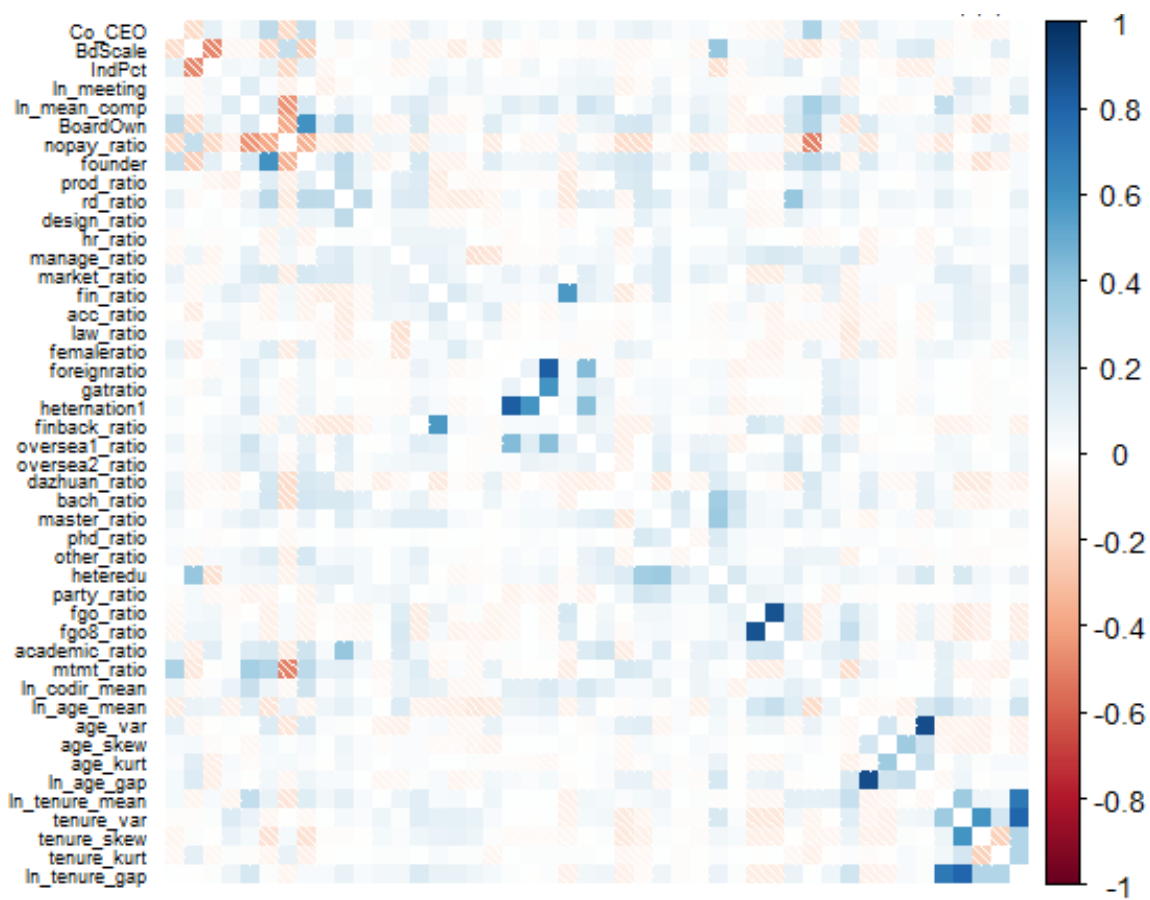
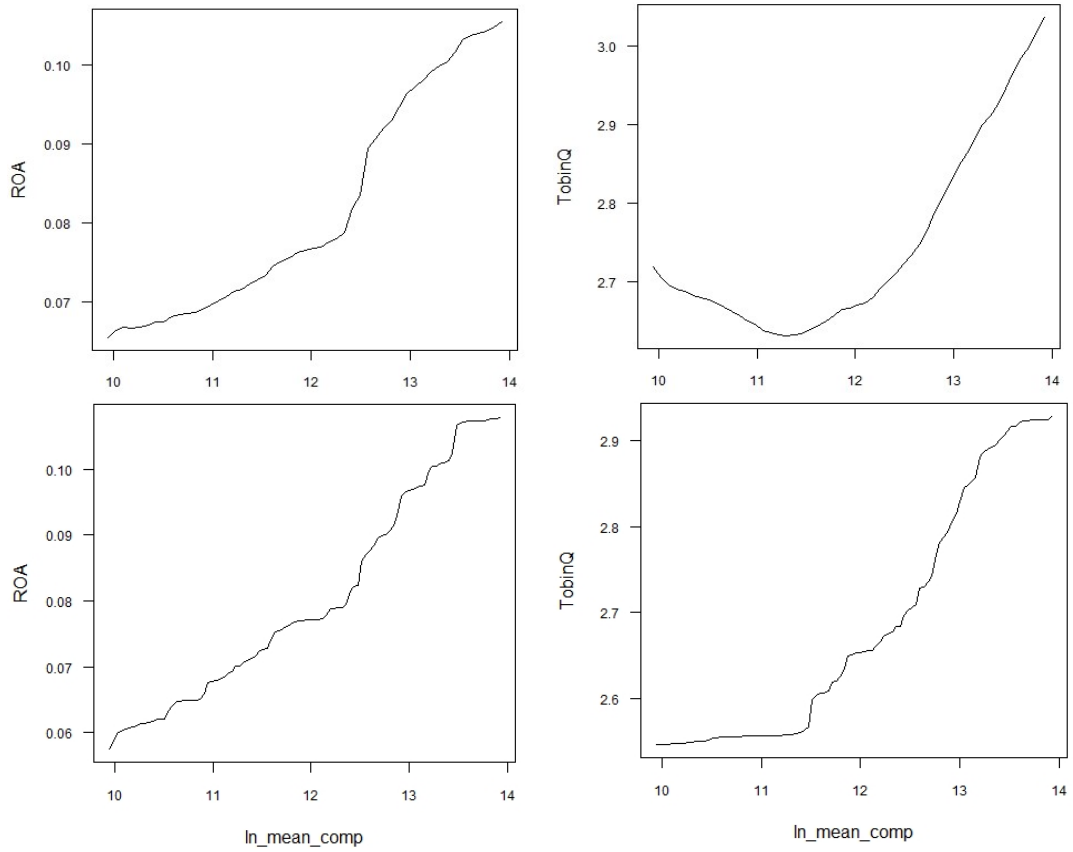


图 1 相关系数图



**图 2 董事平均薪酬的部分依赖图**

注：左上图为用随机森林预测 ROA 的部分依赖图；右上图为用随机森林预测托宾 Q 的部分依赖图；左下图为用 boosting 回归树预测 ROA 的部分依赖图；右下图为用 boosting 回归树预测托宾 Q 的部分依赖图，下同。

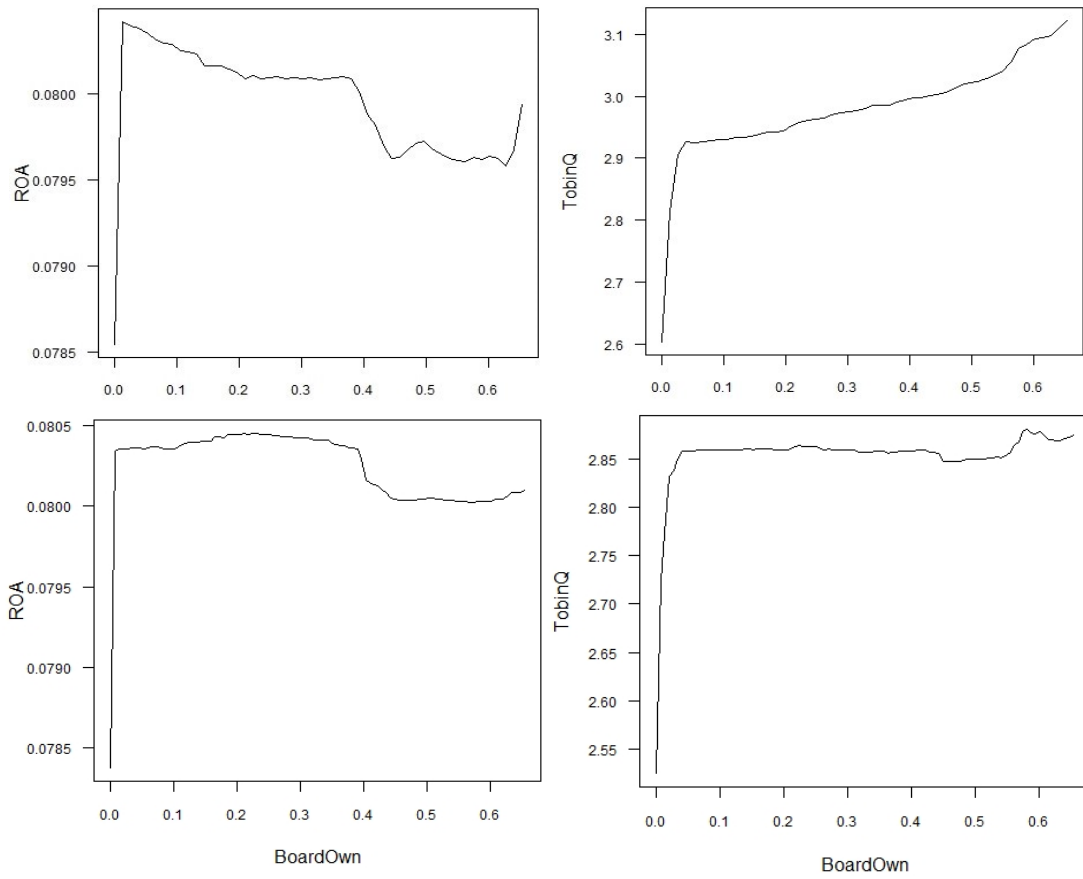


图3 董事会持股比例的部分依赖图



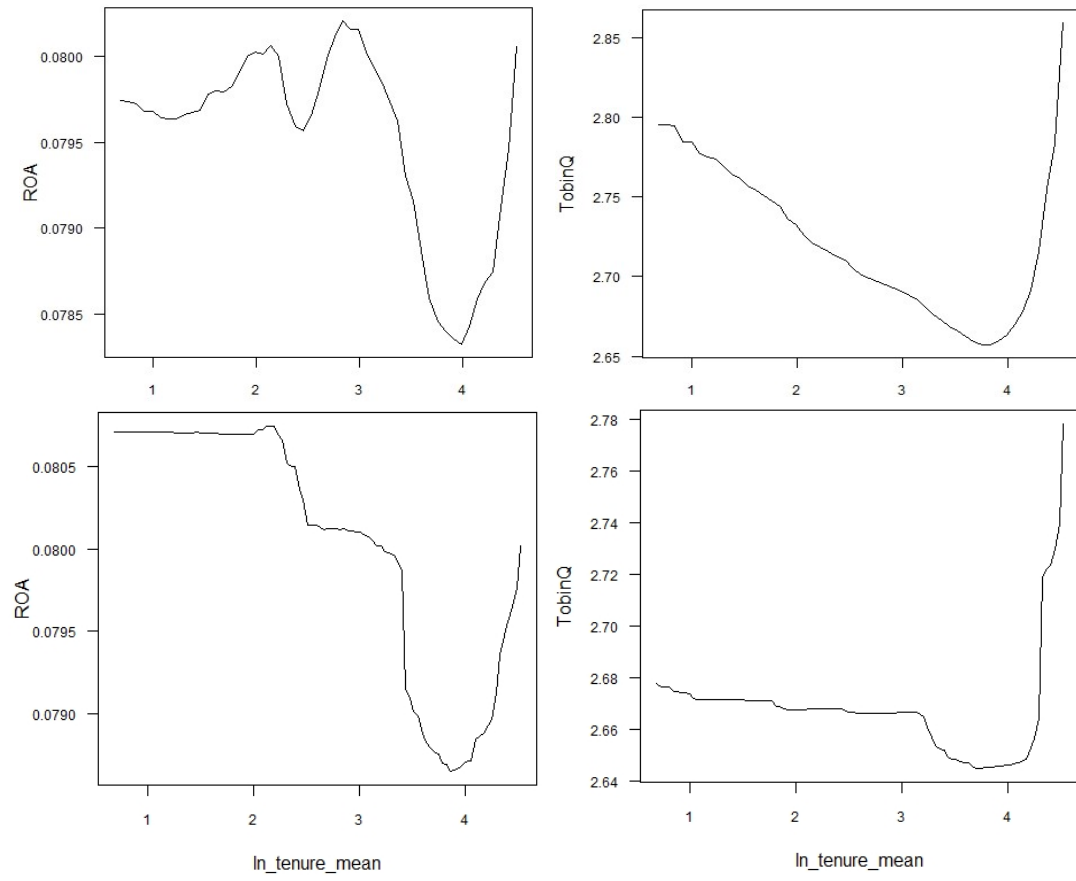


图 4 董事平均任期的部分依赖图

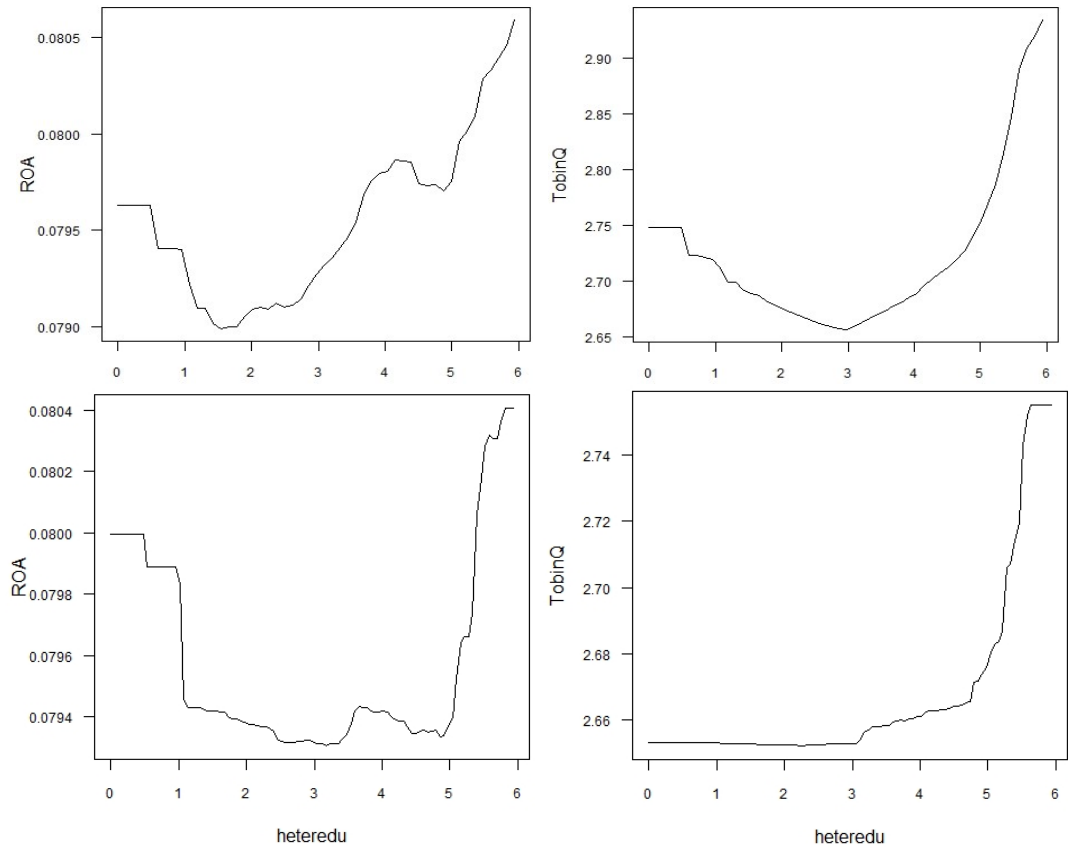


图5 董事学历异质性的部分依赖图

## 附录

附表 1 变量重要性排序 (第 21 至 54 名)

变量排序	ROA	ROA	TobinQ	TobinQ
	随机森林	Boosting	随机森林	Boosting
21	<i>BdScale</i>	<i>ln_meeting</i>	<i>dazhuan_ratio</i>	<i>finback_ratio</i>
22	<i>ln_meeting</i>	<i>nopay_ratio</i>	<i>mtmt_ratio</i>	<i>tenure_var</i>
23	<i>age_var</i>	<b><i>ln_age</i></b>	<i>nopay_ratio</i>	<i>tenure_skew</i>
24	<i>ln_age_gap</i>	<i>academic_ratio</i>	<i>tenure_kurt</i>	<i>manage_ratio</i>
25	<i>rd_ratio</i>	<i>acc_ratio</i>	<i>BdScale</i>	<i>oversea1_ratio</i>
26	<i>fgo8_ratio</i>	<i>fin_ratio</i>	<i>finback_ratio</i>	<i>ln_tenure_gap</i>
27	<i>prod_ratio</i>	<i>tenure_kurt</i>	<i>prod_ratio</i>	<b><i>ln_age</i></b>
28	<i>ln_age_mean</i>	<i>market_ratio</i>	<i>ln_age_gap</i>	<i>fgo_ratio</i>
29	<i>bach_ratio</i>	<i>master_ratio</i>	<i>fin_ratio</i>	<i>master_ratio</i>
30	<i>IndPct</i>	<i>bach_ratio</i>	<i>fgo8_ratio</i>	<i>tenure_kurt</i>
31	<i>ln_codir_mean</i>	<i>rd_ratio</i>	<i>age_kurt</i>	<i>founder</i>
32	<i>foreignratio</i>	<i>fgo_ratio</i>	<i>master_ratio</i>	<i>nopay_ratio</i>
33	<i>age_skew</i>	<i>finback_ratio</i>	<i>age_var</i>	<i>mtmt_ratio</i>
34	<i>age_kurt</i>	<i>femaleratio</i>	<i>oversea1_ratio</i>	<i>femaleratio</i>
35	<i>fgo_ratio</i>	<i>ln_age_gap</i>	<i>ln_codir_mean</i>	<i>market_ratio</i>
36	<i>Co_CEO</i>	<i>fgo8_ratio</i>	<i>age_skew</i>	<i>bach_ratio</i>
37	<i>oversea1_ratio</i>	<i>law_ratio</i>	<i>fgo_ratio</i>	<i>ln_age_gap</i>
38	<i>dazhuan_ratio</i>	<i>IndPct</i>	<i>manage_ratio</i>	<i>fgo8_ratio</i>
39	<i>finback_ratio</i>	<i>prod_ratio</i>	<i>oversea2_ratio</i>	<i>dazhuan_ratio</i>
40	<i>design_ratio</i>	<i>BdScale</i>	<i>bach_ratio</i>	<i>foreignratio</i>
41	<i>master_ratio</i>	<i>oversea1_ratio</i>	<i>IndPct</i>	<i>law_ratio</i>
42	<i>gatratio</i>	<i>dazhuan_ratio</i>	<i>acc_ratio</i>	<i>prod_ratio</i>
43	<i>fin_ratio</i>	<i>hr_ratio</i>	<i>party_ratio</i>	<i>BdScale</i>
44	<i>academic_ratio</i>	<i>other_ratio</i>	<i>hr_ratio</i>	<i>IndPct</i>
45	<i>market_ratio</i>	<i>founder</i>	<i>Co_CEO</i>	<b><i>SOE</i></b>
46	<i>acc_ratio</i>	<i>foreignratio</i>	<i>foreignratio</i>	<i>party_ratio</i>
47	<i>other_ratio</i>	<i>party_ratio</i>	<i>design_ratio</i>	<i>phd_ratio</i>
48	<i>oversea2_ratio</i>	<i>design_ratio</i>	<i>law_ratio</i>	<i>oversea2_ratio</i>
49	<i>party_ratio</i>	<i>oversea2_ratio</i>	<i>heternation1</i>	<i>heternation1</i>
50	<i>manage_ratio</i>	<i>heternation1</i>	<i>phd_ratio</i>	<i>other_ratio</i>
51	<i>hr_ratio</i>	<i>phd_ratio</i>	<i>gatratio</i>	<i>hr_ratio</i>
52	<i>femaleratio</i>	<b><i>SOE</i></b>	<i>femaleratio</i>	<i>design_ratio</i>
53	<i>phd_ratio</i>	<i>gatratio</i>	<i>market_ratio</i>	<i>Co_CEO</i>
54	<i>law_ratio</i>	<i>Co_CEO</i>	<i>other_ratio</i>	<i>gatratio</i>