

中国上市公司分红的动因研究

——基于机器学习的证据

陈运森, 周金泳, 彭嘉续

[摘要] 企业分红不仅是股东投资回报的直接体现,更是“建设以投资者为本的资本市场”的重要手段,但“股利之谜”一直是理论与实践中的基础性难题。本文采用机器学习方法,评估了不同类别公司特征对分红行为的预测效果,旨在更准确地解读中国企业的分红动因。结果发现:①以渐进梯度回归树和随机森林为代表的集成学习方法在预测公司分红行为方面表现突出,显著优于多元线性回归等传统统计模型。②生命周期特征和公司税率特征对中国资本市场分红实践的预测效果较好,第一类和第二类代理问题相关特征次之,而融资需求特征和投资者情绪特征在企业分红行为的预测效果上相对较弱。③在各类分红动因变量中,留存收益资产比、公司税率特征、其他应收款资产比和上一期股利支付水平对分红水平的预测贡献较大。异质性分析显示,影响企业分红的主要因素在半强制分红政策、股息税改革和《现金分红指引》等股利政策颁布前后存在差异,在不同产权性质以及自由现金流状况的企业之间也有所不同,是否采取股利迎合策略和投资者现金股利情绪也会影响公司的现金分红行为。本文为“加强现金分红监管、增强投资者回报”的监管实践提供了理论支持,也为深入理解中国现代资本市场的分红实践和活跃资本市场提供了启示。

[关键词] 股利政策; 企业分红; 机器学习; 股利之谜

[中图分类号] F275 **[文献标识码]** A **[文章编号]** 1006-480X(2024)05-0155-19

一、引言

近年来,中国证券监督管理委员会(简称中国证监会)反复强调要“建设以投资者为本的资本市场”,而“活跃资本市场,提振投资者信心”也是2023年8月中共中央政治局会议对资本市场工作的重要部署。企业分红是向投资者返还利润的重要形式,有助于增强投资者信心。分红回报正是检

[收稿日期] 2024-01-27

[基金项目] 国家自然科学基金面上项目“随机监管的资本市场经济后果及监管联动性研究:基于证监会双随机检查的准自然实验”(批准号 72272168);国家自然科学基金青年项目“上市公司研发投入的数据缺失:基于披露机制建模和机器学习的方法与应用”(批准号 72201288);国家自然科学基金面上项目“企业数字化转型的赋能机制与赋能效果研究:基于人力资本结构调整视角”(批准号 72272167)。

[作者简介] 陈运森,中央财经大学会计学院教授,博士生导师,管理学博士;周金泳,中央财经大学会计学院博士研究生;彭嘉续,中央财经大学会计学院讲师,管理学博士。通讯作者:彭嘉续,电子邮箱:pengjiaxu@cufe.edu.cn。感谢匿名评审专家和编辑部的宝贵意见,文责自负。

验企业是否真正践行“以投资者为本”的试金石。优化和完善分红制度体系建设,也是响应2024年《政府工作报告》强调的“完善中国特色现代企业制度,打造更多世界一流企业”的重要一环。监管层致力于确保公司制定合理的股利政策,通过加强对上市公司股利政策的监督和指导,充分保护投资者利益,进而促进市场稳定和健康发展。2024年3月,中国证监会发布《关于加强上市公司监管的意见(试行)》,更是强调“加强现金分红监管,增强投资者回报”。2018—2022年,中国A股上市公司累计分红8.20万亿元,其中,2022年共有3291家上市公司进行了现金分红,分红金额达2.13万亿元,相比于全球主要资本市场,企业平均股利支付率处于中上游水平。然而,“企业为何分红”却是理论和实务界悬而未决的基础性问题,即国内外学术界经典的“股利之谜”。尽管目前有许多研究尝试对公司分红行为进行解释,但尚未得出统一结论,中国资本市场企业分红的实际动机仍然是一个“黑匣子”。造成这一困境的重要原因在于公司分红决策受多方面因素影响,如公司治理水平、外部环境和公司基本面特征等。因此,这种复杂性要求从更加综合的角度深入理解公司的分红决策,单一类型的特征及其背后反映的股利理论难以全面解释各种分红行为。基于此,本文的研究目标是评估不同类别公司特征对中国资本市场分红实践的预测能力,进而识别出影响中国特色现代资本市场中上市公司分红行为的重要指标。这不仅有助于为监管部门进一步优化公司分红制度提供启示,还有助于为投资者基于公司分红特征优化投资组合提供建议。

为更有效地对企业分红动机进行研究,本文采用多种先进的机器学习方法展开探索性分析。机器学习方法的优势主要体现在以下两个方面:一方面,已有研究通常采用多元线性回归模型,关注单个特征对分红行为的解释能力,而缺乏对不同类别公司特征的比较性分析。然而,机器学习通过参数正则化、变量筛选和特征重要性分析等技术,能够有效地比较不同特征对公司分红行为预测能力的差异,进而识别出影响股利支付的主要因素。相比之下,多元线性回归在面临多重共线性问题时,难以有效区分各特征对公司分红行为的影响程度(陆瑶和施函青,2022)。另一方面,现有文献基于变量间的线性关系假设,主要关注结果的样本内有效性,忽视了研究结论在样本外的预测能力,可能导致结论的适用性受限。机器学习不依赖于预设的变量间函数关系,而是基于训练集数据拟合变量间的复杂关系。这种灵活性使得机器学习模型在样本外预测性能分析中更具优势,从而更全面地分析公司分红的实际动机。

本文系统梳理了股利分配相关文献,并紧密结合中国资本市场上市公司现金股利支付行为的独特制度背景,主要考察了内外部投资者、公司管理层和治理层等维度特征对公司分红行为预测能力的差异。通过综合理论与实践的考察,能够为本文后续的研究奠定扎实基础。①第一类和第二类代理问题相关特征。由于中国上市公司股权结构相对集中,在第一类代理问题普遍存在的背景下,第二类代理问题也同样突出。考察两类代理问题的相关特征,有助于理解在复杂的股权结构和治理机制下,公司如何平衡各方利益,以及这种平衡如何影响股利支付决策。②生命周期特征。中国证监会等监管部门针对不同发展阶段的企业提出了差异化的现金分红策略,并出台了《上市公司监管指引第3号——上市公司现金分红》(简称《现金分红指引》)等法律法规予以支持。考察公司生命周期特征,既有助于更准确地分析处于不同发展阶段企业的股利支付行为是否存在差异,又有助于分析现有股利监管政策对公司分红行为的影响。③公司税率特征。现有研究表明税收波动性会影响公司的股利支付决策(Amberger, 2023)。由于中国资本市场上市公司会因所属行业以及所处地区的不同而面临不同的税收政策,这些政策会影响公司的税后利润和分红能力。因此,考察公司税率特征,有助于更深入地理解税收政策如何影响公司分红决策。④融资需求特征。在中国资本市场,上市公司融资需求与其股利支付行为密切相关,例如,2008年出台的半强制分红政策将公

司融资行为与股利支付行为挂钩。考察公司的融资需求特征,有助于对比和分析公司在不同融资需求下的股利支付行为。⑤投资者情绪特征。个人投资者是中国资本市场投资者的重要主体,而投资者对股利的非理性需求同样会影响股利支付。探究投资者情绪对公司分红行为的预测能力,能够更全面地了解分红背后的动因。总的来说,本文所探讨的指标可能影响分红行为,以公司层面因素为主,涵盖股东、管理层等公司重要主体,同时还在研究框架中纳入了投资者理性和非理性因素,有助于从一个较为全面的视角对股利支付行为展开研究。

研究表明:①在五种典型的机器学习方法中,渐进梯度回归树和随机森林对于公司分红行为的预测效果最好,且模型性能显著优于多元线性回归等传统实证研究方法。②对比本文所划分的各维度公司特征,生命周期特征和公司税率特征对上市公司分红行为的预测识别效果较好,第一类和第二类代理问题相关特征次之。相比之下,融资需求特征和投资者情绪特征对公司分红行为的预测能力较弱。③在各类分红动因的特征变量中,留存收益资产比、实际税率、纳税波动性、其他应收款资产比和上一期股利支付水平对股利支付水平的预测能力较强。异质性分析显示,影响公司分红行为的主要因素随着2008年半强制分红政策、2013年股息税改革以及《现金分红指引》的颁布发生了改变,且分红特征在不同产权性质以及不同自由现金流状况的企业之间也存在明显差异。此外,是否实施股利迎合策略和投资者的股利情绪也会影响公司的现金分红行为。进一步分析表明,影响公司股利支付意愿(以是否发放现金股利进行度量)与影响公司股利支付水平(以股利支付率进行度量)的主要因素有所不同。④部分特征对于分红行为的预测模式与已有研究发现不同,例如融资约束程度与股利支付水平呈V型关系,而在以往研究中对股利行为解释能力较强的再融资动机和控股股东股权质押比例等指标对股利支付水平的实际预测效果较差。在进行更换响应变量、变更样本划分策略、调整样本期间等检验后,上述结果依然稳健。

本文的边际贡献如下:①在理论层面,现有研究主要基于特定股利理论的视角分析单一公司特征对企业分红行为的解释能力,缺乏对不同维度公司特征的比较性分析,而企业分红的实际动机是亟需探究但尚未解决的重要问题。本文深入剖析了多维度公司特征对中国上市公司分红行为预测能力的相对差异,进而识别出最能揭示分红行为背后逻辑的关键指标。这有助于结合现有股利理论,从更加综合的视角理解上市公司分红决策机制。②在方法论层面,本文创新性地将机器学习方法应用于企业分红行为研究,并对比了不同机器学习方法对预测分红行为这一研究问题的适用性。机器学习出色的拟合能力使其能够精准捕捉变量间的复杂关系,并结合其对变量重要性的深入分析能力以及样本外的泛化能力,能够有效克服多元线性回归在进行高维建模和预测性分析时的局限性。③在实践和政策层面,企业分红是从投资端活跃资本市场的重要手段,本文的研究有助于监管部门更具针对性地督促上市公司开展分红,实现完善上市公司常态化分红机制的目的。这也为监管机构在优化完善《关于加强上市公司监管的意见(试行)》《现金分红指引》等分红相关监管政策时提供了更加直接具体的实证依据,为更好地实现“建设以投资者为本的资本市场”目标提供了政策启示。

二、文献回顾

在探究公司股利支付的影响因素时,现有研究主要基于代理理论、生命周期理论、税收差异理论、信号理论和迎合理论等视角展开分析。这些研究从不同角度解释了公司为何会选择支付股利、支付多少股利,以及这些决策如何受到公司内部和外部因素的影响。

委托代理问题一直是学者们关注的重点之一。在两类代理冲突框架下,学者们对上市公司分红行为的影响因素展开了较为丰富的讨论:对于第一类代理问题,谢军(2006)发现,为避免管理层损害公司价值,第一大股东会表现出支付现金股利的明显动机,许文彬和刘猛(2009)也得到了类似的结论。董艳和李凤(2011)发现,高管持股比例与公司股利支付倾向正相关,马鹏飞和魏志华(2023)发现,数字化转型程度高的企业会降低股利支付水平。在董事特征方面,研究发现独立董事制度能够促使公司更倾向于分配现金股利(冯慧群和马连福,2013)。Ye et al.(2019)发现,董事会中女性比例与公司分红水平正相关。Caliskan and Doukas(2015)则发现,风险厌恶程度高的董事所在的公司更倾向于发放股利。对于第二类代理问题,廖珂等(2018)发现,存在控股股东股权质押的企业具有更低的现金股利水平。与此同时,研究发现机构投资者持股比例与公司分红水平呈正相关关系(姜涛和霍雨佳,2022)。

生命周期特征描绘了企业发展与成长的动态轨迹。随着中国资本市场的不断发展,学者们愈发关注企业在不同成长阶段分红策略的差异。目前,学术界主要采用留存收益资产比作为企业生命周期的代理变量,留存收益资产比越大,说明公司成熟度越高,反之则越低。研究表明,公司股利支付水平与成熟度呈正相关(DeAngelo et al., 2006;宋福铁和梁新颖,2010)。而董理和茅宁(2013)则发现,公司成熟度与现金股利之间的正相关关系会因偿债能力的不同而发生变化:当偿债能力较高时,二者之间的正相关关系消失;而当偿债能力较低时,二者之间的正相关关系会进一步加强。

基于税收差异理论,公司可能会调整其股利政策以最大化股东的税后收益。探讨税收差异理论的研究普遍关注股息税改革对公司分红行为的影响(Li et al., 2017;王国俊和王跃堂,2020),但检验公司税率与股利行为关系的文献较少。Barros et al.(2020)以欧洲证券交易所的上市公司为样本,发现实际税率越高的企业具有更高的股利水平,而避税程度对企业发放股利的意愿和金额没有明显影响。Amberger(2023)则发现,纳税波动性较大的公司具有更低的股利支付意愿和股利支付水平。

关于股利公告能否传递信号,现有研究尚未达成共识。一些学者认为股利公告具有信息含量,如Aggarwal et al.(2012)发现,分析师跟踪人数少的公司具有更强烈的派现动机,并更可能通过增加股息来传递信号。2008年,中国资本市场推出的半强制分红政策,将再融资资格与现金分红水平相挂钩,显著提升了整体资本市场的分红意愿和派现水平(魏志华等,2012;魏志华等,2017;王春飞和郭云南,2021)。此外,He et al.(2016)发现,受到融资约束的企业可能为了在严格的融资环境中增强其外部融资能力而支付股利。而另一些学者则得出了相反的结论,认为信号理论并不能解释上市公司的股利支付行为(De Angelo et al., 1996;吕长江和许静静,2010)。

基于迎合理论的视角,研究显示上市公司股利政策多变、高送转等现象盛行可以归因于企业的现金股利迎合行为(罗琦等,2023)。投资者对股利的非理性需求会影响公司的股利调整(支晓强等,2014;李心丹等,2014)。与此同时,半强制分红政策等股利政策的实施,还促使了企业调整自身股利决策以迎合监管政策(王志强和张玮婷,2012;马鹏飞和董竹,2019)。

综上所述,尽管当前股利理论对分红行为的解释在文献中已颇为详尽,但这些研究通常局限于某一特定理论下的个别特征,忽略了不同理论间的综合对比分析,且常局限于样本内部,结论在全样本中的普适性有待探讨。本文则系统地探究了不同维度公司特征在预测分红行为方面的差异性,这不仅有助于更加深入地理解中国资本市场的分红实践,也有助于识别出驱动公司股利政策的主要指标,为进一步完善分红制度提供参考。与此同时,本文创新性地将机器学习方法运用于预测企业分红水平,这不仅为更加深入地探究企业股利决策提供了便利,也进一步拓宽了机器学习在会计和财务领域的应用范畴。

三、研究设计

1. 研究方法

传统实证研究侧重于探究变量间的因果关系,属于解释性研究;而本文旨在比较不同公司特征对于公司分红行为预测能力的差异,属于预测性研究。鉴于解释性研究和预测性研究存在较大差异,为了更深入地探究公司分红动机,本文采用机器学习方法展开探索性分析。然而,根据“没有免费午餐定理”(Wolpert, 1996),在未经检验的情况下难以得出对于某一研究问题最适宜的研究方法。基于此,本文将全面评估多种主流机器学习方法对公司分红行为的预测效果。具体地,本文将以多元线性回归为基准统计分析方法,采用LASSO、决策树、支持向量机、渐进梯度回归树和随机森林等机器学习方法。其中,LASSO是经典的机器学习方法,通过构造一个惩罚函数压缩一些回归系数,同时将不重要特征的回归系数设定为0,从而对模型进行精简,以降低过拟合风险;支持向量机在许多含有高维数据处理的研究场景中具有优秀的模型性能(李斌等,2019);渐进梯度回归树和随机森林作为集成学习的代表性算法(Bertomeu, 2020; Chen et al., 2022),通过组合多个学习器来获得比单一学习器显著优越的模型性能(Raschka, 2017; 陈强, 2021)。

2. 模型性能评估方法

在评估模型性能时,本文从解释能力和预测精度两个维度综合考量。模型解释能力反映了模型在预测公司分红行为时的拟合程度,本文参考现有研究的做法(Bertomeu et al., 2021; Chen et al., 2022),使用样本内拟合优度 R^2_{in} 考察不同研究方法对训练集样本的拟合效果,样本内拟合优度越高,说明模型训练效果越好。本文采用样本外拟合优度 R^2_{out} 和可解释方差 EVS_{005} 考察各研究方法对于测试集样本的预测效果,其值越大,说明模型对公司分红行为的预测能力越强,样本外拟合效果也越好。模型预测精度则反映了模型在预测分红行为时的准确度,本文选取样本外均方误差 MSE_{005} 衡量模型预测值与实际值之间的偏离程度(Chen et al., 2023)。与此同时,考虑到均方误差可能受到离群值的影响,导致其估计值与实际情况偏离较大,本文还使用平均绝对误差 MAE_{005} 和绝对中位差 $MedAE_{005}$ 评估模型预测公司分红行为时的准确性。^①

四、数据来源和变量定义

1. 数据来源

本文选择2006—2022年A股上市公司作为起始样本,在剔除ST及ST*企业、金融类企业、数据缺失的样本后,最终保留了31469个观测值。选取2006年为样本起点是考虑到本文模型中包含的变量数较多,在此之前的观测数存在较多缺失的情况。为了降低异常值对数据分析的干扰,本文对处于最顶部1%和最底部1%的数据点进行了缩尾处理。

2. 变量定义

参考现有文献的普遍做法(魏志华等,2012;廖珂等,2018;姜涛和霍雨佳,2022),本文选取的响应变量为股利支付率,即公司每股现金股利/每股净利润。在稳健性检验中,本文也采用了股利收益率作为替代响应变量。结合与股利理论相关的实证研究以及中国资本市场分红实践,本文尝试

^① 模型评估指标的具体计算方法参见《中国工业经济》网站(ciejournal.ajcass.com)附件。

探究第一类和第二类代理问题相关特征、生命周期特征、公司税率特征、融资需求特征和投资者情绪特征对公司分红行为预测识别能力的差异。通过归纳整合各类指标,本文能够从一个较为综合的视角分析影响现金分红行为的众多特征,进而着重分析影响公司股利支付行为的关键指标。

对于第一类代理问题,即管理层与股东之间的利益冲突,本文选取相关特征的依据如下:遵从现有文献的做法(董艳和李凤,2011;魏志华等,2012),选取管理费用率($Managefee_ratio_{i,t}$)和管理层持股比例($Manageshare_{i,t}$)纳入模型中,这些指标被认为会影响公司的股利支付政策。此外,研究发现董事特征也会影响公司分红行为(冯慧群和马连福,2013;Caliskan and Doukas, 2015;Ye et al., 2019),因此,本文选取了刻画董事特征的7个变量:独立董事比例($Indep_ratio_{i,t}$)、董事会女性比例($Bgender_{i,t}$)、董事长年龄($Bage_{i,t}$)、任期($Btenure_{i,t}$)、薪酬($Bsalary_{i,t}$)、持股比例($Bshare_{i,t}$)和股权激励虚拟变量($Equity_{i,t}$)。与此同时,研究表明财务报告质量高的企业能够更有效地约束自由现金流问题,从而支付更多的股利(Koo et al., 2017;He et al., 2017),因而本文在模型中也加入了财务报告质量($Da_abs_{i,t}$)。对于第二类代理问题,即大股东与中小股东之间的利益冲突,本文首先选择控股股东股权质押比例($Pledge_{i,t}$)和其他应收款占总资产比重($Tunneling_{i,t}$)加入模型,这些变量对公司现金分红行为存在显著影响(廖珂等,2018;姜涛和霍雨佳,2022)。与此同时,本文还加入了第一大股东持股比例($Top1_{i,t}$)、股权制衡度($Sharebalance_{i,t}$)、中小投资者持股比例($Minorityrate_{i,t}$)和机构投资者持股比例($Institution_{i,t}$)等股权结构特征,尽可能全面地衡量第二类代理问题相关特征对于公司分红行为的影响。

基于生命周期理论的思想,本文沿用一般的做法(DeAngelo et al., 2006;宋福铁和梁新颖,2010;谢德仁,2013),选取留存收益资产比($Retainedearn_ratio_{i,t}$)和自由现金流($Freecash_{i,t}$),考察生命周期特征对公司分红行为的影响。本文还尝试考察公司税率特征对分红行为的影响(Barros et al., 2020;Amberger, 2023),在模型中加入实际税率($Tax_ratio_{i,t}$)、税收规避程度($Tax_avoid_{i,t}$)和纳税波动性($Tax_volatility_{i,t}$),分析企业税收负担以及税收规避倾向对公司分红行为的预测效果。在融资需求方面,本文选取融资约束程度($Constraint_{i,t}$)和再融资动机($Refinance_{i,t}$)两个特征变量,考察融资需求特征对公司分红行为的预测能力。同时,为分析投资者情绪对公司现金分红行为的影响,本文还引入投资者情绪($Sentiment_inv_{i,t}$),以探讨其对公司现金分红行为的作用。

最后,研究发现公司基本面特征也会影响公司股利分配行为(Fama and French, 2001),因此,本文将公司的基本面特征变量作为基准模型的预测变量。本文参考已有文献,在基准模型中控制上一期股利水平($Dividend_lag_{i,t-1}$)、资产收益率($ROA_{i,t}$)、每股经营活动现金流量($Cashflow_{i,t}$)、托宾Q($Tobinq_{i,t}$)、账面市值比($BM_{i,t}$)、资产负债率($Lev_{i,t}$)、产权性质($Soe_{i,t}$)、销售增长率($Growth_{i,t}$)、公司规模($lnsize_{i,t}$)、信息环境质量($Analyst_num_{i,t}$)和公司所在省份市场化程度($Market_idx_{i,t}$)等公司基本面特征。此外,考虑到不同行业间的分红水平存在差异(Li et al., 2017),本文还在模型中控制了行业虚拟变量,变量定义如表1所示。

需要说明的是,在进行因果推断研究时,内生性问题会导致模型估计参数的偏差。然而,本文属于预测性研究,主要关注特征变量与响应变量之间的相关关系而非因果关系。因此,即便变量间存在潜在的内生性问题,也不会对本文评估公司分红行为预测效果产生实质性影响(Raschka, 2017;陈强,2021;陆瑶和施函青,2022)。此外,由于部分股利经典理论在实证方面较难度量,如信号理论、基于投资者个人的税收差别理论,本文未对这些理论进行检验。对于有些可能同时表征多类行为的特征变量,本文则将其归类在基本面特征中。

表1 变量定义

类型	名称	定义
响应变量	$Dividend_ratio_{i,t}$	股利支付率,每股现金股利/每股净利润
第一类代理问题相关特征	$Managefee_ratio_{i,t}$	管理费用率,管理费用与营业收入的比值
	$Manageshare_{i,t}$	管理层持股比例
	$Indep_ratio_{i,t}$	独立董事比例,独立董事人数/董事会总人数
	$Bgender_{i,t}$	董事会女性比例,董事会女性人数/董事会总人数
	$Bshare_{i,t}$	董事长持股比例,董事长持股数/总股本
	$Bage_{i,t}$	董事长年龄加1取对数值
	$Btenure_{i,t}$	董事长任期(月数)加1取对数值
	$Bsalary_{i,t}$	董事长薪酬加1取对数值
	$Equity_{i,t}$	股权激励虚拟变量,若公司当年设置股权激励计划则取1,否则为0
第二类代理问题相关特征	$Da_abs_{i,t}$	财务报告质量,采用DD模型得出的操纵性应计利润的绝对值衡量
	$Tunneling_{i,t}$	其他应收款占总资产比重,衡量大股东对于公司的掏空程度
	$Top1_{i,t}$	第一大股东持股比例,衡量股权集中度
	$Sharebalance_{i,t}$	股权制衡度,第2—5大股东持股比例/第一大股东持股比例
	$Minorityrate_{i,t}$	中小投资者持股比例,除公司董事、监事、高级管理人员以及单独或者合计持有公司5%以上股份的股东以外的其他股东持股比例之和
	$Institution_{i,t}$	机构投资者持股比例
生命周期特征	$Pledge_{i,t}$	控股股东股权质押比例,控股股东股权质押数/控股股东持股数
	$Retainedearn_ratio_{i,t}$	留存收益资产比,(盈余公积+未分配利润)/总资产
公司税率特征	$Freecash_{i,t}$	自由现金流,(经营活动现金流+投资活动现金流-利息支出)/总资产
	$Tax_avoid_{i,t}$	税收规避程度,(税前会计利润-应纳税所得额)/期末总资产
	$Tax_ratio_{i,t}$	实际税率,(所得税费用-递延所得税费用)/税前利润,衡量企业的所得税负担
融资需求特征	$Tax_volatility_{i,t}$	纳税波动性,参考Amberger(2023)的做法,采用公司近 $t-4$ 年到 t 年实际税率的标准差进行衡量,并且为了提高不同公司间纳税波动性的可比性,在计算标准差之前将各年度的实际税率除以这五年的平均实际税率
	$Constraint_{i,t}$	融资约束程度,采用SA融资约束指数衡量
投资者情绪特征	$Refinance_{i,t}$	再融资动机,若上市公司当年(补充)发布了再融资(包括增发和配股)预案,则认为该公司具有再融资动机并取值为1,否则为0
	$Sentiment_inv_{i,t}$	分离企业估值水平中的内在价值部分和市场错误定价部分来得到投资者情绪
公司基本面特征	$Dividend_lag_{i,t-1}$	公司上一年度股利支付率
	$ROA_{i,t}$	资产收益率,净利润/总资产
	$Cashflow_{i,t}$	每股经营活动现金流量,经营活动产生的现金流量净额/实收资本
	$Tobinq_{i,t}$	托宾Q
	$BM_{i,t}$	公司账面市值比
	$Lev_{i,t}$	公司资产负债率,总负债/总资产
	$Soe_{i,t}$	产权性质,若公司属于国有企业取1,否则取0
	$Growth_{i,t}$	公司销售增长率
	$lnsize_{i,t}$	公司规模,公司总资产的自然对数
	$Analyst_num_{i,t}$	信息环境质量,采用分析师跟踪人数的自然对数衡量
	$Market_idx_{i,t}$	公司所在省份市场化程度
$Industry$	公司所在行业虚拟变量,采用2012年版中国证监会行业分类代码进行划分	

五、实证结果和分析

1. 描述性统计

基于本文统计,公司股利支付率($Dividend_ratio_{i,t}$)的均值和中位数分别为0.27和0.21,标准差为0.32,这表明不同公司在股利分配决策上存在明显差异。其余变量的分布情况与现有研究基本一致(魏志华等,2012;董理和茅宁,2013;姜涛和霍雨佳,2022),且数值均在合理范围内。此外,由于不同变量的数量级和分布特征存在差异,为提高变量间的可比性,本文参考李斌等(2019)的做法,在拟合模型前对训练集中所有特征变量进行标准化处理,以使得每个变量均值为0,标准差为1。^①

2. 不同机器学习方法下模型对分红行为的预测效果

本文使用样本内拟合优度 R_{in}^2 、样本外拟合优度 R_{out}^2 以及可解释方差 EVS_{out} 考察模型在训练集中的样本内拟合效果和测试集中的样本外泛化能力。表2第(1)列的结果显示,LASSO的样本内拟合优度与多元线性回归接近,决策树的模型表现远不如多元线性回归,支持向量机、渐进梯度回归树和随机森林的样本内拟合优度则明显高于多元线性回归。其中,随机森林具有最高的 R_{in}^2 ,说明以随机森林为代表的集成学习方法能够获得比线性研究方法更高的样本内拟合效果。对比样本外预测能力,表2第(2)、(3)列的结果表明,随机森林具有最高的样本外拟合优度和可解释方差,渐进梯度回归树的模型表现也较好。相比于多元线性回归,渐进梯度回归树和随机森林的样本外拟合优度分别提高51.44%和66.84%。本文综合使用样本外均方误差 MSE_{out} 、平均绝对误差 MAE_{out} 和绝对中位差 $MedAE_{out}$ 考察不同方法下模型的预测精度。表2第(4)列的结果显示,集成学习方法在进行样本外预测时的均方误差较低。与多元线性回归相比,采用渐进梯度回归树和随机森林这两种方法时,样本外均方误差分别降低9.30%和12.86%,第(5)列的平均绝对误差和第(6)列的绝对中位差也展现了类似的结果。总的来说,集成学习方法能够灵活地采用更合适的函数形式对数据进行拟合,从而构造出预测能力更强、精度更高的企业股利分配预测模型。

表2 模型拟合结果

	R_{in}^2	R_{out}^2	EVS_{out}	MSE_{out}	MAE_{out}	$MedAE_{out}$
	(1)	(2)	(3)	(4)	(5)	(6)
多元线性回归	0.2490	0.1565	0.1812	0.0871	0.1834	0.1254
LASSO	0.2386	0.1805	0.1928	0.0850	0.1794	0.1229
决策树	0.0474	0.0106	0.0220	0.1016	0.2137	0.1823
支持向量机	0.5964	0.1613	0.1675	0.0868	0.1788	0.1177
渐进梯度回归树	0.6075	0.2370	0.2520	0.0790	0.1646	0.1009
随机森林	0.9025	0.2611	0.2782	0.0759	0.1605	0.0977

3. 不同类别公司特征对公司现金分红行为预测能力的表现差异

为探究不同类别公司特征对公司分红行为预测能力的表现差异,本文以公司基本面特征作为基准模型,比较在基准模型中加入不同类别公司特征时的模型拟合效果。同时,本文列示了包含所有类

^① 描述性统计结果参见《中国工业经济》网站(ciejournal.ajcass.com)附件。

别公司特征时的模型(简称综合模型)结果作为对照。可以看到,基准模型中包含了反映公司盈利能力和增长潜力的指标,这些公司基本面特征对预测和解释公司分红行为起到了重要作用。在此基础上,对比在基准模型中加入某一类别公司特征时的模型拟合性能,本文发现生命周期理论特征的加入显著提升了基准模型的表现,公司税率特征的表现次之,继而是第一类和第二类代理问题相关特征,而融资需求特征和投资者情绪特征带来的预测性能提升幅度相对较小,如表3所示。这表明,以往研究中关注较少的生命周期特征和公司税率特征对公司现金分红行为的影响程度反而较大,而两类代理问题相关特征、融资需求特征和投资者情绪特征对股利支付行为的预测解释能力相对较弱。

表3 不同类别公司特征对公司现金分红行为的预测表现差异(R_{008}^2)

公司特征组合	多元线性回归	LASSO	决策树	支持向量机	渐进梯度回归树	随机森林
	(1)	(2)	(3)	(4)	(6)	(7)
基准模型	0.1654	0.1641	0.0030	0.1720	0.2049	0.2184
基准+第一类代理	0.1647	0.1691	0.0093	0.1578	0.2146	0.2304
基准+第二类代理	0.1661	0.1679	0.0190	0.1648	0.2180	0.2331
基准+生命周期	0.1678	0.1681	0.0066	0.1730	0.2320	0.2472
基准+公司税率	0.1691	0.1642	0.0049	0.1677	0.2242	0.2407
基准+融资需求	0.1651	0.1688	0.0352	0.1723	0.2143	0.2307
基准+投资者情绪	0.1404	0.1624	0.0147	0.1722	0.2141	0.2280
综合模型	0.1565	0.1805	0.0106	0.1613	0.2370	0.2611

4. 不同特征变量对公司分红行为预测能力的差异

鉴于渐进梯度回归树和随机森林在预测公司分红行为方面展现出了显著的优势,本文计划基于这两种集成学习算法深入探索不同特征变量对公司分红行为预测效果的影响差异。然而,由于集成学习模型的复杂性,其预测结果通常不能像单一学习器那样被直观解释。为了揭示集成学习模型背后的经济学意义,本文将采用特征相对重要性的方法,量化不同特征在预测模型中的重要性程度(周志华,2016)。表4的结果显示,相对重要性最高的特征变量分别为上一期股利支付水平($Dividend_lag_{i,t-1}$)、资产收益率($ROA_{i,t}$)、留存收益资产($Retainedearn_ratio_{i,t}$)、实际税率($Tax_ratio_{i,t}$)、

表4 变量相对重要性 单位:%

排序	随机森林		渐进梯度回归树	
	变量	相对重要性	变量	相对重要性
1	$Dividend_lag_{i,t-1}$	17.8244	$Dividend_lag_{i,t-1}$	30.4594
2	$ROA_{i,t}$	7.9011	$ROA_{i,t}$	12.7493
3	$Retainedearn_ratio_{i,t}$	6.3092	$Retainedearn_ratio_{i,t}$	6.6660
4	$Tax_ratio_{i,t}$	4.4097	$Tax_ratio_{i,t}$	4.9492
5	$Tax_avoid_{i,t}$	3.9103	$Lev_{i,t}$	2.7796
6	$Tax_volatility_{i,t}$	3.2962	$Growth_{i,t}$	2.6070
7	$Growth_{i,t}$	2.9151	$Tax_volatility_{i,t}$	2.1892
8	$Lev_{i,t}$	2.8889	$Da_abs_{i,t}$	2.0731
9	$Tunneling_{i,t}$	2.6110	$Institution_{i,t}$	2.0386
10	$Freecash_{i,t}$	2.5321	$Tunneling_{i,t}$	1.9986

纳税波动性($Tax_volatility_{i,t}$)、资产负债率($Lev_{i,t}$)、销售增长率($Growth_{i,t}$)和其他应收款资产比($Tunneling_{i,t}$)。与之相对,结果表明,尽管已有研究发现再融资动机($Refinance_{i,t}$)和控股股东股权质押比例($Pledge_{i,t}$)等指标对公司股利分配具有一定的解释力,但这些指标的相对重要性较低,说明其对公司股利分配行为的实际预测能力不佳。^①最后,考虑到不同类别特征数量的差异可能影响本文的研究结论,本文选取各自类别中相对重要性最大的一个特征加入基本面特征中,重新训练模型并计算各类别特征的相对重要性。结果表明,尽管随着时间推移,不同类别的公司特征累计相对重要性有所变动,但整体而言,生命周期特征和公司税率特征的相对重要性仍显著高于其他类别的公司特征,进一步验证了表3的研究结论。^②

5.重要特征对公司分红行为的预测模式

进一步地,本文采用累计局部效应图(Accumulated Local Effects Plot, ALE图)研究单个变量对股利支付水平的预测模式(Apley and Zhu, 2020)。借助这一工具,本文将重点分析相对重要性排序靠前的公司特征变量,以及那些与已有研究存在明显差异的指标,深入探究其如何影响公司股利支付水平的变动。

(1)留存收益资产比。图1是留存收益资产比的ALE图。可以看到,随着留存收益资产比的增加,公司股利支付水平呈现出先缓慢增加,然后迅速增加,再恢复缓慢增加的趋势。这表明,公司在成立早期倾向于支付较少的股利,可能是为了留存更多的资金以供自身发展;随着公司逐渐步入成熟期,其发放的股利水平也越来越高;当公司处于成熟期乃至衰退期时,其会选择较为稳健的股利政策并发放较高的股利。这一模式与《现金分红指引》中对于企业不同发展阶段提出的差异化分红政策思路一致,体现了《现金分红指引》相关规定的合理性。

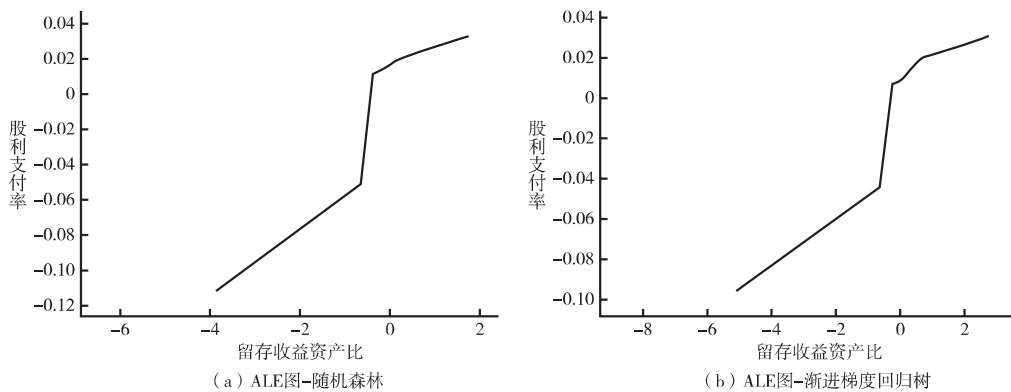


图1 留存收益资产比的累计局部效应

(2)实际税率和纳税波动性。图2是实际税率的ALE图。从中可以看到,中国上市公司的实际税率与股利支付水平之间呈现V型关系,意味着实际税负较低或较高的企业更倾向于支付股利。对于税负较轻的企业而言,这种低税负状态不仅意味着企业能够拥有更多的自由现金流来支持股利的支付,而且还可能有助于降低企业的权益资本成本。这是因为,较低的税负往往能够减少投资者对企业风险的感知,进而提升企业的市场价值。这种积极的市场反应使得企业更加有能力并且愿意支

^① 详细变量重要性排序结果参见《中国工业经济》网站(ciejournal.ajcass.com)附件。

^② 具体结果参见《中国工业经济》网站(ciejournal.ajcass.com)附件。

付股利,为股东创造更大的价值;对于税负较高的企业,尽管高税负会压缩企业的税后利润和现金流,但此时公司也可能通过支付股利尽可能维持与投资者的关系。图3则是纳税波动性的ALE图,可以看到,随着公司纳税波动性增加,公司的股利支付水平会迅速降低,然后呈现出缓慢增加的趋势。这意味着当公司的税后现金流不稳定时,公司会更可能维持一个较低的股利支付水平,从而避免后续因股利支付下降而导致负面市场反应的可能。随着纳税波动性持续处于一个较高的水平,公司可能会逐渐适应这一波动,进而逐渐调整其财务政策,开始提高股利支付水平。这种调整可能是基于对公司未来现金流更加准确的预测,以及对公司财务状况更加深入的分析。

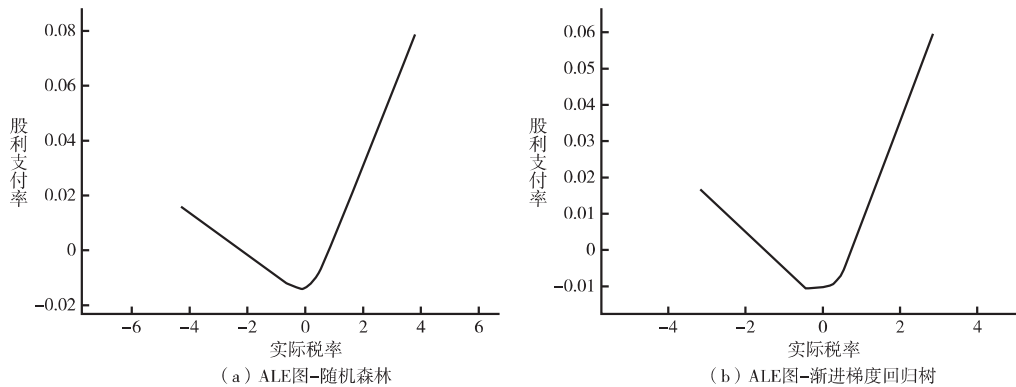


图2 实际税率的累计局部效应

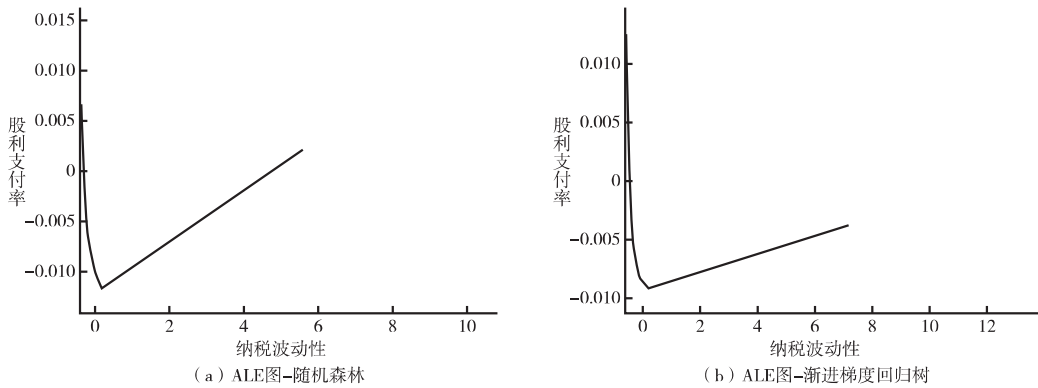


图3 纳税波动性的累计局部效应

(3)其他应收款资产比。其他应收款资产比反映了第二类代理成本。图4的结果表明其他应收款资产比越高,股利支付水平越低,这表明大股东可能通过资金占用等方式侵害中小股东利益,同时也导致了低派现股利政策。

(4)融资约束程度。本文采用SA融资约束指数衡量公司的融资约束程度。由于SA指数为负数,因而其值越小时,融资约束程度越大。图5的结果显示^①,融资约束程度与公司股利支付水平呈现出V型关系:一方面,在公司面临严重融资约束的情况下,公司仍然会维持较高的分红水平,可能

^① 横轴为经标准化处理后的SA指数,因而存在大于0的观测值。

是由于此类企业急需融资,而半强制分红政策规定了再融资所需的最低分红比例,这一发现支持了 He et al.(2016)的结论;另一方面,当公司的融资约束程度相对较低时,同样会倾向于发放更多的股利,这可能是由于此时企业的现金流较为充裕,公司有足够的资金发放股利。

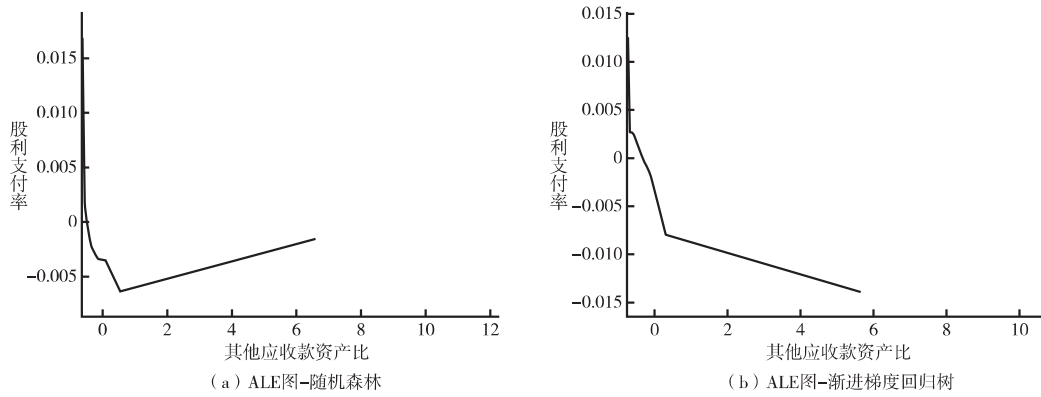


图4 其他应收款资产比的累计局部效应

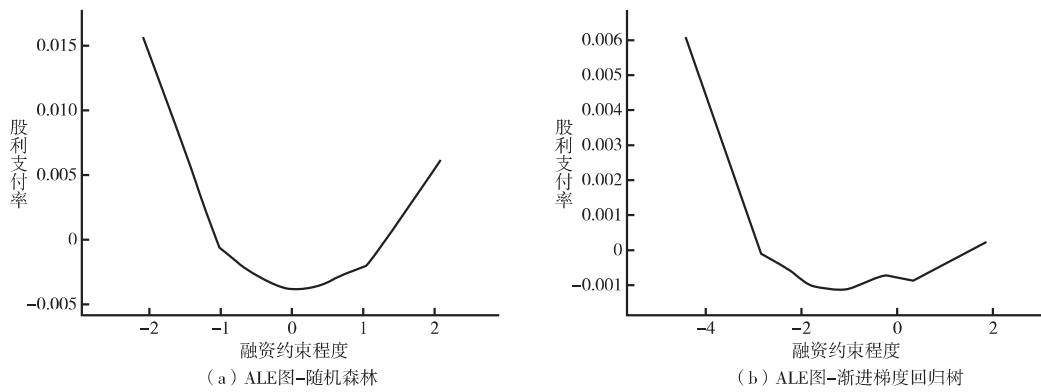


图5 融资约束程度的累计局部效应

(5)上一期股利支付水平。由图6可以看到,上一期股利支付水平越高,公司当期股利支付水平也越高,这与理论预期一致,说明公司股利支付行为具有一定的持续性和稳定性。

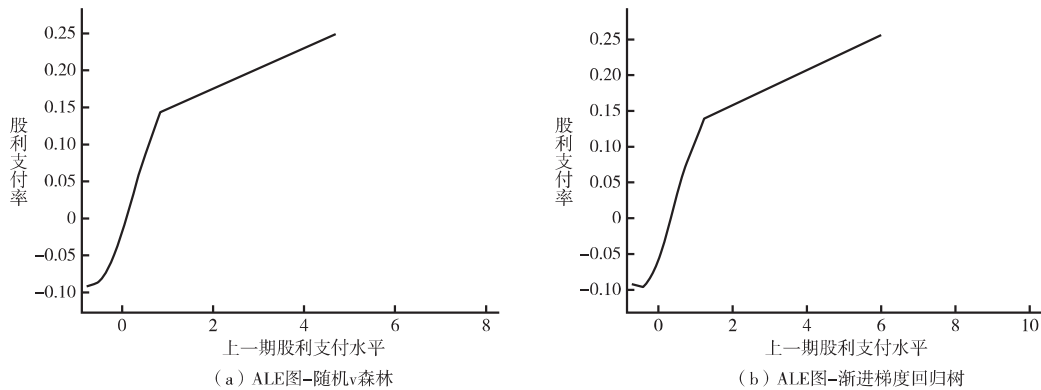


图6 上一期股利支付水平的累计局部效应

(6)其余特征变量。结果显示,除虚拟特征变量以及部分连续变量外,其余特征与公司股利支付水平之间的关系呈现出非线性的特点,这表明线性假定不能完全解释特征变量与公司股利水平之间的相关关系,也在一定程度上解释了为何多元线性回归和LASSO等线性研究方法在公司分红行为的预测识别效果方面不如随机森林等非线性研究方法。^①

6. 进一步分析^②

(1)股利支付行为在分红政策变更前后的差异。为便于探究分红政策变更前后公司分红行为的差异,本文将样本划分为以下三个子样本,分别为2007年、2008—2011年、2013年及之后。这一划分的原因如下:一方面,2008年10月出台的《关于修改上市公司现金分红若干规定的决定》意味着半强制分红政策的正式实施。探究这一时间节点前后公司分红行为的变化,有助于分析半强制分红政策对于企业现金股利支付行为的影响。由于2007年中国会计准则的改变可能导致2006年的数据结果与2007年及以后的数据结果不完全可比,因而本文选择以2007年样本作为半强制分红政策实施前的考察期间。另一方面,对于上市公司而言,2013年11月颁布的《现金分红指引》提出差异化分红政策,旨在通过进一步规范上市公司现金分红,从而维护投资者合法权益。对于投资者而言,2013年1月实施的股息红利税差异化政策则鼓励投资者进行长期投资(潘越等,2013)。《现金分红指引》以及股息税改革均可能影响公司分红行为。考虑到这两项政策出台时间接近但不完全一致,本文剔除了政策变更当年的观测样本并共同检验两项政策带来的效应。^③

本文计算并对比三组子样本中的变量相对重要性排序,结果显示,一方面,与主检验结果相似的是,无论是在股利政策变更前或变更后,上一期股利支付水平($Dividend_lag_{i,t-1}$)、留存收益资产比($Retainedearn_ratio_{i,t}$)、资产收益率($ROA_{i,t}$)、资产负债率($Lev_{i,t}$)和实际税率($Tax_ratio_{i,t}$)均是对上市公司股利支付率预测效果较好的变量。另一方面,异质性分析获得了更进一步的发现:对比2007年与2008—2011年这两个阶段,自由现金流($Freecash_{i,t}$)对于2007年样本期间的相对重要性较高。结果还表明,随着自由现金流增加,公司的股利支付率呈现出先下降再上升的趋势。本文认为这可能是由于当自由现金流处于较低水平时,公司更可能选择将增加的资金用于再投资。因此,即使自由现金流增加,公司也可能选择降低股利支付率,以满足再投资的需求。当自由现金流处于较高水平时,公司有足够的闲置资金进行分红,并且通过提高股利支付率,公司可以向股东传递出积极的信号,增强股东对公司的信任,进而促进公司股价的稳定和增长。相比之下,机构投资者持股比例($Institution_{i,t}$)则在2008—2011年的相对重要性排序较高,并且随着机构投资者持股比例增加,股利支付水平会先降后增。本文认为出现这一现象的原因可能是,当机构投资者持股比例较低时,公司股利支付决策主要受大股东影响。根据利益输送假说,如果大股东持股比例较高的公司支付较高的股利,可能意味着大股东存在较为严重的利益侵占现象。中国资本市场的机构投资者能够发挥监督并改善公司治理的作用,因而此时机构持股比例与股利支付水平负相关;当机构投资者持股比例处于较高水平时,通常意味着机构投资者在公司分红行为上具有显著的影响力。此时,机构投资者占比越高的公司更可能进行分红,以迎合机构投资者获取股息的需求。

对比2008—2011年与2013年之后这两个阶段,结果显示自由现金流和机构投资者持股比例的相对重要性排序发生了变化。在2013年之后这一阶段,自由现金流的相对重要性提高,而机构投

① 具体结果参见《中国工业经济》网站(ciejournal.ajcass.com)附件。

② 进一步分析结果参见《中国工业经济》网站(ciejournal.ajcass.com)附件。

③ 由于公司通常在下一年度发放当前年度的股利,2013年初股息税改革也会影响2012年的股利发放情况,因此本文将2012年、2013年样本进行剔除。

资者持股比例的相对重要性则有所降低。与此同时,本文还发现,在2012年之前,融资约束程度越低,股利支付水平越高,这与在全样本时期范围内融资约束程度与股利支付水平呈现的V型关系有所不同。对此,本文认为一个可能的原因是,在2012年之前中国上市公司对分红的重视程度尚不足,企业缺乏足够的动机进行股利分配。随着中国关于企业分红规定的逐渐完善,中国证监会对分红政策的具体落实提出了更为详细的要求,上市公司的分红意愿和水平普遍提高,使得即使面临较高的融资约束程度,公司也可能进行更多的分红以获得再融资机会。整体而言,随着中国资本市场公司分红制度体系的逐渐完善,对企业现金分红的要求逐渐细化,影响上市公司进行股利支付的主要特征也发生了一定程度的变化。

此外,在半强制分红政策背景下,有再融资需求的上市公司会调整其股利支付策略,以确保符合监管部门规定的分红条件(王志强和张玮婷,2012;马鹏飞和董竹,2019)。本文参考马鹏飞和董竹(2019)的做法,按是否达到半强制分红政策的最低要求对样本进行分组,考察实施股利迎合策略和未实施股利迎合策略时公司分红行为的差异。结果表明,管理费用率的相对重要性在采取股利迎合策略的企业中的排序更高。相比于采取股利迎合策略的企业,杠杆率和自由现金流在未实施此策略的企业中的相对重要性排序更高。并且,杠杆率越高,股利支付率越低。这意味着,在没有实施股利迎合策略的企业,其股利支付决策在更大程度上受到公司财务结构的影响。高杠杆率的企业通常面临更高的财务风险,为了保持财务稳定,企业可能采取保守的股利政策,即降低股利支付率,以减少对外部资金的依赖,并保留更多的现金流以应对潜在的风险。

(2)公司分红行为在不同现金流状况企业间的差异。考虑到以《现金分红指引》为代表的股利政策基于公司特点提出了差异化的分红政策,为进一步分析不同类型企业在近年来一系列分红政策出台后,其分红行为是否发生变化,本文基于自由现金流的行业年度中位数将企业分为现金流状况好与现金流状况差的两组子样本,进而分析这两类企业在公司分红行为方面是否存在差异。结果显示,自由现金流这一指标本身在现金流状况好的企业中具有较高的相对重要性,而其他应收款资产比在现金流状况较差的企业中的相对重要性较高,并且此时自由现金流和其他应收款资产比的累计局部效应与前文结果基本一致。

(3)公司分红行为在不同股权性质企业间的差异。本文进一步探究了公司分红行为在国有企业与非国有企业间的差异。结果表明,相比于国有企业,非国有企业中其他应收款资产比对股利支付水平的预测效果更为显著。该变量的预测模式分析显示,其他应收款资产比越高时,股利支付水平越低。这说明在非国有企业中,大股东会更可能通过资金占用等方式侵害中小股东利益,同时也导致了低派现股利政策。对比国有企业与非国有企业中变量排序的差异,结果显示资产负债率在非国有企业中的排序比国有企业低。这可能是由于国有企业还承担着一些社会责任和公共项目,这些项目通常需要大量的资金投入,进而资产负债率较高,因而资产负债率对国有企业股利支付水平的影响权重更大。此外,销售增长率在非国有企业中的排序相对更高。一种可能的解释是,非国有企业在市场竞争中通常更加灵活和敏锐,而国有企业由于企业性质以及经营目的的不同,其分红政策更为刚性(王国俊和王跃堂,2020),因此,销售增长率对分红行为的影响程度可能较小。公司特征对不同股权性质企业的分红行为预测所产生的影响具有显著差异,这一发现为深入研究不同股权性质企业的分红行为提供了探索性依据。

(4)投资者股利情绪对分红行为的影响。本文进一步考察宏观层面的投资者现金股利情绪对企业现金分红行为的影响。参考姜富伟等(2021)、罗琦等(2023)的做法,本文度量了中国资本市场投资者现金股利情绪,然后将样本按年度现金股利情绪指数的中位数进行分组,进而分析现金股利

情绪高的年份与现金股利情绪低的年份下公司分红行为是否存在明显差异。结果显示,相比于现金股利情绪低的年份,其他应收款资产比在现金股利情绪高的年份具有更高的相对重要性,意味着大股东对于公司的资金占用行为在股利情绪高的年份更能影响企业股利支付水平。

(5)影响公司股利支付意愿的主要特征。针对不分红公司,中国证监会已明确表示将通过强化披露要求来督促其分红。基于此,本文以公司是否发放股利作为响应变量,进一步探究影响公司股利支付意愿的关键因素。这有助于针对性地引导更多未分红企业制定更合理的分红政策,进而提高整体分红水平。相较主检验,本文发现经营活动现金流($Cashflow_{i,t}$)和分析师跟踪人数($Analyst_num_{i,t}$)对公司发放股利意愿的预测能力较强。ALE图的结果表明,分析师跟踪人数更多、经营活动净现金流越高的企业会更倾向于分红。对此,一个可能的解释是,分析师的关注有助于提高公司信息透明度,从而推动提高企业进行分红的意愿;而经营活动现金流量高的企业往往具有更强的盈利能力和资金实力,也更有能力进行分红。

7. 稳健性检验

本文采用变更响应变量、更换样本划分方法、验证累计局部效应图的可靠性和更换样本期间等方式进行稳健性检验,考察结果是否可靠。^①

(1)变更响应变量。在主检验中,本文采用股利支付率作为响应变量。为进一步提高研究结论的普适性,本文还采用股利收益率作为替代性响应变量。股利收益率是指每股现金股利与每股总资产的比值。结果表明,替换响应变量后的研究结论与主检验一致。

(2)变更样本划分方法。在主检验中,本文采用一年训练窗口期、一年测试窗口期的方式进行滚动预测。为避免因样本划分方法而导致的结果偏误,本文还使用以下样本划分方法重新拟合模型:①按7:3的比例随机划分训练集和测试集;②按5:5的比例随机划分训练集和测试集。变更样本划分方法后,以渐进梯度回归树和随机森林为代表的集成学习方法对股利支付行为的预测识别能力仍显著高于多元线性回归,本文的研究结论保持不变。

(3)关于累计局部效应图的稳健性检验。除了利用累计局部效应图外,集成学习还常使用部分依赖图分析特征变量对于响应变量的影响模式。该方法的核心思想是,在其他特征保持恒定的条件下,观察某一特定变量的变化如何影响企业的股利支付水平,并将这种关系以直观的图像形式展示出来。由于部分依赖图需要假设所考察的特征与其他特征不相关,而这一条件在现实中常常不符合,因而本文在主检验中采用累计局部效应图进行详细分析。为了进一步增强累计局部效应图的稳健性,本文将其与部分依赖图的图像模式进行了对比。结果显示,部分依赖图与累计局部效应图的图像模式基本一致,进一步增强了结果的稳健性。^②

(4)变更样本期间。鉴于2007年会计准则发生了变化,为避免因准则变化导致结果发生改变,本文还将样本起点调整至2007年。变更样本期间后的结果依旧稳健。

六、结论与启示

企业分红行为是“建设以投资者为本”的特色现代资本市场的研究重点,其中,股利支付的实际动机一直是备受各界关注但尚未达成共识的重要基础性问题。本文以多元线性回归为基准,

① 稳健性检验结果参见《中国工业经济》网站(ciejournal.ajcass.com)附件。

② 考虑到部分依赖图受变量相关性影响,当部分依赖图与累计局部效应图的图像存在细微差别时,仍以累计局部效应图的图像为准。具体结果参见《中国工业经济》网站(ciejournal.ajcass.com)附件。

采用LASSO、决策树、支持向量机、渐进梯度回归树和随机森林五种机器学习方法探究多元公司特征对上市公司分红行为的预测能力。这既能为监管部门提供改进上市公司现金分红制度的参考依据,又有助于投资者基于公司分红行为优化投资组合,提高投资效率。研究发现:在众多机器学习方法中,渐进梯度回归树和随机森林对股利支付行为的预测效果最佳,且模型性能显著优于多元线性回归;对比不同类别特征,结果显示生命周期特征和公司税率特征对公司分红行为的预测能力最强,第一类和第二类代理问题相关特征也展现出了较好的预测效果,而融资需求特征和投资者情绪特征的预测效果相对较差。在不同公司特征变量中,上一期股利支付水平、留存收益资产比、实际税率、纳税波动性和其他应收款资产比对股利支付水平的预测能力较强。异质性分析显示,公司股利支付行为的影响因素在半强制分红政策、2013年股息税改革以及《现金分红指引》出台前后发生了变化,在不同股权性质和现金流状况企业之间也存在差异。与此同时,是否采取股利迎合策略和投资者的现金股利情绪也是影响公司现金分红行为的重要因素。进一步分析表明,影响公司股利支付意愿和股利支付水平的主要因素有所区别。在变更响应变量、改变样本划分方法、调整样本期间等稳健性检验后,研究结论保持不变。

本文的研究结论具有较强的研究意义和政策启示。在研究意义方面,当前研究大多局限于特定的股利理论框架内,对单一公司特征在分红行为上的解释作用进行探讨,缺乏对多元公司特征的横向对比分析。企业的分红实际驱动力是一个紧迫但尚待解决的关键议题。本文立足于现有的与股利理论或公司分红相关的实证研究以及中国资本市场制度背景,汇总了影响公司股利支付的多方面特征,分析了不同类别的公司特征对股利支付行为预测识别能力的差异,并进一步着重细致刻画了那些对股利支付预测具有重要影响的公司特征及其与股利支付之间的复杂关系。这一研究不仅为理解企业分红动机提供了更为宽广的理论视野,也进一步揭示了资本市场分红实践中所蕴含的复杂机制。在政策启示方面:对监管部门而言,本文的研究为促进企业通过分红方式向资本市场注入活力提供了实证支持。①监管部门可以加强关注影响公司股利支付行为的重要特征,从而制定更为精准的股利监管政策。例如,本文发现其他应收款资产比越高的企业具有越低的分红水平。考虑到这一指标反映了公司资金被占用的程度,因此,监管机构应对此类行为保持高度警惕,持续严厉打击资金占用行为,以缓解代理问题,从而为企业创造更健康、公平的分红环境。2023年10月,为进一步健全上市公司常态化分红机制,中国证监会对《现金分红指引》进行了重新修订,本次修订旨在鼓励公司制定明确的分红政策,并通过强化披露要求引导合理分红。2024年3月,中国证监会进一步出台《关于加强上市公司监管的意见(试行)》,提出对分红采取强约束措施,要求上市公司制定积极、稳定的现金分红政策,明确投资者预期。基于本文的研究结论,监管部门在后续修订《现金分红指引》等相关法制规章时,可以进一步细化对不同类型的企业的分红要求。例如,可以针对不同其他应收款资产比水平的企业设定不同的分红标准,以促进合理分红,防止企业过度占用资金,损害投资者利益。②监管部门还可以进一步加强对处于不同发展阶段企业的分级监管,本文的结果表明,生命周期特征能够较好地对中国资本市场分红实践进行预测,在现有差异化现金分红政策的基础上,可以尝试对上市公司进行更细致的划分,这不仅有助于推动上市公司分红行为的规范化、透明化,还有助于提升整个资本市场的稳定性和健康性。总的来说,本文的研究为上市公司分红监管政策改革提供了实证证据,为后续进一步健全常态化现金分红机制奠定了基础。对于投资者而言,现金分红是投资者获取投资收益的重要途径。投资者同样可以加强对那些能够显著影响企业现金分红水平指标的关注,精准识别出符合自身分红预期的企业,从而做到对持有资金的合理规划,更早地分享企业成长红利。此外,投资者还应积极了解代理理论、生命周期理论等股利理论,

这些理论能帮助他们更深入地了解资本市场分红实践,从而做出更明智的投资决策。对于上市公司而言,深入理解自身的发展情况是至关重要的,其可以更好地促进自身提高分红水平,为投资者带来稳定的回报。更持续稳定的分红不仅能增强投资者的回报预期,还能推动市场形成更为平和、理性的投资氛围,从而进一步促进资本市场的平稳健康发展。

本文认为,后续研究可以利用机器学习等方法实现大数据分析,尝试对信号理论、剩余股利理论等难以在实证上进行度量的股利理论进行检验,这有助于更好地分析这些理论在中国资本市场分红实践中的重要性。需要说明的是,尽管本文立足于现有的与股利理论或公司分红相关的实证研究以及中国资本市场制度背景,对影响公司股利支付的公司特征进行了分类,这有助于对研究结果进行相对有序清晰解读,但指标分类方式并不唯一,后续研究也可基于其他依据归纳整合。

〔参考文献〕

- [1]陈强.机器学习及Python应用[M].北京:高等教育出版社,2021.
- [2]董理,茅宁.公司成熟度、剩余负债能力与现金股利政策——基于财务柔性视角的实证研究[J].财经研究,2013,(11):59-68.
- [3]董艳,李凤.管理层持股、股利政策与代理问题[J].经济学(季刊),2011,(3):1015-1038.
- [4]冯慧群,马连福.董事会特征、CEO权力与现金股利政策——基于中国上市公司的实证研究[J].管理评论,2013,(11):123-132.
- [5]姜富伟,马甜,张宏伟.高风险低收益?基于机器学习的动态CAPM模型解释[J].管理科学学报,2021,(1):109-126.
- [6]姜涛,霍雨佳.分红动因识别、机构持股与信号传递[J].南开管理评论,2022,(4):142-156.
- [7]李斌,邵新月,李玥阳.机器学习驱动的基本面量化投资研究[J].中国工业经济,2019,(8):61-79.
- [8]李心丹,俞红海,陆蓉,徐龙炳.中国股票市场“高送转”现象研究[J].管理世界,2014,(11):133-145.
- [9]廖珂,崔宸瑜,谢德仁.控股股东股权质押与上市公司股利政策选择[J].金融研究,2018,(4):172-189.
- [10]陆瑶,施函青.我国科技企业融资的决定因素研究——基于科创板企业的机器学习分析[J].金融研究,2022,(9):132-151.
- [11]罗琦,张志达,吴希梅,喻天琦.股利情绪、股利迎合与股价崩盘风险——基于百度指数平台搜索量的经验证据[J].管理科学学报,2023,(2):87-103.
- [12]吕长江,许静静.基于股利变更公告的股利信号效应研究[J].南开管理评论,2010,(2):90-96.
- [13]马鹏飞,董竹.股利折价之谜——基于大股东掏空与监管迎合的探索[J].南开管理评论,2019,(3):159-172.
- [14]马鹏飞,魏志华.企业数字化转型如何影响现金股利政策:“信息面”还是“资金面”[J].南开管理评论,2023,(9):1-27.
- [15]潘越,王宇光,戴亦一.税收征管、政企关系与上市公司债务融资[J].中国工业经济,2013,(8):109-121.
- [16]宋福铁,梁新颖.企业生命周期理论与上市公司现金股利分配实证研究[J].财经研究,2010,(9):123-133.
- [17]王春飞,郭云南.半强制股利政策与股权融资成本[J].金融研究,2021,(8):172-189.
- [18]王国俊,王跃堂.红利税改革对现金分红宣告效应的影响研究[J].中国工业经济,2020,(11):137-155.
- [19]王志强,张玮婷.上市公司财务灵活性、再融资期权与股利迎合策略研究[J].管理世界,2012,(7):151-163.
- [20]魏志华,李常青,吴育辉,黄佳佳.半强制分红政策、再融资动机与经典股利理论——基于股利代理理论与信号理论视角的实证研究[J].会计研究,2017,(7):55-61.
- [21]魏志华,吴育辉,李常青.家族控制、双重委托代理冲突与现金股利政策——基于中国上市公司的实证研究[J].金融研究,2012,(7):168-181.
- [22]谢德仁.企业分红能力之理论研究[J].会计研究,2013,(2):22-32.
- [23]谢军.股利政策、第一大股东和公司成长性:自由现金流理论还是掏空理论[J].会计研究,2006,(4):51-57.

- [24]许文彬,刘猛.我国上市公司股权结构对现金股利政策的影响——基于股权分置改革前后的实证研究[J].中国工业经济,2009,(12):128-138.
- [25]支晓强,胡聪慧,童盼,马俊杰.股权分置改革与上市公司股利政策——基于迎合理论的证据[J].管理世界,2014,(3):139-147.
- [26]周志华.机器学习[M].北京:清华大学出版社,2016.
- [27]Aggarwal, R., J. Cao, and F. Chen. Information Environment, Dividend Changes, and Signaling: Evidence from ADR Firms[J]. Contemporary Accounting Research, 2012, 29(2):403-431.
- [28]Amberger, H. J. Volatility of Tax Payments and Dividend Payouts[J]. Contemporary Accounting Research, 2023, 40(1):451-487.
- [29]Apley, D. W., and J. Zhu. Visualizing the Effects of Predictor Variables in Black Box Supervised Learning Models[J]. Journal of the Royal Statistical Society Series B: Statistical Methodology, 2020, 82(4):1059-1086.
- [30]Barros, V., P. V. Matos, and J. M. Sarmiento. What Firm's Characteristics Drive the Dividend Policy? A Mixed-Method Study on the Euronext Stock Exchange[J]. Journal of Business Research, 2020, 115:365-377.
- [31]Bertomeu, J. Machine Learning Improves Accounting: Discussion, Implementation and Research Opportunities[J]. Review of Accounting Studies, 2020, 25:1135-1155.
- [32]Bertomeu, J., E. Cheynel, E. Floyd, and W. Pan. Using Machine Learning to Detect Misstatements[J]. Review of Accounting Studies, 2021, 26:468-519.
- [33]Caliskan, D., and J. A. Doukas. CEO Risk Preferences and Dividend Policy Decisions [J]. Journal of Corporate Finance, 2015, 35:18-42.
- [34]Chen, C., B. Ke, and Q. Zhao. Measuring Firm Quality Using Machine Learning[R]. SSRN Working Paper, 2023.
- [35]Chen, X., Y. H. Cho, Y. Dou, and B. Lev. Predicting Future Earnings Changes Using Machine Learning and Detailed Financial Data[J]. Journal of Accounting Research, 2022, 60(2):467-515.
- [36]DeAngelo, H., L. DeAngelo, and D. J. Skinner. Reversal of Fortune: Dividend Signaling and the Disappearance of Sustained Earnings Growth[J]. Journal of Financial Economics, 1996, 40(3):341-371.
- [37]DeAngelo, H., L. DeAngelo, and R. M. Stulz. Dividend Policy and the Earned/Contributed Capital Mix: A Test of the Life-Cycle Theory[J]. Journal of Financial Economics, 2006, 81(2):227-254.
- [38]Fama, E. F., and K. R. French. Disappearing Dividends: Changing Firm Characteristics or Lower Propensity to Pay[J]. Journal of Financial Economics, 2001, 60(1):3-43.
- [39]He, W., L. Ng, N. Zaiats, and B. Zhang. Dividend Policy and Earnings Management across Countries[J]. Journal of Corporate Finance, 2017, 42:267-286.
- [40]He, Z., X. Chen, W. Huang, R. Pan, and J. Shi. External Finance and Dividend Policy: A Twist by Financial Constraints[J]. Accounting & Finance, 2016, 56(4):935-959.
- [41]Koo, D. S., S. Ramalingegowda, and Y. Yu. The Effect of Financial Reporting Quality on Corporate Dividend Policy[J]. Review of Accounting Studies, 2017, 22:753-790.
- [42]Li, O. Z., H. Liu, C. Ni, and K. Ye. Individual Investors' Dividend Taxes and Corporate Payout Policies[J]. Journal of Financial and Quantitative Analysis, 2017, 52(3):963-990.
- [43]Raschka, S. Python Machine Learning[M]. Birmingham: Packt Publishing - ebooks Account, 2017.
- [44]Wolpert, D. H. The Lack of A Priori Distinctions between Learning Algorithms[J]. Neural Computation, 1996, 8(7):1341-1390.
- [45]Ye, D., J. Deng, Y. Liu, S. H. Szewczyk, and X. Chen. Does Board Gender Diversity Increase Dividend Payouts? Analysis of Global Evidence[J]. Journal of Corporate Finance, 2019, 58:1-26.

Motivation for Dividends of Listed Companies in China: Evidence Based on Machine Learning

CHEN Yun-sen, ZHOU Jin-yong, PENG Jia-xu

(School of Accountancy, Central University of Finance and Economics)

Abstract: Corporate dividends are not only a direct reflection of shareholders' return on investment but also serve to bolster investor confidence. It is an important means to "build an investor-oriented capital market". However, the "dividend puzzle" remains a fundamental challenge in both theory and practice, as existing theories fail to fully explain the true motivations behind dividend payouts.

This paper employs a variety of machine learning methods to analyze the mechanisms of corporate dividends under the influence of multiple factors, and evaluates the predictive effect of different types of company characteristics on dividend behavior, aiming to more accurately interpret the dividend behavior of China's enterprises. The results are as follows. Firstly, the ensemble learning methods represented by the gradient boosting regression tree and the random forest are outstanding in predicting and explaining corporate dividend behavior, which significantly exceed conventional statistical models such as multiple linear regression. Secondly, among the characteristics of multiple companies, the life cycle characteristics and corporate tax rate characteristics have a better prediction effect on the dividend practice of China's capital market, followed by the first and second types of agency problems, while the financing demand and investor sentiment characteristics have a weaker prediction effect on the dividend behavior of enterprises. Thirdly, among the characteristic variables of various dividend drivers, the ratio of retained earnings to assets, the characteristics of corporate tax rate, the ratio of other receivables to assets, and the level of dividend payment in the previous period contribute the most to the prediction of dividend level. Heterogeneity analysis shows that the main factors affecting corporate dividends are significantly different before and after the promulgation of dividend policies such as semi-mandatory dividend policy, dividend tax reform, and the Cash Dividend Guidelines. There are significant differences between enterprises with different ownership types and different free cash flow conditions. At the same time, whether to adopt a dividend catering strategy and investor dividend sentiment will also affect corporate cash dividend behavior.

The results show that the regulatory authorities can pay more attention to the important characteristics that affect corporate dividend payment behavior, to achieve the purpose of improving corporate cash dividend level. This paper helps investors to identify enterprises that meet their dividend expectations, thus achieving reasonable planning of holding funds and sharing the growth dividend of enterprises earlier. This paper provides new evidence for "strengthening the supervision of cash dividends and enhancing investor returns", and offers important insights for deeply understanding the dividend practices and activating China's modern capital market.

Keywords: dividend policy; corporate dividends; machine learning; dividend puzzle

JEL Classification: G10 G35 M14

[责任编辑:李鹏]