

企业数据资产信息披露与资本市场定价效率

李世刚, 邵宏彬, 方芳, 卢福财

【摘要】 数据资产作为数字经济时代衡量企业竞争力的核心资产,其信息价值尚未得到充分关注。本文以 2007—2022 年 A 股上市公司为样本,利用文本分析法构建数据资产信息披露指标,检验数据资产信息披露水平对资本市场定价效率的影响。研究发现,数据资产信息披露能降低股价同步性,提高资本市场定价效率,该结论经过一系列稳健性检验后依然成立。机制分析发现,数据资产信息披露通过提高信息供给、增强信息扩散、加快信息融入三个方面来提升企业资本市场定价效率。异质性检验发现,在文本相似度更高、地区信息化水平更高及机构投资者注意力更高的样本中,其促进效果更为显著。上述研究为提升企业数据资产信息披露意愿、发挥数据资产价值提供了经验证据,也为监管机构进一步完善数据资产信息披露规范提供了政策启示。

【关键词】 数据资产信息披露; 资本市场定价效率; 股价同步性; 公司特质信息; 信息效率

【中图分类号】 F275 **【文献标识码】** A **【文章编号】** 1006-480X(2025)07-0138-18

一、引言

党的十九届四中全会首次将数据增列为生产要素,数据资源在驱动经济高质量发展中的作用不断提升。2024 年实施的《企业数据资源相关会计处理暂行规定》(简称《暂行规定》)指出,企业应根据数据资源的持有目的、形成方式、业务模式,以及与数据资源有关的经济利益的预期消耗方式等,对数据资源相关交易和事项进行会计确认、计量和报告。但由于会计准则下确认标准的限制,企业在报表中所披露的数据资产主要通过购买价款和开发支出等历史成本计量,难以全面反映数据资源的附加价值与未来潜力。上海交通大学上海高级金融学院发布的《中国企业数据资产入表情况跟踪报告》2024 年年度报告数据显示,截至 2025 年 4 月 30 日,A 股的 5000 多家上市公司中,只有 100 家企业披露了数据资源入表的相关事项,披露数据资源入表信息的上市公司比例不足 2%,无法完全反映企业数据资产的真实情况。同时,一些专业化程度较高的数据资源需要特定企业在

【收稿日期】 2024-11-15

【基金项目】 国家社会科学基金重大招标项目“数据要素驱动经济增长的理论与政策研究”(批准号 23&ZD073);国家自然科学基金地区项目“民营企业 IPO 前‘掏空式分红’、审核问询与原始股东减持”(批准号 72462021);江西省教育科学“十四五”规划课题“江西优秀传统文化与当代大学生社会主义核心价值观塑造:理论逻辑与作用机理”(批准号 22YB051)。

【作者简介】 李世刚,江西财经大学会计学院教授;邵宏彬,江西财经大学会计学院硕士研究生;方芳,江西财经大学会计学院副教授;卢福财,江西财经大学数字经济学院教授,博士生导师。通讯作者:方芳,电子邮箱:fangfang@jxufe.edu.cn。感谢匿名评审专家和编辑部的宝贵意见,文责自负。

特定场景下才能发挥作用,例如,医学领域的医疗影像诊断数据资源、工业物联网领域的设备数据资源等,这也使得通过资产负债表来量化数据资源的价值存在困难。随着文本挖掘和处理技术的进步,文本信息作为市场信息的增量补充同样具有决策效应(Bochkay et al., 2023),《暂行规定》也指出,企业可以根据实际情况,自愿披露数据资源(含未作为无形资产或存货确认的数据资源)在应用场景、应用情况等方面的相关信息,这就为通过文本向外部投资者传递关于企业数据资源的发展潜力、数据资产的发展状况等信息提供了契机。

股价同步性是衡量资本市场定价效率的重要指标,反映了公司特质信息在股价中的融入程度。根据“信息效率假说”,高股价同步性表明股价更多地反映市场或行业信息而非企业本身的运营状况,从而导致资本市场定价效率偏低。已有研究发现,社会责任报告(王艳艳等, 2014)、内部控制报告(赵纳晖和刘瑾, 2024)、管理层讨论与分析(宋昕倍等, 2024)等文本信息供给增加具有提高资本市场定价效率的积极作用,但与传统的社会责任报告或ESG报告等以履责或合规为主要披露目标的信息不同,数据资产信息直接体现了企业在数字经济时代对数据要素应用的能力,管理层可以自主决定是否在年报中披露数据资产信息,以向外界传递企业数字化能力、技术实力以及创新潜力等难以量化的动态竞争优势信息(苑泽明等, 2022)。在数字经济时代,独特的数据资源不仅能作为资产衡量未来经济利益的流入,而且越来越成为企业价值的重要影响因素,例如,Facebook凭借其庞大且粘性高的用户资源实现了上市时超出账面价值20倍的估值,Uber通过搜集用户的出行数据、行为数据和动态定价算法,实现了“互联网+闲置资源”的价值创造模式,成为全球用时最短、估值最高的非上市科技公司。数据资产信息披露不仅可以提高资本市场上公司特质信息的含量,还是企业向投资者展现企业在数据能力建设方面的竞争实力和吸引并帮助投资者评估企业未来成长潜力的重要途径(危雁麟等, 2022),理应能够促进企业信息在资本市场上的流动,提高市场配置效率。为此,探究数据资产信息披露对资本市场定价效率的影响,对于进一步挖掘数据资产价值、发挥资本市场定价作用具有重要意义。

本文的研究贡献主要体现在三个方面:①拓展了数据资产信息披露的经济后果研究。以往文献较多关注数据资产在提高企业竞争力方面的作用,而本文从资本市场定价效率角度切入,揭示数据资产信息披露通过降低信息不对称、提升特质信息含量等途径进而影响市场信息效率的经济后果,建立了企业微观行为与市场配置效率的关联机制。②丰富了信息披露对资本市场定价效率影响的研究成果。现有文献主要探讨了审计报告、内部控制报告、管理层讨论与分析等传统信息披露对资本市场定价效率的影响,对关于数据资产这一具有时代特征信息的披露展开的实证研究还相对较少,本文为非定量信息披露尤其是文本信息在提升资本市场信息效率方面提供了新的实证证据支撑。③为认识数据要素在推动经济高质量发展中的作用提供理论与实证支持。企业通过披露数据资产信息,可以将其对数据要素资源的整合与应用能力传递至资本市场,从而增强投资者信心并优化市场资源配置,在宏观层面促进数字经济发展与资本市场功能的融合。

二、文献回顾

1. 数据资产信息披露的经济后果

企业将数据资源融入具体应用场景及商业模型,能够实现降本增效、提升产品和服务质量、创造经济效益等多重目标,且由于数据资产的非消耗性、时效性以及创造价值的非货币交易模式,现有文献倾向于将数据资产与无形资产相联系(胡亚茹和许宪春, 2022),并认为其信息披露与无形资

产一样具有揭示企业未来成长性、树立良好声誉的关键作用,从而获得投资者青睐并对资本市场资源配置产生积极影响。然而,目前有能力将数据资源形成数据资产并在财务报表中列示的企业还相对较少,这导致财务报表往往难以全面反映企业数据资源的价值,需要通过额外的信息披露帮助外部投资者评估企业的数据资源价值。与此同时,中国资本市场个人投资者数量较多,往往对市场热点形成较高的关注度,而数据资产作为当前市场上的热点话题,其信息披露可以充分调动市场情绪、吸引投资者特别是个人投资者的广泛关注,进而影响其他市场主体的信息搜寻行为。为此,相关研究从上市公司年度报告入手,利用文本分析方法构建数据资产的代理变量,探究数据资产信息披露的价值创造作用,发现企业披露数据资产信息不仅能够强化特质信息供给和提高企业信息透明度,在吸引分析师关注和降低审计费用方面发挥作用(危雁麟等,2022;牛彪等,2024),还可以帮助企业向外界传递自身高水平的数字化能力信号,从而提高人力资本水平以改善劳动力结构(苑泽明等,2024),展示企业良好的发展前景。进一步地,数据资产信息披露能够通过揭示未来的发展潜力获得投资者偏好并缓解企业融资约束(牛彪和于翔,2024;何瑛等,2024),在赋能企业高质量发展以及提升企业价值方面发挥重要作用(苑泽明等,2022)。

2. 资本市场定价效率的影响因素

(1)信息披露与资本市场定价效率。公司内部与外界的信息不对称问题被认为是资本市场资源配置效率较低的重要原因之一(伊志宏等,2019),如何减少资本市场与公司之间的信息不对称成为许多学者关注的重点。现有文献主要从“信息披露水平”和“信息披露特征”两个维度探讨了信息披露对资本市场定价效率的影响。公司的信息披露水平越高,投资者可获取的公司特质信息就越多,从而降低其对市场及行业层面宏观信息的依赖程度,并通过交易将特质信息融入股价从而提高资本市场定价效率。已有研究从社会责任报告披露(王艳艳等,2014)、关键审计事项披露(王木之和李丹,2019)、社交网站信息披露(胡军和王甄,2015)等多个角度证实了上述观点。然而,信息披露水平影响市场信息效率的本质在于是否增加了公司特质信息含量,如果公司的披露行为更多提供的是市场或行业信息,反而会降低资本市场定价效率(王艳艳等,2014),因此公司信息披露质量的改善(如公司特质信息透明度的提高)也能提高资本市场定价效率(Jin and Myers,2006)。伴随大数据挖掘与文本分析技术的发展,越来越多的文献开始关注信息披露文本特征所提供的增量信息的作用。例如,文本相似度可衡量当期文本相对于上一期的特质信息含量变化,文本可读性影响投资者对信息的理解程度,而文本语调则反映管理层对企业未来经营状况的预判。研究表明,当文本相似度较低(宋昕倍等,2024;赵纳晖和刘瑾,2024)、可读性较高(Bai et al.,2019)、语调情感更为真实(吴武清等,2020)时,资本市场定价效率趋于上升。但是,管理层也可能基于私利操纵信息披露,通过不真实的积极报道加剧信息不对称而降低资本市场定价效率(Yu et al.,2013)。

(2)信息环境与资本市场定价效率。作为新兴资本市场,中国证券市场具有个人投资者占比较高的特征,由于缺乏可靠的信息来源及信息挖掘能力,个人投资者常表现出显著的非理性投资行为,需要依赖分析师、新闻媒体等其他市场信息主体的协助。现有文献关于分析师作用发挥的观点存在分歧。部分学者认为,分析师跟踪是挖掘企业特质信息的重要途径,可以通过降低信息不对称来提高资本市场定价效率(朱红军等,2007),并且在分析师报告信息含量越高、分析师影响力越大的情况下,其提高资本市场定价效率的作用更明显(伊志宏等,2019)。另一部分学者则认为,分析师只是信息传递的媒介(胡军和王甄,2015),并非通过挖掘公司特质信息来提高资本市场定价效率,因而分析师传递市场或行业信息的行为反而会降低市场信息效率(黄俊和郭照蕊,2014)。随着通信设施改善和通信技术进步,新闻报道在信息传递过程中所起的作用日益凸显。投资者可通过

财经报刊、电视及互联网等渠道获取更多公司信息,并利用此类信息进行交易,从而提高资本市场定价效率(黄俊和郭照蕊,2014)。机构投资者作为资本市场的重要参与者,对股市信息效率也具有显著影响。相较于个人投资者,机构投资者在信息搜寻、解读、处理及大规模交易能力方面更具优势,能够通过交易机制推动信息融入股价(侯宇和叶冬艳,2008)。同时,机构投资者还能发挥监管作用,通过提升信息透明度来提高资本市场定价效率(王亚平等,2009)。然而,机构投资者行为并非总是理性的,其羊群行为也会降低资本市场定价效率(许年行等,2013)。

3. 文献述评

通过上述文献回顾可以发现,数据资产信息作为一项重要的前瞻性非财务信息,能在一定程度上缓解信息不对称,提升企业透明度,在改善融资约束、降低审计费用等方面起到积极作用(何瑛等,2024;牛彪等,2024),但仍有以下几个方面值得进一步研究:①在研究视角方面,已有文献多聚焦于数据资产信息披露对企业经营效率、竞争实力等微观层面的探讨,较少关注其是否能够被资本市场有效识别,并进一步影响市场信息效率。本文以股价同步性为切入点,实证检验数据资产信息披露对资本市场定价效率的影响,这有助于深化对数据资产外部经济后果的认识。②在信息披露对资本市场定价效率的影响方面,当前研究主要围绕财务报告、审计报告和社会责任报告等文本中的合规信息展开(王艳艳等,2014;王木之和李丹,2019),部分研究开始关注文本相似度、语调、可读性等文本特征所传递的增量信息(Bai et al., 2019; 吴武清等, 2020; 宋昕倍等, 2024),但对数据资产信息这一新型披露内容的研究仍显不足。本文将数据资产纳入信息披露研究范畴,检验其在提升特质信息含量、改善定价效率方面的作用,拓展了非财务信息内容与市场信息效率间关系研究的视角。③在外部信息环境对资本市场定价效率的影响方面,已有研究指出,分析师、媒体和机构投资者等外部主体在企业信息识别、传播和融入股价过程中发挥着关键作用(侯宇和叶冬艳,2008;黄俊和郭照蕊,2014;伊志宏等,2019),在一定程度上提升资本市场的定价效率。基于此,本文进一步论证上述外部信息因素在信息传导路径中的作用,深化对信息环境在企业信息披露影响资本市场反应中的作用机制的理解。

三、理论分析与研究假说

相对于其他国家和地区的投资者,中国资本市场的投资者较多关注国家及行业层面信息,使得中国资本市场的股价同步性处于全球较高水平(Jin and Myers, 2006)。根据“信息效率假说”,数据资产信息能够帮助其他市场主体知悉企业数据资产发展状况,通过提供差异化的信息吸引投资者关注,提高公司特质信息供给;同时,数据资产信息作为衡量企业数字技术发展实力和未来发展潜力的重要来源,容易引起新闻媒体等信息中介的积极报道,从而发挥信息扩散效应,促进公司特质信息的流动;此外,机构投资者具有更多的信息来源和专业的分析能力,可以挖掘并推动公司特质信息快速融入股价。因此,本文从提高信息供给、增强信息扩散、加快信息融入三个方面分析数据资产信息披露降低股价同步性的理论逻辑,进而说明数据资产信息披露对资本市场定价效率的影响。

从数据资产信息披露提高信息供给的角度看,数据资产信息披露可以提高市场上公司特质信息含量,通过影响投资决策来提高资本市场定价效率。当企业处于信息不对称的市场环境中,管理层可以通过披露重要信息,向市场传递企业价值的正面信号,数据资产是衡量企业未来核心竞争力的重要资产,其披露能够帮助投资者获取更多关于企业数字技术实力、数据管理能力和未来增长潜

力的详细信息(危雁麟等,2022),进而对投资者的投资决策产生重要影响。但受限于研发周期较长及入表实践刚起步,其相关信息并不能在财务报表中被完全披露,因而年报的文本信息可以作为对公司特质信息的“增量”补充,弥补数据资产相关信息的缺失。同时,数据资产的形成能够优化企业各个环节的信息处理流程,提升内部信息传递的效率和透明度,从而提高会计信息的质量(牛彪和于翔,2024),高质量的会计信息有助于降低投资者的信息鉴别成本,提高市场和投资者对企业的信任,为企业营造良好的声誉,进一步降低信息不对称程度。分析师作为企业与投资者之间信息传递的中介,由于具备更强的搜集企业深层次信息的能力(Piotroski and Roulstone,2004),可以对企业披露的数据资产信息做进一步挖掘分析,并通过分析师报告将信息传递给投资者。因此,企业数据资产信息数量和质量的提升能够显著改善分析师对公司特质信息的获取能力,减少其对宏观市场和行业信息的依赖,并针对不同公司释放的数据资产信息作出“定制化”的分析报告,提高预测的准确度,为投资者的决策提供可靠的依据,进而提升企业的资本市场定价效率。

从数据资产信息披露增强信息扩散的角度看,数据资产信息披露可以引起新闻媒体等信息中介的关注,通过提高信息流动性来提高资本市场定价效率。在信息技术高度发展的数字时代,首先,新闻媒体不仅是投资者获取信息的重要渠道,还是信息传播的核心媒介,由于当前数据资产在实务界和理论界都备受关注,其披露会迅速吸引新闻媒体的追踪报道,进而将数据资产相关信息传播至更为广泛的投资者群体,加速市场对企业数据资产价值的认知,使得更多投资者能够及时获取企业的特质信息,并通过投资者交易让股价能够更充分地反映企业的特质信息。其次,中国股票市场以个人投资者为主,许多投资者都缺乏专业的信息获取与分析能力,而相较于其他信息中介,新闻媒体是投资者获取信息的主要来源。通过对数据资产信息的解读与分析,并以直观、易懂的方式进行传播,新闻媒体能够帮助非专业投资者更好地理解企业的经营策略和未来潜力,提升信息的流动性。此外,随着对企业数据资产的相关报道数量逐渐上升,新闻媒体还能发挥监督治理作用,通过鉴别企业披露的数据资产信息,帮助投资者识别不同公司数据资产的“优劣”,避免虚假信息误导投资决策。因此,通过新闻媒体的持续评估与跟踪,企业的信息得到扩散并被逐步整合进市场认知中,从而推动更多投资者基于公司特质信息进行投资决策,提升资本市场定价效率。

从数据资产信息披露加快信息融入的角度看,数据资产信息披露能够吸引机构投资者持股,通过发挥机构投资者的交易作用来提高资本市场定价效率。一方面,机构投资者通常偏好信息透明度高、公司治理结构完善及经营风险较低的上市公司(Mccahery et al.,2016),企业披露数据资产信息,有利于机构投资者更全面评估目标企业的数据资产运营状况与潜在风险;另一方面,数据资产作为企业的重要战略资源,代表着企业的核心竞争力和可持续发展潜力,这与稳定型机构投资者对长期投资价值的重视高度契合。因此,企业可以通过数据资产信息披露展现企业在数字化转型与技术创新过程中的竞争优势,进而提升企业对机构投资者的吸引力,促使机构投资者增加持股比例。与此同时,机构投资者在市场中的交易行为是信息融入股价的重要渠道之一,由于机构投资者管理的资金规模庞大,其交易活动能够迅速而明显地影响股价走势。更重要的是,机构投资者在信息收集与分析方面具备强大的专业能力,能够深入挖掘公司特质信息,并在其交易决策中快速整合这些信息(侯宇和叶冬艳,2008)。这种基于深度信息的大规模交易加快了公司特质信息融入股价的过程,使得股价能够更加准确地反映企业的内部信息,减少对宏观市场或行业信息的依赖,从而有效降低股价同步性,提升了资本市场定价效率。

通过上述分析可以发现,数据资产信息披露为市场提供了更多的企业特质信息,帮助投资者更准确地评估企业的真实状况,减少其对宏观市场和行业信息的依赖。同时,数据资产作为企业的重

要资源,对其披露会吸引新闻媒体等信息中介的广泛关注与传播,使其扩散到更广泛的投资者群体,从而提升股价对公司特质信息的反应效率。机构投资者则凭借其专业的分析能力和大规模交易行为,将这些信息迅速反映在市场定价过程中,推动企业特质信息更快融入股价。因此,数据资产信息披露在提高公司特质信息供给、增强公司特质信息扩散、加快公司特质信息融入方面发挥着重要作用,能够显著提升资本市场定价效率。根据上述分析,本文提出:

假说:限定其他条件不变的情况下,上市公司数据资产信息披露水平越高,资本市场定价效率越高。

四、研究设计

1. 样本选择与数据来源

由于2007年新会计准则的实施使得部分年度财务指标的计算发生变化,为了保证会计数据的可比性和连贯性,本文选取2007—2022年全部A股上市公司为初选样本,并对该样本进行如下筛选:①剔除金融类、保险类公司的样本数据;②剔除上市期间曾被ST、*ST、退市等特殊处理的公司数据;③剔除资不抵债即资产负债率大于1的公司样本;④剔除存在缺失值的样本。同时,为剔除部分极端值对研究结果的影响,本文在1%和99%分位对所有连续变量进行Winsorize缩尾处理。经过上述筛选以后,本文最终获得30600个公司年度观测值。本文的数据资产披露信息由企业年报中文本信息加工而成,新闻媒体相关数据来源于中国研究数据服务平台(CNRDS),其他财务、公司治理数据主要来源于国泰安数据库(CSMAR)。

2. 模型及变量说明

为检验数据资产信息披露对资本市场定价效率的影响,借鉴黄俊和郭照蕊(2014)的做法,构建如下模型:

$$SYN_{i,t+1} = \alpha_0 + \alpha_1 DA_{i,t} + \gamma Controls_{i,t} + \sum Firm + \sum Year + \sum Area + \varepsilon_{i,t} \quad (1)$$

其中, i 代表企业, t 代表年份,被解释变量 SYN 为资本市场定价效率,解释变量 DA 为企业数据资产信息披露水平, $Controls$ 表示控制组变量。

(1)被解释变量:资本市场定价效率。资本市场定价效率采用股价同步性指标衡量, SYN 值越高,代表股价同步性越高,股价中包含的公司特质信息越少,资本市场定价效率越低。由于不同行业公司的股价表现不同,本文参考Durnev et al.(2003)的方法,运用(2)式和(3)式计算个股的股价同步性。其中, $R_{i,t}$ 为个股 i 第 t 周的收益率; $R_{m,t}$ 为市场指数第 t 周收益率; $R_{l,t}$ 为行业 l 第 t 周收益率,市场及行业收益率采用总市值加权平均法计算,行业分类方式参照中国证券监督管理委员会2012年公布的分类标准。具体而言,通过(2)式回归得到拟合优度 R^2 ,表示股票 i 的个股收益率随市场收益率和行业整体收益率变动的幅度,即股价中能够被市场和行业信息解释的部分。为使拟合优度指标更接近正态分布,采用(3)式对拟合优度指标进行对数化处理,得到股价同步性指标 SYN 。上述个股、分市场及行业收益率均采用考虑现金红利再投资的回报率。

$$R_{i,t} = \alpha_i + \beta_{i,1} R_{m,t} + \beta_{i,2} R_{l,t} + \varepsilon_{i,t} \quad (2)$$

$$SYN = \log\left(\frac{R^2}{1 - R^2}\right) \quad (3)$$

(2)解释变量:企业数据资产信息披露水平。借鉴苑泽明等(2022)“种子词+Word2Vec相似词

汇扩充”的方法,本文利用文本分析技术,挖掘上市公司年报内容作为企业数据资产信息披露水平度量范围,构建两项指标。具体而言,本文以大数据技术标准推进委员会和中国信息通信研究院2021年发布的《数据资产管理实践白皮书(5.0版)》对数据资产的定义为基准,将“数据资产”和“数据资源”作为种子词汇,利用 Word2Vec 构建相似词集得到“数据资产”相关关键词,统计相似度最高的前十个关键词在年报中出现的频次总数,由于该词频总数呈右偏态分布,对数据资产总词频进行对数化处理得到变量 *LNDA*,并计算“数据资产”相关关键词词频占当年年报总词频的比例得到变量 *PERDA*。

(3)控制变量。借鉴王木之和李丹(2019)、宋昕倍等(2024)、赵纳晖和刘瑾(2024)的研究,控制变量主要包括:企业规模(*Size*)、上市年限(*Age*)、产权性质(*SOE*)、是否为“四大”审计(*Big4*)、资产负债率(*LEV*)、营业收入增长率(*Grow*)、总资产净利率(*ROA*)、市账比(*MB*)、股票换手率(*Turnover*)、两权分离度(*Seperation*)、审计意见(*Opition*)、股权集中度(*TOP1*)、两职合一(*Dual*)、管理层持股比例(*Manhold*)、董事会规模(*Boardsize*)、独立董事占比(*Indir*)、地区市场化指数(*Market*)。为保证回归结果的可靠性,本文进一步控制了公司、年份和省份固定效应,并在公司层面进行聚类调整。具体变量定义与计算方式如表1所示。

表1 主要变量定义		
变量名称	变量符号	变量定义
资本市场定价效率	<i>SYN</i>	下一期的股价同步性指标,使用市场和行业收益率根据(2)式和(3)式计算得到
数据资产信息披露水平	<i>LNDA</i>	数据资产信息披露指标之一,使用“数据资产”相关关键词词频统计总数+1的自然对数度量
	<i>PERDA</i>	数据资产信息披露指标之二,使用“数据资产”相关关键词词频统计总数占年报中总词频的比例度量
企业规模	<i>Size</i>	总资产的自然对数
上市年限	<i>Age</i>	$\ln(\text{样本当前年份}-\text{上市年份}+1)$
产权性质	<i>SOE</i>	企业是否为国有企业,是为1,否则为0
是否为“四大”审计	<i>Big4</i>	财务报告是否由“四大”会计师事务所审计,是为1,否则为0
资产负债率	<i>LEV</i>	总负债/总资产
营业收入增长率	<i>Grow</i>	$(\text{当年主营业务收入}-\text{上年主营业务收入})/\text{上年主营业务收入}$
总资产净利率	<i>ROA</i>	净利润/总资产
市账比	<i>MB</i>	企业市值/账面价值
股票换手率	<i>Turnover</i>	流通股数计算的年内日均换手率取对数
两权分离度	<i>Seperation</i>	控制权与所有权之间的差值
审计意见	<i>Opition</i>	财务报告的审计意见是否为标准无保留意见,是为1,否则为0
股权集中度	<i>TOP1</i>	第一大股东的持股比例
两职合一	<i>Dual</i>	董事长与总经理是否为同一人,是为1,否则为0
管理层持股比例	<i>Manhold</i>	董监高持股数量占总股数量之比
董事会规模	<i>Boardsize</i>	董事会董事数量
独立董事占比	<i>Indir</i>	独立董事人数/董事会总人数
市场化指数	<i>Market</i>	公司注册地所在省份“市场化总指数评分”

3.描述性统计

表2报告了本文主要变量的描述性统计结果。可以发现,SYN的均值约为-0.45,标准差约为1.05,最小值与最大值分别约为-3.92和1.87,说明不同公司之间的股价同步性差异较大。LNDA的均值约为0.53,中位数为0,最小值与最大值分别为0和3.97,说明样本公司目前的数据资产信息披露水平总体偏低;PERDA的均值约为0.003,说明样本企业的数据资产相关信息平均披露水平约占年报词频总体水平的0.3%,总体看数据资产信息披露不足。

表2 主要变量的描述性统计结果

变量符号	观察值	平均值	标准差	最小值	P25	中位数	P75	最大值
SYN	30600	-0.4473	1.0517	-3.9233	-1.0478	-0.3499	0.2739	1.8669
LNDA	30600	0.5274	0.9038	0.0000	0.0000	0.0000	0.6931	3.9703
PERDA	30600	0.0031	0.0065	0.0000	0.0000	0.0000	0.0028	0.0253
Size	30600	22.1703	1.2623	19.8881	21.2648	21.9854	22.8961	26.0988
Age	30600	2.1406	0.7705	0.6931	1.6094	2.3026	2.7726	3.3322
SOE	30600	0.3837	0.4863	0.0000	0.0000	0.0000	1.0000	1.0000
Big4	30600	0.0563	0.2306	0.0000	0.0000	0.0000	0.0000	1.0000
LEV	30600	0.4281	0.2008	0.0586	0.2684	0.4223	0.5789	0.8821
Grow	30600	0.3573	1.0143	-0.7108	-0.0394	0.1190	0.3850	7.3003
ROA	30600	0.0535	0.0633	-0.2033	0.0274	0.0514	0.0830	0.2413
MB	30600	3.9626	2.7768	1.2707	2.3187	3.1898	4.5681	19.2529
Turnover	30600	6.1864	0.7879	4.1038	5.6631	6.2248	6.7600	7.8472
Seperation	30600	4.7701	7.3862	0.0000	0.0000	0.0000	7.6639	28.1632
Opition	30600	0.9765	0.1516	0.0000	1.0000	1.0000	1.0000	1.0000
TOP1	30600	34.7191	14.6280	9.1818	23.2367	32.5915	44.6389	74.2950
Dual	30600	0.2735	0.4457	0.0000	0.0000	0.0000	1.0000	1.0000
Manhold	30600	12.7933	19.1266	0.0000	0.0006	0.3540	23.4489	67.5000
Boardsize	30600	8.5916	1.6846	5.0000	7.0000	9.0000	9.0000	15.0000
Indir	30600	37.3984	5.2352	33.3300	33.3300	33.3300	42.8600	57.1400
Market	30600	9.6106	1.7950	4.1410	8.6680	9.8130	10.8270	12.8640

五、实证结果分析

1.数据资产信息披露与资本市场定价效率基准回归结果

表3列示了数据资产信息披露与资本市场定价效率的基准回归结果。本文采用递进式回归方式,第(1)、(3)列只加入聚类效应,第(2)、(4)列在此基础上加入控制变量以及公司、年份和省份固定效应。结果显示,各列的数据资产信息披露水平指标(LNDA、PERDA)与股价同步性的回归系数均在1%的水平上显著为负。这表明,数据资产信息披露通过提高企业特质信息供给、增强公司特质信息扩散、加快公司特质信息融入,能够有效降低企业股价同步性,提高资本市场定价效率,支持了本文研究假说。

表 3 数据信息披露与资本市场定价效率基准回归结果

	(1)	(2)	(3)	(4)
	<i>DA=LND</i>		<i>DA=PERDA</i>	
<i>DA</i>	-0.1038*** (-11.4776)	-0.0289*** (-2.8105)	-10.9503*** (-8.8193)	-3.7855*** (-2.5999)
控制变量	否	是	否	是
公司固定效应	否	是	否	是
年份固定效应	否	是	否	是
省份固定效应	否	是	否	是
观测值	30600	30070	30600	30070
调整后的 R^2	0.0079	0.4857	0.0045	0.4857

注：***、**、*分别表示 1%、5%、10% 的显著性水平；括号内为 t 值，标准误经公司层面聚类调整，以下各表同。
基准回归的控制变量结果参见《中国工业经济》网站(ciejournal.ajcass.com)附件。

2. 稳健性检验^①

(1) 替换数据资产信息披露衡量变量。为了保障回归结果的可靠性和稳健性，文本选取年报中“数据资产”相关关键词词频出现的绝对数(*TDA*)和年报中是否披露数据资产信息的虚拟变量(*DDA*)替换原解释变量重新回归，结果与基准回归结果保持一致。

(2) 替换资本市场定价效率的代理变量。与基准回归模型中使用总市值加权平均法计算的市场及行业收益率不同，本文在稳健性检验中采用流通市值加权平均法计算市场及行业收益率，再根据(2)式和(3)式进行拟合得到股价同步性 *SYN_M*。同时，借鉴伊志宏等(2019)的做法，在(2)式的基础上加入均滞后一期的市场收益率以及行业收益率进行拟合求出拟合优度 R^2 ，再利用(3)式进行正态化处理得到 *SYN_TIM*，将 *SYN_M* 和 *SYN_TIM* 按照前述研究模型进行稳健性检验，结果与基准回归结果保持一致。

(3) 双向聚类调整。考虑到本文数据为面板数据，回归结果可能仍然存在公司截面相关和时间序列上的自相关问题未能缓解，本文在重新回归时对样本同时进行公司层面和年度层面的双向聚类调整。回归结果表明，在进行双重聚类调整以后，研究结果依然稳健。

(4) 改变样本区间的测试。2019 年年末突发公共卫生事件对全球经济发展造成较大冲击，一定程度上影响了样本企业的经营活动，可能会对数据资产信息披露与资本市场定价效率的关系产生影响。为此，本文剔除了受影响较为严重的 2020 年和 2021 年样本重新进行回归。回归结果表明，在剔除影响因素之后，回归结果依然与基准回归结果保持一致。

(5) 剔除高科技企业样本。考虑到相对于一般企业，高科技企业需要运用数字化技术来进行系统改造和升级，由此具备更强的创新动能，也更容易形成数据资产并在企业年报中披露，因而可能存在行业性差异对本文的结论造成影响。因此，本文将高科技企业从样本中剔除并重新进行回归。检验结果显示，在剔除高科技企业影响以后，回归结果依然是稳健的。

3. 内生性问题控制^②

(1) 倾向得分匹配。由描述性统计可知，尽管大部分企业已经实施数字化转型战略，但能够

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

② 内生性问题控制结果参见《中国工业经济》网站(ciejournal.ajcass.com)附件。

将数据资产信息在年报中披露并展示的企业还相对较少,可能导致实验组样本过少。同时,能够披露数据资产的企业可以输出更加结构化、标准化的有效信息(苑泽明等,2024),降低利益相关者的信息获取成本,提升股价信息含量,可能会对研究结论产生影响。为此,本文采用倾向得分匹配(PSM)进行内生性检验,将披露了数据资产信息的样本作为处理组,未披露数据资产信息的样本为对照组,选取基准回归模型中的控制变量作为协变量集,使用最近邻匹配方法分别进行混合面板数据匹配和分年度数据匹配。回归结果显示,匹配后的估计结果与基准回归结果保持一致。

(2)熵平衡检验。为了缓解遗漏变量和样本选择问题导致的结论偏差,本文采用熵平衡检验来控制实验组和对照组样本之间的系统性差异。同样地,本文选取基准模型中的控制变量作为协变量集,对该协变量集构建矩阵并赋予相应权重,利用加权后的样本重新回归。回归结果显示,匹配后的估计结果与基准回归保持一致。

(3)考虑时间变化趋势。由于不同地区和不同行业企业的数据资产发展水平有所差异,可能存在地区和行业间的固有差异随时间变化导致的估计偏误问题,本文在基准模型的基础上分别加入地区与时间趋势的交互项($Area \times Year$)和行业与时间趋势的交互项($Ind \times Year$)并重新回归。回归结果显示,在考虑地区随时间变化的趋势和行业随时间变化的趋势以后,回归结果与基准结果保持一致。

(4)处理效应模型。考虑到不同行业企业的数字化进程存在差异,如信息技术和软件等相关高科技行业企业能更好地运用数字技术,更容易形成数据资产并在年报中披露,本文采用处理效应模型控制可能存在的样本自选择问题。检验结果与基准回归结果保持一致,说明本文的结果是稳健的。

六、机制检验

1.提高信息供给

根据前文分析,数据资产信息披露一方面能提升企业信息透明度,通过提高公司特质信息供给来缓解与外界的信息不对称程度并吸引投资者及分析师的关注度;另一方面,可以提高分析师报告中的公司特质信息含量和预测准确性,并通过分析师的信息获取和传递能力进一步帮助投资者进行交易决策,从而提高资本市场定价效率。本文对数据资产信息披露提高市场信息供给的机制进行检验,分别使用分析师关注度($Follow$)和分析师预测偏差($Bias$)两个变量来衡量数据资产信息披露提高信息披露水平和提升信息质量程度。借鉴危雁麟等(2022)的研究,使用在一年内对该公司进行过跟踪分析的分析师(团队)数量刻画企业分析师关注度($Follow$),并使用所有分析师当年最近一次每股盈余预测的标准差与公司期初股票价格的比值刻画分析师的预测偏差程度($Bias$)。

表4的检验结果表明, $LNDA$ 和 $PERDA$ 与分析师关注度变量($Follow$)的系数均在5%的水平上显著为正,与分析师预测偏差变量($Bias$)的系数分别在5%和1%的水平上显著为负,说明企业披露数据资产信息的确通过降低信息不对称提高了分析师对企业的关注程度,并通过提高信息披露质量降低了分析师预测的偏差程度,进而提高股价中公司特质信息的含量,提升资本市场定价效率。

表 4 提高信息供给机制检验

	(1)	(2)	(3)	(4)
	<i>DA=LND A</i>		<i>DA=PERDA</i>	
	<i>Follow</i>	<i>Bias</i>	<i>Follow</i>	<i>Bias</i>
<i>DA</i>	0.0272** (2.2366)	-0.0004** (-2.5762)	3.8045** (2.2727)	-0.0545*** (-3.0189)
控制变量	是	是	是	是
公司固定效应	是	是	是	是
年份固定效应	是	是	是	是
省份固定效应	是	是	是	是
观测值	30070	16794	30070	16794
调整后的 R ²	0.7031	0.2561	0.7031	0.2562

2. 增强信息扩散

新闻媒体报道是广大投资者,尤其是个人投资者,获取上市公司信息的重要途径,作为数字经济领域的热门话题,数据资产信息一经披露便易引发新闻媒体的广泛报道。媒体不仅将企业披露的数据资产信息传播至更广泛的受众,还通过“二次加工”帮助投资者更好地理解这些信息,并协助其识别信息的“真伪”。通过加速信息的扩散,企业的特质信息得以更快地在市场中流通,从而更迅速地反映在股价中,提高资本市场定价效率。基于此,本文使用中国研究数据服务平台(CNRDS)中网络财经新闻库的“网络新闻报道数”和“财经报刊新闻报道数”衡量媒体对上市公司数据资产信息扩散的影响。具体地,本文采用网络新闻报道内容中出现上市公司的新闻总数(*News_cont1*)以及财经报刊和网络新闻报道内容中出现上市公司的新闻总数(*News_cont2*)两个指标来度量新闻媒体对上市公司的数据资产信息扩散效应。

由表 5 可知,解释变量 *LND A* 和 *PERDA* 与网络新闻内容报道数(*News_cont1*)、财经报刊和网络新闻内容报道数(*News_cont2*)的系数均在 5% 的水平上显著为正,说明企业披露数据资产信息以后,确实吸引了更多新闻媒体的关注。随着企业披露的信息进一步扩散至更广泛的投资者群体,特质信息被更为广泛地纳入投资者的决策过程,股价信息也就能更充分地反映企业的自身运营状况,进而提高资本市场定价效率。

表 5 增强信息扩散机制检验

	(1)	(2)	(3)	(4)
	<i>DA=LND A</i>		<i>DA=PERDA</i>	
	<i>News_cont1</i>	<i>News_cont2</i>	<i>News_cont1</i>	<i>News_cont2</i>
<i>DA</i>	7.6413** (2.1507)	10.3493** (2.1564)	1086.9240** (2.1313)	1377.6680** (2.0274)
控制变量	是	是	是	是
公司固定效应	是	是	是	是
年份固定效应	是	是	是	是
省份固定效应	是	是	是	是
观测值	30034	29328	30034	29328
调整后的 R ²	0.7486	0.7843	0.7486	0.7843

3. 加快信息融入

前文认为,机构投资者通常倾向于选择信息透明度高、公司治理结构健全且经营风险相对较低的上市公司。数据资产信息披露不仅提升了企业的透明度,使机构投资者能够更准确评估其运营状况和潜在风险,还展示了企业在数字化转型与技术创新中的竞争优势,与机构投资者对长期投资价值的追求相契合。凭借其信息分析能力和大规模交易手段,机构投资者能够迅速将企业特质信息融入股价,推动特质信息更快被市场反应,减少对宏观信息的依赖。基于此,本文采用机构投资者持股比例作为衡量机构投资者加快信息融入效应的代理指标。

表 6 的检验结果显示, *LNDA* 和 *PERDA* 与机构投资者持股比例的系数均在 5% 的显著性水平上呈正相关。说明企业披露数据资产信息,能够提升机构投资者的关注度,并促使其增加持股比例。随着机构投资者对企业特质信息的深入挖掘和交易决策,股价更快速地反映出企业的真实经营状况,减少了对宏观市场或行业信息的依赖,从而有效提高资本市场定价效率。

表 6	加快信息融入机制检验	
	(1)	(2)
	<i>DA=LNDA</i>	<i>DA=PERDA</i>
<i>DA</i>	2.5963** (2.0324)	407.9662** (2.3200)
控制变量	是	是
公司固定效应	是	是
年份固定效应	是	是
省份固定效应	是	是
观测值	30070	30070
调整后的 R ²	0.4461	0.4461

七、异质性分析

本文的理论框架揭示了数据资产信息披露通过“信息供给—信息扩散—信息融入”三阶段路径提升资本市场定价效率的内在逻辑。然而,上述机制的有效性可能受制于企业自身信息供给、外部信息流动环境以及机构投资者行为特征的异质性影响。为有效识别不同机制的作用条件,本文基于信息供给、信息扩散、信息融入这三条作用路径,分别选取文本信息相似度、地区信息化水平以及机构投资者分心程度作为调节变量进行实证检验。文本信息相似度反映了企业当年与上年文本信息含量的差异程度,可揭示在企业特质信息供给程度不同时数据资产信息披露带来的作用差异;地区信息化水平影响着信息扩散的效率,能够体现信息扩散机制下数据资产信息披露在不同环境中的扩散效果;机构投资者分心程度刻画了机构投资者注意力的分散程度,代表了在信息融入机制过程中,数据资产信息能否被机构投资者有效获取并通过交易融入股价。通过对这三个方面的异质性检验,可有效辨析在不同情境下数据资产信息披露对资本市场定价效率影响的差异化表现,并进一步验证和丰富了前文机制检验的结论,使研究更具科学性和说服力。

1. 文本信息相似度

企业年度财务报告作为上市公司强制性定期披露报告,是投资者了解企业发展状况和未来

前景的重要途径,而管理层讨论与分析(MD&A)是年报中最重要、信息含量最多的部分(王雄元等,2018),其内容中关于管理层对重大事项的理解和未来前景的展望能够满足信息使用者对相关性和可靠性和前瞻性信息的需求(Clarkson et al.,1999)。文本相似度是财务报告的重要特征之一,当年年报与上一年年报的余弦相似度,可以反映企业当年的文本特质信息含量高低。根据前文论述,数据资产信息披露可以通过提高公司特质信息供给、帮助投资者了解企业未来发展前景来提高资本市场定价效率。因此,本文预期,如果企业当年的MD&A与上一年计算出的文本相似度较高,说明企业当年的MD&A提供的特质信息较少,此时在年报中的数据资产信息披露作为非财务信息的重要补充,能够在帮助投资者缓解信息不对称、提高资本市场定价效率方面发挥更加显著的作用。

为检验MD&A文本相似度对数据资产信息披露的异质性作用,本文借鉴王雄元等(2018)的做法,使用企业年报中MD&A的文本相似度作为公司特质信息披露的衡量指标,并根据行业一年度中位数对样本进行划分,分为文本相似度高组和文本相似度低组并分别进行回归分析,回归结果如表7所示。可以发现,在文本相似度更高的样本中,LNDA和PERDA的系数均显著为负;在文本相似度更低的样本中,数据资产信息披露变量系数均不显著。上述结果表明,在文本相似度更高的样本中,数据资产信息披露可以提供更多的公司特质信息,从而更有效地降低股价同步性,提高资本市场定价效率。

表 7 文本信息相似度异质性影响

	(1)	(2)	(3)	(4)
	DA=LNDA		DA=PERDA	
	文本相似度高	文本相似度低	文本相似度高	文本相似度低
DA	-0.0400*** (-2.6453)	-0.0138 (-0.9227)	-4.1802* (-1.9548)	-2.8746 (-1.3513)
控制变量	是	是	是	是
公司固定效应	是	是	是	是
年份固定效应	是	是	是	是
省份固定效应	是	是	是	是
观测值	14548	14432	14548	14432
调整后的 R ²	0.4947	0.4849	0.4945	0.4850

2.地区信息化水平

投资者能否获取股市的相关信息对其投资决策有重要的影响,从信息渠道影响股市投资行为的视角看,电脑以及互联网等信息技术的普及不仅降低了交易成本,还减少了投资者获取股市信息的成本,从而提高投资者对股市的参与程度。考虑到中国股市的实际情况,网络信息的可获得性在很大程度上体现了一般化的股市信息的可获得情况(郭士祺和梁平汉,2014)。根据前文论述,数据资产信息披露可以吸引新闻媒体等信息中介的关注,进而将信息扩散到更广泛的投资者群体以提高资本市场定价效率,而投资者的信息获取渠道则决定了媒体的新闻报道能否真正传播到有需要的投资者手中,因而以互联网普及状况为代表的地区信息化发展水平决定了个人投资者能否及时获取公司特质信息的传播。本文预期,在信息化水平更高的地区,投资者获取公司特质信息的方式更加便捷,信息获取和传播成本显著降低,使公司特质信息得以通过媒体等信息中介迅速传播至市

场各方,数据资产信息披露对资本市场定价效率的促进作用更加明显。

为检验地区信息化水平对数据资产信息披露的异质性作用,本文使用样本企业所在省份的每百人使用计算机数作为地区信息化水平的代理变量,并根据省份一年度中位数对样本企业进行划分,分为地区信息化水平高组和地区信息化水平低组,分别进行回归检验。表8列示了相关检验结果,可以发现,在地区信息化水平高的企业中,*LNDA*和*PERDA*的系数均显著为负;而在地区信息化水平低的组别中,数据资产信息披露各变量系数均不显著。上述结果联合表明,信息化水平较高的地区信息传播效率更高,公司特质信息流动更快,企业数据资产信息披露对资本市场定价效率的提升作用更加显著。

表 8 地区信息化水平异质性影响

	(1)	(2)	(3)	(4)
	<i>DA=LNDA</i>		<i>DA=PERDA</i>	
	地区信息化水平高	地区信息化水平低	地区信息化水平高	地区信息化水平低
<i>DA</i>	-0.0259* (-1.7996)	-0.0145 (-0.5222)	-3.7358* (-1.7997)	-2.2395 (-0.5609)
控制变量	是	是	是	是
公司固定效应	是	是	是	是
年份固定效应	是	是	是	是
省份固定效应	是	是	是	是
观测值	15985	5905	15985	5905
调整后的 R ²	0.4287	0.5242	0.4287	0.5242

3. 机构投资者注意力

根据前文论述,数据资产信息披露通过展现企业在数字化技术能力与运用上的竞争优势,可以吸引机构投资者增加持股比例,并通过机构投资者的交易行为将数据资产所体现的公司价值信息融入股价当中。然而,已有研究发现,由于有限关注的限制,机构投资者无法在同一时间对所有的信息给予充分关注并进行应对,因而对某个投资标的的关注会减少对其他投资标的的注意力(袁玉等,2024),从而产生“机构投资者分心”现象。一方面,机构投资者的分心程度越大,机构投资者越不可能对市场中的所有信息和投资组合公司的私有信息进行充分理解和消化,也无法进一步挖掘公司层面的有价值的特质信息;另一方面,分心的机构投资者会表现出更少的交易行为,难以将其掌握的公司层面的私有信息融入股价中,导致公司股价中信息融入的含量减少,无法向其他外部投资者传递更多有价值的增量信息。因此,本文预期,当企业的机构投资者分心程度更高时,该企业披露的数据资产信息并不能及时得到机构投资者的关注,机构投资者通过进一步挖掘和交易将数据资产信息融入股价中的作用也无法充分发挥。

为检验机构投资者分心对数据资产信息披露的异质性作用,本文借鉴袁玉等(2024)的做法,构建公司层面的机构投资者分心程度指标,按照行业一年度中位数将样本划分为机构投资者分心程度高组和机构投资者分心程度低组,回归结果如表9所示。可以发现,在机构投资者分心程度更低的样本中,*LNDA*和*PERDA*的系数分别在5%和1%的水平上显著为负;在机构投资者分心程度更高的样本中,数据资产信息披露变量系数均不显著。上述结果表明,在机构投资者分心程度更低的公司中,机构投资者可以更及时地关注到企业数据资产信息披露的状况,并通过进一步挖掘和交易将

该企业的数据资产特质信息融入股价当中,从而更有效地降低股价同步性、提高资本市场定价效率。

表 9 机构投资者注意力异质性影响

	(1)	(2)	(3)	(4)
	DA=LND A		DA=PERDA	
	分心程度高	分心程度低	分心程度高	分心程度低
DA	-0.0230 (-1.5864)	-0.0410** (-2.5222)	-2.7977 (-1.3739)	-5.9758*** (-2.6461)
控制变量	是	是	是	是
公司固定效应	是	是	是	是
年份固定效应	是	是	是	是
省份固定效应	是	是	是	是
观测值	14665	14323	14665	14323
调整后的 R ²	0.4923	0.4657	0.4923	0.4657

八、结论与启示

数据资产作为数字经济时代衡量企业核心竞争力的重要资产,其信息披露水平不仅影响投资者对其价值信息能否获取、反应以及获取和反应程度,更影响资本市场对数据资产信息的定价效率。本文基于信息效率假说,选取 2007—2022 年沪深 A 股上市公司数据,运用文本分析方法构建企业数据资产信息披露水平指标,探讨了企业数据资产信息披露对资本市场定价效率的影响。研究结果显示,企业披露数据资产信息能降低股价同步性,提升资本市场定价效率,在进行一系列稳健性检验和内生性问题处理以后,上述结论依然成立。机制研究发现,数据资产信息披露可以通过提高信息供给、增强信息扩散、加快信息融入等方式提高资本市场定价效率。通过异质性分析发现,数据资产信息披露对资本市场定价效率的促进效果在文本相似度高、地区信息化水平高及机构投资者注意力更高的样本中更加显著。基于上述研究发现,本文提出如下启示:

(1)上市公司应提高数据资产信息披露水平,通过提升企业信息披露的数量和质量,促进资本市场对企业定价效率的提升。本文的研究结论表明,数据资产信息披露对公司特质信息的供给、扩散和融入股价具有积极的作用,这对提高企业竞争力、提升资本市场效率至关重要。为此,企业应建立标准化的披露机制,确保信息披露的规范性和透明度,并采用云平台等数字化工具对数据资产信息进行实时更新,确保投资者能及时获取关键信息;同时,企业可以根据数据资产的形成过程将信息披露阶段化,从而使投资者保持信心。例如,早期阶段可以披露技术基础、资源投入及预期效益,中期阶段披露项目进展与市场反馈,数据资产形成后则披露其实际应用及财务影响,使投资者增进对数据资产的了解,并保持对企业的信心。此外,企业可以结合行业对比,通过提供数据资产管理的对标分析,展示企业在数据资产应用和价值转化上的优势,帮助投资者更全面评估企业的数据资产创新能力和长期增长潜力。

(2)上市公司在编制年报以及其他定期报告时,应特别关注数据资产信息披露的内容差异化和文本特质信息的丰富性,提高企业特质信息的供给。本文的研究发现,数据资产信息披露在企业年报 MD&A 文本相似度较高的情境下,能更好地弥补报告中信息供给不足的问题,显著提高资本市场

的定价效率。为此,企业在编制各类涉及数据资产对外报告时,应注重通过多样化、差异化的数据资产信息文本展示企业独有的特质信息,尤其要避免单纯的内容复述和格式化、模板化的更新,充分利用数据资产等信息作为高附加值的披露内容,凸显企业数字化能力和创新驱动优势。同时,可通过结合数据资产的应用场景、数字化转型成果及预期商业化潜力,弥补财务报表无法直接通过报表完整呈现数据资产信息价值的缺点。此外,企业还应建立起更灵活和动态化的数据资产信息披露更新机制,鼓励管理层及时披露数据资产利用和商业化路径上的新进展,确保对外报告能够真实反映企业的最新发展动态和战略方向,帮助投资者形成更准确的价值判断并树立长期信心。

(3)有关部门应加快区域信息化水平建设,推动资本市场对企业特质信息快速反应,提升区域内的信息传播效率。信息化水平直接影响区域内信息流动的速度,较高的信息化水平能够加速公有信息和私有信息的传递,增强市场对企业特质信息的反应速度。因此,政府应优先加大对通信基础设施的投入,提升网络覆盖和数据传输能力,确保企业数据资产信息能够更快速、更广泛地传递至市场各方。与此同时,政府可以积极推动建立区域性信息共享平台,促进企业、投资者和信息中介机构之间的高效互动和信息交流,通过降低信息获取和传播成本,帮助市场更迅速地吸收和处理公司的特质信息,从而提升资本市场效率。此外,政府还可以支持跨区域的信息合作,通过推动不同地区之间的数字化平台互联互通,提升信息跨区域流动,这不仅能帮助区域内企业获取更广泛的市场信息,还能促使公司特质信息在全国范围内迅速传播,增强资本市场对公司信息的敏感度,提升资本市场对公司数据资产信息的定价效率,从而促进区域经济的可持续增长与高质量发展。

(4)机构投资者应提升信息挖掘与获取能力,科学合理地分配有限的注意力资源,以加快对数据资产信息等高价值特质信息的获取和市场融入,提升资本市场定价效率。本文研究结果表明,数据资产信息披露能够通过吸引机构投资者关注并通过机构投资者的交易将信息融入股价中,从而提高资本市场定价效率,但上述效果仅在机构投资者注意力集中时才能显现。因此,机构投资者应积极运用大数据分析、文本挖掘等技术手段,对企业年报、管理层讨论与分析(MD&A)等非财务文本进行深度剖析,识别企业核心数据资产战略及其潜在经济价值,增强对数字化转型和数据要素化利用的研究能力,形成更加前瞻和深入的投资判断。在投资组合管理中,应优化注意力分配策略,重点关注核心持仓企业及其高附加值的数据资产信息披露,避免因注意力分散而忽视关键数据资产信息。具体而言,应动态调整投资组合的研究重心和覆盖面,合理安排跟踪频率和研究深度,确保有限注意力被更充分地用于挖掘高质量、差异化的数据资产特质信息。通过上述举措,机构投资者不仅能够更有效地识别和利用企业数据资产信息的潜在价值,还能通过积极的交易行为加速信息在资本市场中的传递与定价,进一步推动资本市场资源配置的优化和高效运行。

〔参考文献〕

- [1]郭士祺,梁平汉. 社会互动、信息渠道与家庭股市参与——基于2011年中国家庭金融调查的实证研究[J]. 经济研究, 2014, (S1): 116-131.
- [2]何瑛,陈丽丽,杜亚光. 数据资产化能否缓解“专精特新”中小企业融资约束[J]. 中国工业经济, 2024, (8): 154-173.
- [3]侯宇,叶冬艳. 机构投资者、知情人交易和市场效率——来自中国资本市场的实证证据[J]. 金融研究, 2008, (4): 131-145.
- [4]胡军,王甄. 微博、特质性信息披露与股价同步性[J]. 金融研究, 2015, (11): 190-206.
- [5]胡亚茹,许宪春. 企业数据资产价值的统计测度问题研究[J]. 统计研究, 2022, (9): 3-18.
- [6]黄俊,郭照蕊. 新闻媒体报道与资本市场定价效率——基于股价同步性的分析[J]. 管理世界, 2014, (5): 121-130.

- [7]牛彪,于翔.数据资产获得投资者偏好了吗?——基于权益资本成本视角[J].证券市场导报,2024,(7):68-79.
- [8]牛彪,于翔,苑泽明,丁亚楠.数据资产信息披露与审计师定价策略[J].当代财经,2024,(2):154-164.
- [9]宋昕倍,陈莹,逯东,程杰.信息环境、上市公司增量信息披露与资本市场定价效率——基于MD&A文本相似度的研究[J].南开管理评论,2024,(5):30-39.
- [10]王木之,李丹.新审计报告和股价同步性[J].会计研究,2019,(1):86-92.
- [11]王雄元,高曦,何捷.年报风险信息披露与审计费用——基于文本余弦相似度视角[J].审计研究,2018,(5):98-104.
- [12]王亚平,刘慧龙,吴联生.信息透明度、机构投资者与股价同步性[J].金融研究,2009,(12):162-174.
- [13]王艳艳,于李胜,安然.非财务信息披露是否能够改善资本市场信息环境?——基于社会责任报告披露的研究[J].金融研究,2014,(8):178-191.
- [14]危雁麟,张俊瑞,汪方军,程茂勇.数据资产信息披露与分析师盈余预测关系研究——基于文本分析的经验证据[J].管理工程学报,2022,(5):130-141.
- [15]吴武清,赵越,闫嘉文,汪寿阳.分析师文本语调会影响股价同步性吗?——基于利益相关者行为的中介效应检验[J].管理科学学报,2020,(9):108-126.
- [16]许年行,于上尧,伊志宏.机构投资者羊群行为与股价崩盘风险[J].管理世界,2013,(7):31-43.
- [17]伊志宏,杨圣之,陈钦源.分析师能降低股价同步性吗——基于研究报告文本分析的实证研究[J].中国工业经济,2019,(1):156-173.
- [18]袁玉,吴战旒,廖佳.机构投资者分心会加剧管理层语言膨胀吗?——基于年报文本语调的实证研究[J].南开管理评论, <http://kns.cnki.net/kcms/detail/12.1288.F.20240530.0908.002.html>, 2024.
- [19]苑泽明,于翔,李萌.数据资产信息披露、机构投资者异质性与企业价值[J].现代财经(天津财经大学学报), 2022,(11):32-47.
- [20]苑泽明,于翔,李萌,刘冠辰.数据资产促进了中国企业人力资本水平提升吗?——基于文本分析法经验证据[J].南开管理评论, <http://kns.cnki.net/kcms/detail/12.1288.f.20240919.0839.002.html>, 2024.
- [21]赵纳晖,刘瑾.内部控制评价报告文本特征与资本市场定价效率[J].审计研究,2024,(3):148-160.
- [22]朱红军,何贤杰,陶林.中国的证券分析师能够提高资本市场的效率吗——基于股价同步性和股价信息含量的经验证据[J].金融研究,2007,(2):110-121.
- [23] Bai, X. L., Y. Dong, and N. Hu. Financial Report Readability and Stock Return Synchronicity[J]. Applied Economics, 2019, 51(4): 346-363.
- [24] Bochkay, K., S. V. Brown, A. J. Leone, and J. W. Tucker. Textual Analysis in Accounting: What's Next[J]. Contemporary Accounting Research, 2023, 40(2): 765-805.
- [25] Clarkson, P. M., J. L. Kao, and G. D. Richardson. Evidence That Management Discussion and Analysis (MD&A) Is a Part of a Firm's Overall Disclosure Package[J]. Contemporary Accounting Research, 1999, 16(1): 111-134.
- [26] Durnev, A., R. Morck, B. Yeung, and P. Zarowin. Does Greater Firm-specific Return Variation Mean More or Less Informed Stock Pricing[J]. Journal of Accounting Research, 2003, 41(5): 797-836.
- [27] Jin, L., and S. C. Myers. R² Around the World: New Theory and New Tests[J]. Journal of Financial Economics, 2006, 79(2): 257-292.
- [28] McCahery, J. A., Z. Sautner, and L. T. Starks. Behind the Scenes: The Corporate Governance Preferences of Institutional Investors[J]. Journal of Finance, 2016, 71(6): 2905-2932.
- [29] Piotroski, J. D., and B. T. Roulstone. The Influence of Analysts, Institutional Investors, and Insiders on the Incorporation of Market, Industry, and Firm-Specific Information into Stock Prices[J]. The Accounting Review, 2004, 79(4): 1119-1151.
- [30] Yu, Z., B., L. C. Li, G. L. Tian, and H. Zhang. Aggressive Reporting, Investor Protection and Stock Price Informativeness: Evidence from Chinese Firms[J]. Journal of International Accounting, Auditing and Taxation, 2013, 22(2): 71-85.

Firm Data Asset Disclosure and Capital Market Pricing Efficiency

LI Shi-gang, SHAO Hong-bin, FANG Fang, LU Fu-cai
(School of Accounting, Jiangxi University of Finance and Economics)

Abstract: With the rapid development of the digital economy and its increasingly prominent role in driving China's high-quality economic growth, data resources have emerged as a critical component within industries such as finance, healthcare, and energy. Unique data resources not only serve as an important asset in forecasting future economic inflows but also constitute a key determinant of enterprise value. Through proactively disclosing information about data assets, organizations can convey their dynamic comparative advantages in digital capabilities and innovation potential to external stakeholders, thereby attracting investors and enhancing capital market information efficiency. The Provisional Regulations on Accounting Treatment for Corporate Data Resources, implemented by the Ministry of Finance of China in 2024, clarify that data resources that meet the criteria for asset recognition may be identified as assets and quantified in financial statements, providing a potential pathway for realizing the value of data assets. However, under current accounting standards, only a limited number of firms can report data resources as assets. Certain specialized data cannot be accurately valued due to constraints on usage scenarios, making it difficult for investors to obtain a clear picture of a firm's data assets solely from financial statements. To alleviate the constraints, the regulation allows companies to voluntarily disclose additional information on data resources not recognized as assets, including their application scenarios, usage status, and methods for value creation. This allows firms to communicate the utilization and development potential of their data resources through financial reports.

Advances in text mining and information processing technologies enable textual content to be quantified and leveraged to support investor decision-making. Using a sample of A-share listed companies in China from 2007 to 2022, this study employs textual analysis to construct the data assets disclosure index. Then, this study empirically examines the effect of data assets disclosure on capital market pricing efficiency. The results show that data assets disclosure significantly reduces stock price synchronicity, thereby enhancing pricing efficiency. The conclusion is robust across a series of robustness checks and remains significant after addressing potential endogeneity issues. Mechanism analysis reveals that data assets disclosure improves pricing efficiency by enhancing firm-specific information supply, facilitating information dissemination, and accelerating its incorporation into stock prices. Further heterogeneity analysis shows that the effect is more pronounced among firms with higher textual similarity in management discussion and analysis (MD&A), among those located in regions with higher informatization, and those receiving greater attention from institutional investors.

This study contributes to existing literature in several ways. First, it extends the understanding of the economic consequences of data assets disclosure. Prior researches focus primarily on its impact on alleviating firms' financing constraints and enhancing competitiveness, and few studies investigate whether such disclosures can be effectively interpreted by capital markets to improve resource allocation. By examining its role in reducing stock price synchronicity, this study sheds light on the external economic consequences of data assets disclosure from the perspective of information efficiency. Second, it enriches the body of research on the determinants of capital market pricing efficiency and provides new empirical evidence on how the disclosure of qualitative textual data, especially non-financial information, can enhance market efficiency. Lastly, this study provides both theoretical framework and empirical support for the role of data assets as a productive factor in driving high-quality economic development. By voluntarily disclosing data assets information, firms can demonstrate their capacity for data integration and application, bolster investor confidence, and contribute to efficient resource allocation in capital markets, thereby facilitating the integration of the digital economy with the functions of the capital market.

Keywords: information disclosure of data assets; capital market pricing efficiency; stock price synchronicity; company-specific information; information efficiency

JEL Classification: M40 G10 M21

[责任编辑:王燕梅]