

都披露等于都没披露？数据资产信息披露、投资者有限注意力与资本市场

定价效率

彭林婧，杨雨坤[†]

摘要：本文利用中国A股上市公司数据，基于有限注意力理论，重新审视了数据资产信息披露对资本市场定价效率的影响。研究发现，数据资产信息披露的经济后果并非静态不变，而是随着市场整体披露水平的提升呈现显著的边际递减特征。具体而言，在市场整体披露水平较低时，企业披露数据资产信息能够通过增加特质信息供给、增强信息扩散及加速信息融入来显著降低股价同步性，提高资本市场定价效率。然而随着近年来披露行为的普遍化，没有证据表明上述机制仍然有效。进一步分析表明，仅在市场整体数据资产信息披露水平较低，且企业披露的特质信息相对较少、信息传播环境优越或机构投资者注意力未被过度分散的情境下，披露行为才能维持其正向效应。本文拓展了有限注意力理论的边界，将研究视角延伸至同质化信息对投资者注意力的分散这一维度，揭示了数据资产信息披露在市场普遍化情境下面临的价值稀释问题。

关键词：数据资产信息披露、有限注意力理论、资本市场定价效率、股价同步性、价值稀释

1 引言

近年来，随着数据被正式确立为关键生产要素，企业对数据资源的管理、开发与价值实现已成为数字经济时代的重要议题(Jones and Tonetti, 2020)。中国2024年实施的《企业数据资源相关会计处理暂行规定^①}（以下简称《暂行规定》）为企业数据资源的会计处理提供了规范指导(黄蓉 et al., 2025)。然而，受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(王艳 and 杨达, 2024)。因此，企业通过年报文本自愿披露数据资产相关信息的行为能够为外部投资者提供重要的增量信息(Bochkay et al., 2023; 李世刚 et al., 2025)。《暂行规定》亦鼓励企业根据实际情况，自愿披露未确认为无形资产或存货的数据资源在应用场景、使用成效等方面的信息。在此背景下，李世刚、邵宏彬、方芳和卢福财(李世刚 et al., 2025)利用机器学习方法分析了2007到2022年中国A股上市公司的年报，测度了它们的数据资产信息披露水平(以及股价同步性)。他们的研究表明，企业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。然而，笔者在利用李世刚等(李世刚 et al., 2025)在《中国工业经济》官网^②上公开提供的数据进行研究时却发现，数据资产信息披露与资本市场定价效率的关系出现显著的时变特征，当仅使用最近几年(例如2020-2022年)样本进行检验时，数据资产信息披露对资本市场定价效率的影响不再具有统计显著性。这一发现挑战了现有研究对数据资产信息披露效果的普遍认知，提示我们可能需要重新审视数据资产信息披露的经济后果。

股价同步性是衡量资本市场定价效率的重要代理变量，反映了公司特质信息在股价中的

[†] 作者信息：彭林婧，目前就读于北京理工大学管理学院会计学专业且预计于2027年7月毕业，邮箱：Linda13677066050@163.com。杨雨坤，目前就读于北京理工大学管理学院信息管理与信息系统专业且预计于2027年7月毕业，邮箱：yangyukun2005@163.com。杨雨坤是该论文的通讯作者。(作者信息截至2026年2月7日)

访问杨雨坤的个人网页 <https://yangyukun2005.github.io/publications/> 可获取该论文的最新版本。

^① https://www.gov.cn/zhengce/zhengceku/202308/content_6899395.htm

^② 《中国工业经济》是一本中文顶级经济学期刊，他们论文的网页为：
<https://ciejournal.ajcass.com/Magazine/show/?id=119653>

融入程度(Chan and Hameed, 2006)。根据“信息效率假说”，较高的股价同步性意味着股价更多反映市场或行业信息，而非企业本身运营状况，这代表了较低的资本市场定价效率(Morck et al., 2000)。已有研究表明，社会责任报告(王艳艳 et al., 2014)、内部控制报告(赵纳晖 and 刘瑾, 2024)以及管理层讨论与分析(宋昕倍 et al., 2024)等文本信息披露有助于提升资本市场定价效率。但与以合规或履责为主要目标的传统披露不同，数据资产信息披露直接体现了企业在数字经济时代对数据要素的整合与应用能力，是企业向外界传递数字化能力、技术实力以及创新潜力的重要信号(苑泽明 et al., 2022)。然而，现有研究更多地关注单个企业数据资产信息披露的独立效应，忽视了市场整体信息披露水平的作用。事实上，投资者获取和处理信息并非没有成本(Blankespoor et al., 2020)。根据有限注意力理论，投资者的认知资源和处理能力自然是有限的，他们不可能同时关注和处理所有可用信息(Hirshleifer et al., 2011; Hirshleifer and Teoh, 2003; Peng and Xiong, 2006)。因此，当市场中企业普遍披露某一类信息时，投资者的注意力资源就有可能会被分散，进而导致单个企业披露的边际价值下降。同时，随着披露企业数量增加，信息同质化现象加剧，单个企业的特质信息也可能被淹没在整体市场噪声中，难以被投资者有效识别和利用，从而降低了信息融入股价的效率。

随着数字经济的蓬勃发展，市场环境和信息披露实践正在发生深刻变化，上市公司披露数据资产的情况正在变得越来越普遍(Liu et al., 2026)。李世刚等(李世刚 et al., 2025)、张娆和张雅婷(张娆 and 张雅婷, 2025)、危雁麟等(危雁麟 et al., 2022)采用了相似的基于机器学习模型分析财务报告的方法，测度中国A股上市公司的数据资产信息披露水平，且李世刚等(李世刚 et al., 2025)、张娆和张雅婷(张娆 and 张雅婷, 2025)均引用了危雁麟等(危雁麟 et al., 2022)的论文作为参考。张娆和张雅婷的测度结果在2007到2022年的年度均值如图1所示；李世刚等(李世刚 et al., 2025)采用了两种指标度量企业数据资产披露水平，它们在2007到2022年的年度均值如图2所示^⑤。他们的研究都表明，随着时间的推移，企业的数据资产信息披露水平明显地提高了。这种变化可能改变数据资产信息披露的信息价值和市场反应：当所有上市公司都披露相似内容的数据资产信息时，投资者便可能降低对单个企业披露的关注度，或者将数据资产信息视为行业常态而非企业特质，从而削弱了其作为差异化信号的价值。这一现象正可以概括为“都披露等于都没披露”。

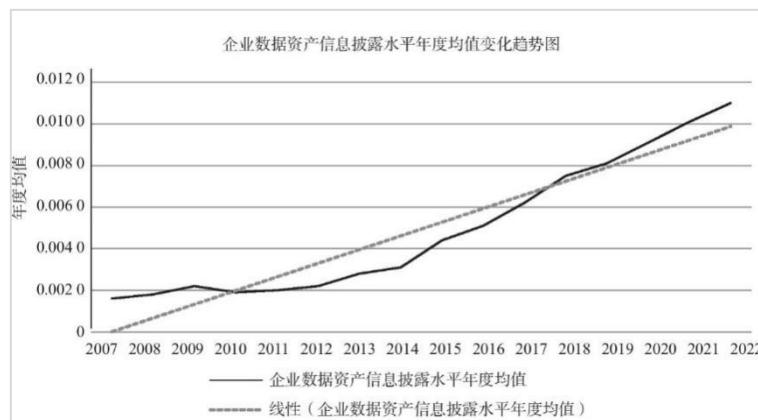


图1 张娆和张雅婷对企业数据资产披露水平测度结果的年度均值

注：此图系笔者直接从张娆和张雅婷的论文中截取，未作任何修改。

^⑤ 张娆和张雅婷(张娆 and 张雅婷, 2025)首先将“数据资源”和“数据资产”作为种子词；再利用Word2Vec神经网络模型获得相似词的词集，保留相似度高于0.5的相似词构成词典；最后基于上市公司年报文本，统计词典中各词汇的出现频率，通过加权求和（权重为与种子词的相似度）再除以总词频计算出企业数据资产信息披露水平指标。李世刚等(李世刚 et al., 2025)的方法和所构建的两种具体指标会在后文详细介绍。

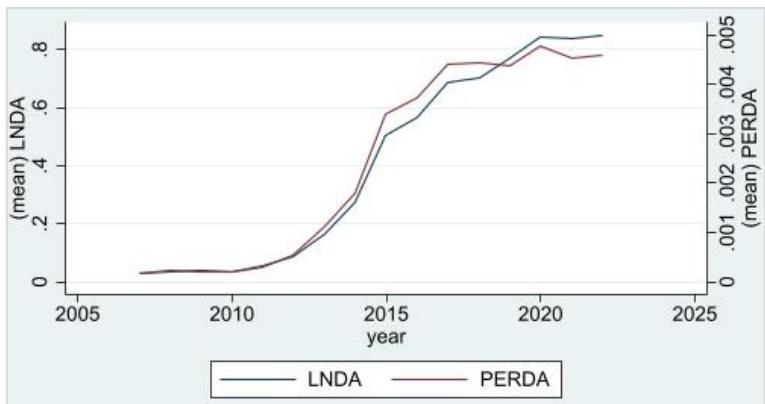


图 2 李世刚等对企业数据资产披露水平测度结果的年度均值

注：此图系笔者自行根据李世刚等(李世刚 et al., 2025)所公布数据绘制而成。

故为了检验这一命题，本文也同样基于 2007-2022 年 A 股上市公司数据，继承李世刚等 (李世刚 et al., 2025) 的方法构建市场整体数据资产信息披露水平指标，考察整体披露质量对单个企业披露效果的“调节作用”。实证结果表明，随着市场整体披露水平的不断提高，单个企业数据资产信息披露对资本市场定价效率的边际影响显著下降，且在近年来整体披露水平较高的阶段，其独立效应已不再显著。机制检验表明，在整体披露水平较低时，企业数据资产信息披露能够通过提升信息供给、增强信息扩散和加快信息融入等途径降低股价同步性；而当整体披露水平较高时，上述传导机制均明显弱化。异质性分析进一步揭示，只有在整体披露水平较低且信息同质化程度较低、信息传播环境较优以及机构投资者注意力未被分散的情境下，数据资产信息披露才能显著降低股价同步性。总体而言，本文从有限注意力约束与动态信息环境的视角，刻画了数据资产信息披露经济后果的情境依赖性与演化特征，把有限注意力理论拓展到了“市场整体信息披露密度对信息价值的稀释效应”角度，并为理解信息披露的经济后果提供了新的经验证据。

笔者认为，本文的边际贡献主要体现在以下方面：

(1) 在理论层面，本文通过考察同质化信息环境如何影响投资者注意力的分散，拓展了有限注意力理论的边界。既有研究更多地关注投资者在异质信息之间的注意力配置问题，强调注意力约束如何影响投资者对单一信息事件的反应。本文将有限注意力理论引入到信息环境演化的动态视角，指出当某类信息在市场中由稀缺转向普遍时，同质化的信息可能会分散投资者的注意力，从而削弱单个企业披露所能产生的边际信息价值。

(2) 在方法论层面，本文突破了现有信息披露文献主要基于企业层面披露指标进行分析的研究范式，引入并刻画了市场整体信息披露水平这一关键状态变量。本文通过构建市场层面的数据资产信息披露指标，并将其纳入门槛回归与分组回归框架，刻画了企业披露效果在不同信息环境下的变化。这一方法使得本文能够区分“披露行为本身的经济后果”与“信息环境变化带来的边际效应衰减”，为解释既有文献中结论差异提供了新的识别路径，也为后续研究分析其他类型信息披露的情境依赖性提供了可推广的经验范式。

(3) 在结论层面，本文提供了关于数据资产信息披露经济后果的一系列的、具有挑战性的经验证据。本文发现，数据资产信息披露仅在整体披露水平较低、信息相对稀缺的阶段，能够显著提升资本市场定价效率；而随着上市公司普遍披露该类信息，其边际定价作用则会减弱甚至消失。这一结果表明，信息披露并非越多越好，其市场功能存在“内生边界”。本文由此揭示了信息披露可能面临的“价值稀释”问题，为理解上市公司信息披露与资本市场定价效率之间的关系提供了新的经验证据与启示。

2 文献评述与研究假说

2.1 投资者的有限注意力

有限注意力理论认为，人类认知资源和信息处理能力存在固有约束，不可能同时关注和处理所有可用信息(Egeth and Kahneman, 1975)。在资本市场中，这一约束导致投资者必须对海量信息进行选择性关注，从而影响其决策行为与市场定价效率(Odean, 1999)。Peng 和 Xiong(Peng and Xiong, 2006)受到“注意力是一种稀缺的认知资源”这一心理学证据的启发，开创性地建立了投资者在学习中的注意力分配模型，并研究了其对资产价格动态的影响。他们发现，有限注意力会导致“类别学习行为”(category-learning behavior)，即投资者倾向于更关注市场和行业层面信息，而忽视企业特有信息。Hirshleifer 等(Hirshleifer et al., 2011)提出一个关于投资者有限注意力的理论模型，解释了为什么市场对盈余总额反应不足，却对盈余的组成部分反应过度。结果表明，投资者对收益信息的忽视导致收益公告后价格漂移，对收益组成部分的忽视则导致应计收益和现金流异常。Lu(Lu, 2022)提出了一个关于财务报告信息披露的理论模型，表明当投资者的信息处理能力有限时，提供更细粒度的详细信息可能反而导致信息过载，使其难以有效提取关键信号，从而降低决策质量；相比之下，汇总性信息虽损失部分潜在信息量，但因结构简洁、认知负荷低，在低注意力条件下可实现更高信息提取效率。Liu(Liu et al., 2023, 2014)等人构建了一个理论模型，说明 IPO 抑价是发行公司吸引注意力有限的投资者的战略手段——通过承诺将抑价股票分配给参与路演的机构投资者，补偿其注意力机会成本。他们发现这种注意力驱动的抑价与媒体关注度呈正相关且不对称，仅在价格上调的 IPO 中显著，并受股份保留比例和扩张程度的影响。

大量实证结果支持了上述理论：Hirshleifer 等(Hirshleifer et al., 2009)发现当市场存在大量竞争性信息时，投资者对盈余公告的反应会减弱。Aboody 等(Aboody et al., 2010)发现，过去表现优异的股票在盈余公告前呈现显著正回报，公告后则出现强烈反转，其原因正是大幅上涨的股票往往会吸引投资者的注意力和投资，尤其是在它们公布收益之前。DellaVigna 和 Pollet(Dellavigna and Pollet, 2009)将投资者对发布在周五财报的反应与其他工作日进行了比较，发现投资者对在周五发布的盈余公告反应更弱。Ferracuti 和 Lind(Ferracuti and Lind, 2025)发现，当许多公司同时发布盈余公告时，投资者会理性地将有限的注意力从公司特定信息转向宏观经济信息，因为集中发布的盈余公告能更有效地预测宏观经济状况。Kwan 等(Kwan et al., 2026)发现机构投资者在高市场波动时期会战略性地将注意力从个股新闻转向宏观新闻，而能够有效进行这种注意力再分配的基金（尤其是对冲基金）能获得更高的回报。王英允等(王英允 et al., 2019)探讨了上市公司管理层如何利用投资者注意力约束进行战略性、选择性信息披露。蒋丹凌等(蒋丹凌 et al., 2023)基于有限注意力假说，利用分析师共同覆盖构建公司关联网络，发现中国 A 股市场中存在显著且持久的关联公司股票动量溢出效应。方军雄等(方军雄 et al., 2018)发现，竞争性信息的存在分散了投资者对分析师研究报告的注意力。当同一天发布的分析师研究报告越多（即竞争性信息越多），投资者对分析师评级报告的即时反应也越弱。

尽管既有研究已从有限注意力约束出发，考察了多源异质信息并存情境下投资者注意力在不同信息之间的竞争与配置，并揭示了竞争性信息冲击如何削弱投资者对特定公司事件或披露信息的反应，但是却（很可能）尚未关注当市场中大量企业同时披露高度相似的信息时，同质化信息本身是否会通过分散投资者注意力，从而削弱单个企业披露的边际信息价值。

2.2 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

资本市场定价效率受到企业信息披露质量与数量的显著影响。根据“信息效率假说”，

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(Morck et al., 2000; 伊志宏 et al., 2019)。现有研究主要从“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机制(李世刚 et al., 2025)。在信息披露水平方面,企业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,提高资本市场定价效率(李世刚 et al., 2025)。大量实证研究从社会责任报告披露(Dhaliwal et al., 2011; 王艳艳 et al., 2014)、关键审计事项披露(Goh et al., 2024; 王木之 and 李丹, 2019)、社交网站信息披露(Bartov et al., 2018; Blankespoor et al., 2014; Guo et al., 2025; 胡军 and 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关键在于是否增加了公司特质信息含量(李世刚 et al., 2025),若上市公司的披露行为更多提供的是市场或行业信息而非其特质信息时,反而会降低资本市场定价效率(王艳艳 et al., 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)。随着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作用(Brown and Tucker, 2011)。研究表明,当文本相似度较低(宋昕倍 et al., 2024; 赵纳晖 and 刘瑾, 2024)、可读性较高(Bai et al., 2019)、语调情感更为真实(吴武清 et al., 2020)时,资本市场定价效率趋于上升。但需要警惕的是,管理层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)。现有研究认为,企业数据资产信息披露作为一项重要的非财务信息,能够通过提高信息供给、增强信息扩散、加快信息融入三个途径显著降低股价同步性,提高资本市场定价效率(李世刚 et al., 2025)。但与此同时,企业对自身数据资产的披露情况也愈发普遍,当市场整体披露水平提高时,单个企业披露的边际价值也可能因投资者有限注意力约束而下降(Peng and Xiong, 2006),出现“都披露等于都没披露”的现象。

(2) 信息环境与资本市场定价效率

资本市场定价效率还受到外部信息环境的深刻影响(宋昕倍 et al., 2024)。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(贺佳 et al., 2025; 陈雨恬 et al., 2024; 马勇 and 喻逸伟, 2023)。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(黄俊 and 郭照蕊, 2014)。现有文献关于分析师作用发挥的观点存在一定分歧: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(朱红军 et al., 2007),可以通过降低信息不对称来提高资本市场定价效率(Piotroski and Roulstone, 2004),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 et al., 2019);另一部分学者则认为,分析师只是信息传递的媒介(胡军 and 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 and 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 and 郭照蕊, 2014)。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 and 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^④(王亚平 et al., 2009)。

已有文献相对充分地验证了单个企业信息披露对资本市场定价效率的影响,但忽视了市场整体信息披露密度这一情境因素。在整体信息密度较高的环境下,单个企业披露的效果也可能被稀释。

^④ 但机构投资者的行为也并非总是理性的,其羊群行为也会降低资本市场定价效率(许年行 et al., 2013)。

2.3 数据资产信息披露的经济后果

现有研究表明，数据资产信息披露对企业价值创造和资本市场效率具有多维度的积极影响。在微观企业层面，数据资产信息披露能够显著缓解融资约束，何瑛等(何瑛 et al., 2024)发现数据资产化通过信号传递效应和抵押品效应，有效缓解了“专精特新”中小企业的融资难题；同时，数据资产信息披露能获得投资者偏好，牛彪和于翔(牛彪 and 于翔, 2024)发现数据资产信息通过降低信息不对称和提高企业透明度，显著降低了权益资本成本。在资本市场层面，李世刚等(李世刚 et al., 2025)基于 2007-2022 年 A 股上市公司数据，系统论证了数据资产信息披露能通过提高信息供给、增强信息扩散和加快信息融入三个机制，显著降低股价同步性，提升资本市场定价效率。此外，危雁麟等(危雁麟 et al., 2022)的研究表明，数据资产信息披露能提高分析师关注度，降低分析师预测偏差，从而增强市场信息效率；苑泽明等(苑泽明 et al., 2022)则发现数据资产信息披露能提升企业价值，且这一效应在机构投资者持股比例较高的企业中更为显著。

然而，信息披露的经济后果并非一成不变，而是高度依赖于披露所处的情境。袁玉等(袁玉 et al., 2025)研究表明，机构投资者能直接有效地监督上市公司管理层行为，其注意力被分散会使管理层在年报中以策略性误导投资者为目的的语调操纵行为增多。随着近年来上市公司披露数据资产信息的程度提升(张娆 and 张雅婷, 2025; 李世刚 et al., 2025)，数据资产信息披露降低股价同步性的效果很可能随之减小。基于李世刚等公开提供的数据，笔者重新检验了数据资产信息披露的经济后果，发现当仅使用最近几年（如 2020-2022 年）数据进行回归，数据资产信息披露对资本市场定价效率的提升作用减弱了；而使用完整样本

（2007-2022 年）时则具备统计显著性。这一特征提示我们，数据资产信息披露的经济后果可能随市场整体披露环境而动态变化。基于有限注意力理论，当市场整体信息披露水平提高时，投资者的注意力资源被分散，单个企业披露的边际价值就可能随之下降。

2.4 研究假说提出

基于前文理论，笔者认为，数据资产信息披露对资本市场定价效率的影响并非静态不变，而是可能受到市场整体信息披露环境的调节。这一观点源于有限注意力理论与信息环境交互作用的内在逻辑。当市场整体数据资产信息披露水平较低时，单个企业披露的数据资产信息具有较高的信息稀缺性和信号价值。投资者能够将有限的注意力资源集中于这些披露信息，有效识别企业特质信息，并通过交易行为将其融入股价，从而降低股价同步性，提高资本市场定价效率。然而，随着数字经济的深入发展，越来越多的企业开始披露数据资产相关信息。当市场整体数据资产信息披露水平较高时，投资者的注意力资源被大量相似信息分散。此时，企业的特质信息就可能被同质化的信息淹没在市场噪声中。在这种情况下，单个企业数据资产信息披露更难被投资者注意到，其对股价同步性的影响也随之减弱。另外从实证现象看，笔者基于李世刚等公开提供的数据发现：当使用完整样本期（2007-2022 年）数据检验时，数据资产信息披露对降低股价同步性的作用确实显著；但当仅使用近几年市场整体数据资产信息披露水平较高的样本（如 2020-2022 年）时，这一关系不再具有统计显著性。这一现象无法用现有理论完全解释，但与“都披露等于都没披露”的直觉高度吻合。

基于此，笔者提出以下假说：

H1：当市场的整体数据资产信息披露水平较低时，上市公司数据资产信息披露水平的提升确有利于降低其股价同步性、提高资本市场定价效率。

H2：当市场的整体数据资产信息披露水平较高时，上市公司数据资产信息披露水平的提升对降低其股价同步性、提高资本市场定价效率的作用将会减弱。

3 实证设计

3.1 数据来源与样本选择

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与连贯性，本文以 2007—2022 年全部 A 股上市公司作为初始样本，并对它们进行如下筛选和处理：(1) 剔除金融类、保险类公司的样本数据；(2) 剔除在上市期间内曾被实施 ST、*ST、PT 或退市等特殊处理的公司；(3) 剔除资产负债率大于一（即资不抵债）的样本；(4) 剔除存在缺失值的样本。此外，为减轻部分极端值对估计结果的潜在干扰，本文对所有连续变量在 1% 和 99% 分位数处进缩尾处理。经上述步骤筛选后，最终获得 30600 个公司一年度观测值。其中，企业数据资产信息披露信息通过对其年报文本进行加工提取；新闻媒体相关数据来源于中国研究数据服务平台（CNRDS）；其余财务与公司治理类变量主要来自国泰安数据库（CSMAR）。^⑤

3.2 计量模型设定

为了检验数据资产信息披露对股价同步性的影响，我们首先构建如下基本模型。

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

上式中， i 代表企业， t 代表年份，被解释变量 SYN 为股价同步性，核心解释变量 DA 为企业数据资产信息披露水平， $Controls$ 表示一系列控制变量， $\varepsilon_{i,t}$ 代表随机扰动项。另外为提升估计的可靠性，模型同时控制公司、年份和省份固定效应 $\sum Firm$ ， $\sum Year$ 和 $\sum Area$ ，并将标准误在公司层面进行聚类调整。

接下来，我们再参考 Hansen(Hansen, 1999) 的研究，构建如下单门槛面板回归模型(2)。首先基于模型(2)进行门槛检验，在确认存在显著门槛效应后，把 2007-2022 年共 16 年的样本分为两组；之后，再分别把两组子样本代入模型(1)中进行回归；另外参考江艇(江艇, 2022)的研究，我们再引入分组虚拟变量、以及分组虚拟变量与核心解释变量的交互项，以判断两组样本的回归系数是否存在显著差异。

$$\begin{aligned} SYN_{i,t+1} = & \alpha_0 + \alpha_{11} DA_{i,t} \cdot I(\overline{DA}_t \leq \lambda) + \alpha_{12} DA_{i,t} \cdot I(\overline{DA}_t > \lambda) \\ & + \gamma Controls_{i,t} + \sum Firm + \sum Year + \sum Area + \varepsilon_{i,t} \end{aligned} \quad (2)$$

其中，门槛变量 \overline{DA}_t 代表核心解释变量企业数据资产信息披露水平在第 t 年的均值，用以衡量市场的整体数据资产信息披露水平； λ 代表特定的门槛值； $I(\cdot)$ 为示性函数，若满足括号中的条件，则 $I(\cdot) = 1$ ，反之， $I(\cdot) = 0$ ；其他变量的含义保持不变。

基于上文的理论分析和假说，我们预期：市场的整体数据资产信息披露水平低组的回归系数 α_1 显著为负，高组的却不显著。

3.3 变量定义

(1) 被解释变量：资本市场定价效率

本文采用股价同步性作为衡量资本市场定价效率的代理变量，若其值越高，则股价更多反映市场或行业层面的共性信息而非公司特质信息，故资本市场的定价效率越低。

参考 Durnev 等人(Durnev et al., 2003)的方法，首先对个股周收益率进行如下回归。其本

^⑤ 为了保证本文实证结果与现有研究的可比性，笔者直接利用李世刚(李世刚 et al., 2025)等在《中国工业经济》官网上所公开提供的数据进行实证研究。下面这段话本质上就是对他们如何搜集与处理数据的复述。

质上就相当于 Fama-MacBeth 回归(Fama and MacBeth, 1973)中的第一阶段，只不过我们的目的并非是得到因子暴露，而是获取其拟合优度 R^2 。

$$R_{i,t} = \alpha_i + \beta_{i,1} R_{m,t} + \beta_{i,2} R_{I,t} + \varepsilon_{i,t} \quad (3)$$

其中，因子 $R_{i,t}$ 代表个股*i*在第*t*周的收益率；因子 $R_{m,t}$ 代表市场指数在第*t*周的收益率， $R_{I,t}$ 代表行业 I 第*t*周收益率； $\varepsilon_{i,t}$ 代表随机扰动项。以上收益率均基于后复权价格，采用总市值加权平均法计算。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，行业分类依据中国证监会 2012 年标准。

上述回归的拟合优度 R^2 即代表了股票*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。为了使拟合优度指标更接近正态分布，进一步采用如下对数转换，得到最终的股票价格同步性指标 SYN 。

$$SYN = \ln\left(\frac{R^2}{1 - R^2}\right) \quad (4)$$

(2) 解释变量及门槛变量：企业数据资产信息披露水平及其年度均值

借鉴苑泽明等(苑泽明 et al., 2022)提出的“种子词+Word2Vec 相似词扩充”方法，本文基于《数据资产管理实践白皮书（5.0 版）》对数据资产的界定，选取“数据资产”和“数据资源”作为种子词，利用 Word2Vec 模型扩展出语义相近的关键词集，并统计相似度最高的前十个关键词在年报中出现的总频次。鉴于该频次呈右偏分布，分别构建两个代理变量：一是对总词频加 1 后取自然对数，记为 LNDA；二是计算相关关键词频次占年报总词频的比例，记为 PERDA。最后，分别计算这两个变量的年度均值作为门槛变量。

(3) 控制变量

借鉴王木之和李丹(2019)、宋昕倍等(宋昕倍 et al., 2024)、赵纳晖和刘瑾(赵纳晖 and 刘瑾, 2024)的研究，控制变量包括：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，并将标准误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 本文主要变量定义

变量名称	变量符号	变量定义
资本市场定价效率	SYN	下一期的股价同步性指标，使用市场和行业收益率根据式(3)和式(4)计算得到
企业数据资产信息披露水平	LNDA	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
	PERDA	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	Size	总资产的自然对数
上市年限	Age	$\ln(\text{样本当前年份}-\text{上市年份}+1)$
产权性质	SOE	企业是否为国有企业，是为 1，否则为 0
是否为“四大”审计	Big4	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	LEV	总负债/总资产
营业收入增长率	Grow	(当年主营业务收入-上年主营业务收入)/上年主营业务收入
总资产净利率	ROA	净利润/总资产
市账比	MB	企业市值/账面价值
股票换手率	Turnover	流通股数计算的年内日均换手率取对数
两权分离度	Seperation	控制权与所有权之间的差值
审计意见	Opinion	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	TOP1	第一大股东的持股比例
两职合一	Dual	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	Manhold	董监高持股数量占总股数量之比
董事会规模	Boardsize	董事会董事数量
独立董事占比	Indir	独立董事人数/董事会总人数
市场化指数	Market	公司注册地所在省份“市场化指数评分”

3.4 变量描述性统计

本文主要变量的描述性统计结果如表 2 所示。结果显示，被解释变量 SYN 的均值约为 -0.4473，标准差约为 1.0517，最小值与最大值分别约为 -3.9233 和 1.8669，这说明不同公司之间的股价同步性差异较大。核心解释变量一 LNDA 的均值约为 0.5274，中位数为 0，最小值与最大值分别为 0 和 3.9703，说明从全体样本上看上市公司的数据资产信息披露水平偏低；核心解释变量二 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%，这说明从全体样本上看，上市公司的数据资产信息披露程度确实不足。当然，上文的图 1 和图 2 已经展示了企业数据资产信息披露程度随时间推移而提高这一趋势。

表 2 本文主要变量描述性统计

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

4 实证结果

4.1 基准回归结果

本文首先对两个门槛变量进行了门槛检验。如表 3 所示, 变量 \overline{LNDA}_t 的门槛值为 0.8402, 变量 \overline{PERDA}_t 的门槛值为 0.0046, 且二者均在 1% 水平上显著。基于这两个门槛值, 我们可以把 2007-2022 年共 16 年的样本按照“市场的整体数据资产信息披露水平”分为高低两组, 其结果也如表 3 所示。以 \overline{LNDA}_t 作为门槛变量时, 2020-2022 年这三年被划分为了“市场整体数据资产信息披露水平高组”, 其余被划分为了低组; 以 \overline{PERDA}_t 作门槛变量时, 2020 年和 2022 年被划分为了“市场整体数据资产信息披露水平高组”, 其余被划分为了低组。

之后, 我们再分别对两组子样本进行回归, 结果如表 4 所示。可以看出, 不论是使用 LNDA 还是 PERDA 作为解释变量, 我们均发现只有市场整体数据资产信息披露水平较低组的回归系数才显著为负, 市场整体数据资产信息披露水平较高组的回归系数都不具有统计显著性, 同时组间系数也均存在统计上的显著差异。这说明, 在早期上市公司还没有普遍披露数据资产时, 提高数据资产披露程度确有利于降低股价同步性, 然而当后期数据资产的披露已经成为一种普遍现象时, 没有证据表明提高数据资产披露程度仍可以降低股价同步性。

表 3 门槛检验结果以及依据门槛值的分组结果

门槛变量	门槛值	F 统计量 (括号内 为 P 值)	按年份分组结果			
			$\overline{LNDA}_t \leq 0.8402$	$\overline{LNDA}_t > 0.8402$	$\overline{PERDA}_t \leq 0.0046$	$\overline{PERDA}_t > 0.0046$
\overline{LNDA}_t	0.8402	77.57*** (0.0050)	2007-2019	2020-2022		
\overline{PERDA}_t	0.0046	69.48*** (0.0025)			2007-2019、 2021	2020、 2022

注: 除有另外说明, 本表和其他表格遵循以下规定: ***、** 和* 分别代表在 1%、5% 和 10% 水平上显著。

表 4 基于门槛检验的分组回归结果

面板 A. 以 $LNDA$ 作被解释变量

	$LNDA_t \leq 0.8402$	$LNDA_t > 0.8402$
	(1)	(2)
$LNDA$	-0.021* (-1.72)	-0.022 (-0.81)
控制变量	是	是
三项固定效应	是	是

组间系数差异检验 P 值	0.002	
观测值	21,017	8,669
调整后的 R ²	0.511	0.686
面板 B. 以 PERDA 作被解释变量		
	$\overline{PERDA}_t \leq 0.0046$	$\overline{PERDA}_t > 0.0046$
	(1)	(2)
PERDA	-2.723*	-0.706
	(-1.74)	(-0.13)
控制变量	是	是
三项固定效应	是	是
组间系数差异检验 P 值	0.001	
观测值	23,669	4,938
调整后的 R ²	0.499	0.737

注: 在面板 A 中, 第(1)列为以 LNDA 作解释变量、市场的整体数据资产信息披露水平低组的回归结果, 第(2)列为以 LNDA 作解释变量、市场的整体数据资产信息披露水平高组的回归结果, 组间系数差异检验 P 值表示了第(1)与(2)列系数有无显著差异。在面板 B 中, 第(1)列为以 PERDA 作解释变量、市场的整体数据资产信息披露水平低组的回归结果, 第(2)列为以 PERDA 作解释变量、市场的整体数据资产信息披露水平高组的回归结果, 组间系数差异检验 P 值表示了第(1)与(2)列系数有无显著差异。由于控制了公司、年份和省份三项固定效应, 故样本量可能有所损失。除有另外说明, 本表和其他回归结果的展示表格遵循以下规定: ***、** 和 * 分别代表在 1%、5% 和 10% 水平上显著; 括号内为 t 值, 标准误经公司层面聚类调整。

4.2 稳健性检验

(1) 双向聚类调整

为控制面板数据中可能存在的公司截面相关与时间序列自相关问题, 本文在基准回归的基础上采用了公司层面与年度层面的双向聚类调整, 结果如表 5 所示。可以看出, 进行双向聚类调整后, 核心结论依然成立。具体表现为: 在市场平均披露水平低组, 两项核心解释变量的回归系数显著为负; 而市场披露水平高组的回归系数不具备显著性; 同时, 两组的回归系数存在显著差异。

表 5 双向聚类调整回归结果

面板 A. 以 LNDA 作被解释变量		
	$\overline{LNDA}_t \leq 0.8402$	$\overline{LNDA}_t > 0.8402$
	(1)	(2)
LNDA	-0.0213*	-0.022
	(-2.09)	(-1.02)
控制变量	是	是
三项固定效应	是	是
组间系数差异检验 P 值	0.041	
观测值	21017	8669
调整后的 R ²	0.4251	0.4984
面板 B. 以 PERDA 作被解释变量		
	$\overline{PERDA}_t \leq 0.0046$	$\overline{PERDA}_t > 0.0046$
	(1)	(2)

<i>PERDA</i>	-2.723*	-0.706
	(-2.44)	(-0.13)
控制变量	是	是
三项固定效应	是	是
组间系数差异检验 P 值	0.018	
观测值	23,669	4,938
调整后的 R ²	0.4189	0.4633

(2) 剔除非高科技企业样本

相比一般企业，高科技企业的数据资产更加丰富，也更容易形成可以披露在企业年报中的信息。因此，本文聚焦高科技企业样本做进一步的检验。借鉴彭红星和毛新述(彭红星 and 毛新述, 2017)的做法，将制造业中的 15 个大类，信息传输、软件和信息技术服务业中的 3 个大类和科学研究与技术服务业归类为高科技企业做单独回归，检验结果如表 6 所示。可以看出，在市场平均披露水平较低的阶段，LNDA 和 PERDA 的系数仍然显著为负，而随着市场披露水平升高，这种显著性不再存在。这表明在高科技企业中，企业披露数据资产信息对于提高资本市场定价效率的影响效果依然会受到行业平均披露水平的限制。

表 6 剔除非高科技企业回归结果

面板 A. 以 <i>LNDA</i> 作被解释变量		
	$\overline{LNDA}_t \leq 0.8402$	$\overline{LNDA}_t > 0.8402$
	(1)	(2)
<i>LNDA</i>	-0.025** (-2.37)	-0.012 (-0.92)
控制变量	是	是
三项固定效应	是	是
组间系数差异检验 P 值	0.002	
观测值	26674	18071
调整后的 R ²	0.4884	0.4642
面板 B. 以 <i>PERDA</i> 作被解释变量		
	$\overline{PERDA}_t \leq 0.0046$	$\overline{PERDA}_t > 0.0046$
	(1)	(2)
<i>PERDA</i>	-3.559** (-2.40)	-1.558 (-0.86)
控制变量	是	是
三项固定效应	是	是
组间系数差异检验 P 值	0.001	
观测值	26674	18071
调整后的 R ²	0.4884	0.4642

4.3 内生性处理

(1) 考虑时间趋势变化

为排除不同地区或行业因数据经济发展步调不一而存在的固有趋势对结果的干扰，本文在模型中分别加入了地区与年份 (Area×Year)、行业与年份 (Ind×Year) 的交互固定效应

并重新回归。回归结果显示，在控制了可能随时间和空间演变的遗漏因素后，本文的核心结论依然稳健。

表 7 考虑时间趋势变化回归结果

面板 A. 以 $LNDA$ 作被解释变量				
	$\overline{LNDA}_t \leq 0.8402$		$\overline{LNDA}_t > 0.8402$	
	(1)	(2)	(3)	(4)
$LNDA$	-0.025** (-2.00)	-0.043*** (-3.21)	-0.013 (-0.48)	-0.019 (-0.74)
控制变量	是	是	是	是
三项固定效应	是	是	是	是
省份 \times 年份固定效应	是	否	是	否
行业 \times 年份固定效应	否	是	否	是
组间系数差异检验 P 值	0.002	0.657	-	-
观测值	21017	21017	8669	8669
调整后的 R ²	0.428	0.441	0.503	0.529

面板 B. 以 $PERDA$ 作被解释变量				
	$\overline{PERDA}_t \leq 0.0046$		$\overline{PERDA}_t > 0.0046$	
	(1)	(2)	(3)	(4)
$PERDA$	-3.150* (-1.88)	-5.280*** (-2.93)	-2.476 (-0.61)	-3.471 (-0.89)
控制变量	是	是	是	是
三项固定效应	是	是	是	是
省份 \times 年份固定效应	是	否	是	否
行业 \times 年份固定效应	否	是	否	是
组间系数差异检验 P 值	0.001	0.076	-	-
观测值	21017	21017	8669	8669
调整后的 R ²	0.428	0.441	0.503	0.529

(2) 熵平衡检验

为缓解处理组与对照组之间可观测特征的系统性差异带来的选择偏差，本文采用熵平衡法对样本进行重新加权，使两组公司在所有控制变量特征上达到平衡。在使用熵平衡权重消除可观测协变量的差异后，分组回归的结果模式与基准结论一致，即在构建了一个更可比的分析样本后，平均披露水平“低组有效、高组无效”的动态模式依然稳健。

表 8 熵平衡检验结果

面板 A. 以 $LNDA$ 作被解释变量				
	$\overline{LNDA}_t \leq 0.8402$		$\overline{LNDA}_t > 0.8402$	
	(1)	(2)	(3)	(4)
$LNDA$	-0.025* (-1.89)	-0.027 (-0.99)	-	-
控制变量	是	是	是	是
三项固定效应	是	是	是	是

组间系数差异检验 P 值	<0.001	
观测值	21017	8669
调整后的 R ²	0.442	0.512
面板 B. 以 PERDA 作被解释变量		
	$\overline{PERDA}_t \leq 0.0046$	$\overline{PERDA}_t > 0.0046$
	(1)	(2)
PERDA	-3.027*	-3.773
	(-1.73)	(-0.91)
控制变量	是	是
三项固定效应	是	是
组间系数差异检验 P 值	<0.001	
观测值	21017	8669
调整后的 R ²	0.442	0.512

5 机制再检验

根据李世刚等(李世刚 et al., 2025)的观点，数据资产相关信息的披露有助于外部市场主体更清晰地把握企业数据资产的发展动态。这类信息因其差异化特征，能够有效吸引投资者注意，从而扩充公司特质信息的供给；同时，作为企业数字技术实力与未来增长潜力的重要表征，数据资产信息极易引发新闻媒体等信息中介的积极关注与报道，进而产生显著的信息扩散效应，促进特质信息在市场中的流动；此外，机构投资者凭借其广泛的信息渠道与专业的研判能力，能够深入挖掘此类信息，并加速推动其融入股价。因此，上市公司的数据资产信息披露能够通过提高信息供给、增强信息扩散和加快信息融入三个机制来降低股价同步性、提高资本市场定价效率。然而，上文的基准回归结果表明，数据资产信息披露降低股价同步性的作用会受到市场整体披露水平的影响。那么，这是否同样体现在这三个具体机制上？本节将对这三个机制进行再检验。

5.1 提高信息供给

既有研究认为数据资产信息披露能够通过增加公司特质信息供给来提升资本市场定价效率(危雁麟 et al., 2022; 李世刚 et al., 2025; 牛彪 and 于翔, 2024)。在传统财务报表难以充分反映数据资产价值的背景下，年报文本中的数据资产披露为投资者提供了重要的增量信息(Liu et al., 2025)，使其更全面地了解企业在数字技术应用、数据治理能力及未来成长性方面的状况，从而缓解信息不对称。与此同时，数据资产的建设与披露过程有助于优化企业内部信息流转，提高会计信息质量，降低投资者的信息甄别成本。作为重要的信息中介，分析师能够对企业披露的数据资产信息进行深度解读，并通过研究报告向市场传递，从而减少其对宏观和行业信息的依赖，提高预测的针对性和准确性。因此数据资产信息披露能够显著提升分析师关注度、降低预测偏差，进而提高股价中公司特质信息的含量。然而，分析师同样受到有限注意力约束。当市场整体数据资产信息披露水平不断提高、披露内容趋于同质化时，分析师的认知资源可能被大量相似信息分散，难以对单个企业进行深入挖掘。在这种情形下，单个企业增加数据资产信息披露对提升分析师关注度和改善预测质量的边际作用就可能随之减弱。

因此，我们对数据资产信息披露通过提高信息供给降低股价同步性这一机制进行再检验。

具体而言，我们借鉴危雁麟等(危雁麟 et al., 2022)的做法，使用在一年内对某企业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），并以这两个指标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，再基于上文门槛检验的分组结果，用两组样本分别进行回归，结果如表 9 所示。可以看出，当市场的整体数据资产信息披露水平较低时，企业提高数据资产信息披露程度确实有利于提升分析师关注度、降低分析师预测偏差，但是当市场的整体数据资产信息披露水平较高时，没有证据表明企业提高数据资产信息披露程度仍然有利于提升分析师关注度、降低分析师预测偏差。

表 9 提高信息供给机制检验结果

面板 A. 以 $LNDA$ 作被解释变量

	$LNDA_t \leq 0.8402$		$LNDA_t > 0.8402$	
	(1)	(2)	(3)	(4)
	<i>Follow</i>	<i>Bias</i>	<i>Follow</i>	<i>Bias</i>
$LNDA$	0.033** (2.31)	-0.0005*** (-3.12)	-0.011 (-0.64)	0.0003 (0.58)
控制变量	是	是	是	是
三项固定效应	是	是	是	是
组间系数差异检验 P 值	0.010	0.655	-	-
观测值	21,017	12,555	8,669	3,722
调整后的 R ²	0.748	0.383	0.900	0.635

面板 B. 以 $PERDA$ 作被解释变量

	$PERDA_t \leq 0.0046$		$PERDA_t > 0.0046$	
	(1)	(2)	(3)	(4)
	<i>Follow</i>	<i>Bias</i>	<i>Follow</i>	<i>Bias</i>
$PERDA$	4.177** (2.32)	-0.059*** (-3.23)	0.671 (0.19)	0.143 (1.41)
控制变量	是	是	是	是
三项固定效应	是	是	是	是
组间系数差异检验 P 值	0.107	0.946	-	-
观测值	23,669	13,715	4,938	1,884
调整后的 R ²	0.698	0.235	0.780	0.360

注：在面板 A 中，第(1)列为以 Follow 作被解释变量、以 $LNDA$ 作解释变量、市场整体数据资产信息披露水平低组的回归结果，第(2)列为以 Bias 作被解释变量、以 $LNDA$ 作解释变量、市场整体数据资产信息披露水平低组的回归结果，第(3)列为以 Follow 作被解释变量、以 $LNDA$ 作解释变量、市场整体数据资产信息披露水平高组的回归结果，第(4)列为以 Bias 作被解释变量、以 $LNDA$ 作解释变量、市场整体数据资产信息披露水平高组的回归结果；第(1)列中的组间系数差异检验 P 值表示了第(1)与(3)列系数有无显著差异，第(2)列中的组间系数差异检验 P 值表示了第(2)与(4)列系数有无显著差异。在面板 B 中，第(1)列为以 Follow 作被解释变量、以 $PERDA$ 作解释变量、市场整体数据资产信息披露水平低组的回归结果，第(2)列为以 Bias 作被解释变量、以 $PERDA$ 作解释变量、市场整体数据资产信息披露水平低组的回归结果，第(3)列为以 Follow 作被解释变量、以 $PERDA$ 作解释变量、市场整体数据资产信息披露水平高组的回归结果，第(4)列为以 Bias 作被解释变量、以 $PERDA$ 作解释变量、市场整体数据资产信息披露水平高组的回归结果；第(1)列中的组间系数差异检验 P 值表示了第(1)与(3)列系数有无显著差异，第(2)列中的组间系数差异检验 P 值表示了第(2)

与(4)列系数有无显著差异。

5.2 增强信息扩散

数据资产信息披露还可能通过增强信息扩散来降低股价同步性。既有研究认为相较于传统财务信息，数据资产信息具有较强的新颖性和话题性，更容易引起分析师与媒体的关注。其往往被市场视为企业数字化能力和未来成长潜力的重要信号，因而更可能通过研究报告和新闻报道等渠道在资本市场中被广泛传播(李世刚 et al., 2025)。在这一过程中，原本仅存在于年报文本中的公司特质信息得以向更广泛的投资者群体扩散，加快其融入资产价格。在以个人投资者为主的中国资本市场中，这一机制尤为关键。个人投资者更多地依赖媒体报道和二手信息进行投资决策，媒体对数据资产信息的再加工有助于降低信息理解成本，使更多投资者识别企业在数据资源配置和业务模式方面的差异，从而减少股价对市场和行业信息的依赖，提高公司特质信息在价格中的权重。然而，信息扩散机制的作用同样受整体信息环境影响。当市场整体数据资产信息披露水平较低时，相关信息具有较高稀缺性和辨识度，更容易获得持续关注并实现有效扩散；而随着披露企业数量增加、信息内容趋于同质化，媒体与投资者的边际关注度可能下降，企业披露的信息更易被视为行业共性特征而非公司特质。在有限注意力约束下，信息扩散对股价同步性的改善作用也有可能随之减弱。

因此，我们对数据资产信息披露通过增强信息扩散降低股价同步性这一机制进行再检验。具体而言，我们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（News_cont1）和财经报刊与网络新闻总报道数量（News_cont2）衡量媒体对上市公司数据资产信息扩散的影响，再基于上文门槛检验的分组结果，用两组样本分别进行回归，结果如表 10 所示。可以看出，当市场的整体数据资产信息披露水平较低时，企业提高数据资产信息披露程度确实有利于吸引新闻媒体的关注，但是当市场的整体数据资产信息披露水平较高时，没有证据表明企业提高数据资产信息披露程度仍然有利于吸引新闻媒体的关注。

表 10 增强信息扩散机制检验结果

面板 A. 以 $LNDA$ 作被解释变量

	$\overline{LNDA}_t \leq 0.8402$		$\overline{LNDA}_t > 0.8402$	
	(1)	(2)	(3)	(4)
	News_cont1	News_cont2	News_cont1	News_cont2
$LNDA$	11.953*** (2.60)	15.315** (2.56)	3.019 (0.79)	4.019 (0.83)
控制变量	是	是	是	是
三项固定效应	是	是	是	是
组间系数差异检验 P 值	0.001	0.003	-	-
观测值	20,987	8,663	20,825	8,006
调整后的 R ²	0.779	0.932	0.816	0.950

面板 B. 以 $PERDA$ 作被解释变量

	$\overline{PERDA}_t \leq 0.0046$		$\overline{PERDA}_t > 0.0046$	
	(1)	(2)	(3)	(4)
	News_cont1	News_cont2	News_cont1	News_cont2
$PERDA$	1,404.331** (2.41)	1,680.394** (2.20)	700.075 (1.05)	832.296 (0.95)
控制变量	是	是	是	是

三项固定效应	是	是	是	是
组间系数差异检验 P 值	0.013	0.028	-	-
观测值	23,635	4,934	23,299	4,416
调整后的 R ²	0.773	0.948	0.807	0.959

注: 在面板 A 中, 第(1)列为以 News_cont1 作被解释变量、以 LNDA 作解释变量、市场整体数据资产信息披露水平低组的回归结果, 第(2)列为以 News_cont2 作被解释变量、以 LNDA 作解释变量、市场整体数据资产信息披露水平低组的回归结果, 第(3)列为以 News_cont1 作被解释变量、以 LNDA 作解释变量、市场整体数据资产信息披露水平高组的回归结果, 第(4)列为以 News_cont2 作被解释变量、以 LNDA 作解释变量、市场整体数据资产信息披露水平高组的回归结果; 第(1)列中的组间系数差异检验 P 值表示了第(1)与(3)列系数有无显著差异, 第(2)列中的组间系数差异检验 P 值表示了第(2)与(4)列系数有无显著差异。在面板 B 中, 第(1)列为以 News_cont1 作被解释变量、以 PERDA 作解释变量、市场整体数据资产信息披露水平低组的回归结果, 第(2)列为以 News_cont2 作被解释变量、以 PERDA 作解释变量、市场整体数据资产信息披露水平低组的回归结果, 第(3)列为以 News_cont1 作被解释变量、以 PERDA 作解释变量、市场整体数据资产信息披露水平高组的回归结果, 第(4)列为以 News_cont2 作被解释变量、以 PERDA 作解释变量、市场整体数据资产信息披露水平高组的回归结果; 第(1)列中的组间系数差异检验 P 值表示了第(1)与(3)列系数有无显著差异, 第(2)列中的组间系数差异检验 P 值表示了第(2)与(4)列系数有无显著差异。

5.3 加快信息融入

除提高信息供给和增强信息扩散外, 数据资产信息披露还可能通过加快信息融入过程提升资本市场定价效率。信息融入强调公司特质信息从披露到反映在股价中的速度和充分程度。在信息披露更加充分且结构清晰的情况下, 投资者能够更快识别信息含义并据此调整交易行为, 从而缩短信息进入价格的时间(李世刚 et al., 2025)。相较于隐含于经营结果中的信息, 数据资产信息以更为直接和前瞻的方式揭示企业在数字化投入、数据治理能力及未来战略方向方面的特征, 有助于投资者提前形成预期并在交易中加以反映(McCAHERY et al., 2016)。从市场运行角度看, 系统化的数据资产信息披露有助于降低不确定性和噪声交易比例, 使投资者决策更多基于基本面而非市场共振, 从而提高价格发现效率并降低股价同步性。然而, 该机制同样受信息环境的制约。当市场的整体数据资产信息披露水平较低时, 相关信息具有较强的增量特征, 投资者更容易将其视为有价值的新信息并迅速纳入定价; 而随着披露行为普遍化、信息内容趋于同质化, 其边际信息量可能下降, 投资者对相关披露的反应趋于理性, 从而削弱信息融入速度的提升效应。

因此, 我们对数据资产信息披露通过加快信息融入降低股价同步性这一机制进行再检验。具体而言, 我们使用机构投资者持股比例 (Inshold) 作为衡量机构投资者加快信息融入效应的代理指标, 再基于上文门槛检验的分组结果, 用两组样本分别进行回归, 结果如表 11 所示。可以看出, 当市场的整体数据资产信息披露水平较低时, 企业提高数据资产信息披露程度确实有利于提升机构投资者的关注度, 并促使其增加持股比例; 但是当市场的整体数据资产信息披露水平较高时, 没有证据表明企业提高数据资产信息披露程度仍然有利于提升机构投资者的关注度, 促使其增加持股比例。

表 11 加快信息融入机制检验结果

面板 A. 以 LNDA 作被解释变量

	$\overline{LNDA}_t \leq 0.8402$	$\overline{LNDA}_t > 0.8402$
	(1)	(2)
	Inshold	Inshold

<i>LNDA</i>	1.532 (1.00)	3.660 (1.58)
控制变量	是	是
三项固定效应	是	是
组间系数差异检验 P 值	0.321	
观测值	21,017	8,669
调整后的 R ²	0.547	0.789
面板 B. 以 <i>PERDA</i> 作被解释变量		
	<i>PERDA_t</i> ≤ 0.0046 (1)	<i>PERDA_t</i> > 0.0046 (2)
	<i>Inshold</i>	<i>Inshold</i>
<i>PERDA</i>	309.034 (1.64)	728.750 (1.31)
控制变量	是	是
三项固定效应	是	是
组间系数差异检验 P 值	0.893	
观测值	23,669	4,938
调整后的 R ²	0.519	0.711

注: 在面板 A 中, 第(1)列为以 *Inshold* 作被解释变量、以 *LNDA* 作解释变量、市场的整体数据资产信息披露水平低组的回归结果, 第(2)列为以 *Inshold* 作被解释变量、以 *LNDA* 作解释变量、市场的整体数据资产信息披露水平高组的回归结果; 组间系数差异检验 P 值代表了第(1)与(2)列系数有无显著差异。在面板 B 中, 第(1)列为以 *Inshold* 作被解释变量、以 *PERDA* 作解释变量、市场的整体数据资产信息披露水平低组的回归结果, 第(2)列为以 *Inshold* 作被解释变量、以 *PERDA* 作解释变量、市场的整体数据资产信息披露水平高组的回归结果; 组间系数差异检验 P 值代表了第(1)与(2)列系数有无显著差异。

6 异质性分析

前文的基准回归以及机制检验结果表明,企业数据资产信息披露对资本市场定价效率的影响并非在所有情境下同样成立,而是依赖于市场整体数据资产信息披露水平。当市场整体披露水平较低时,单个企业披露的数据资产信息确实能够通过提高信息供给、增强信息扩散和加快信息融入等途径降低股价同步性;而当市场整体披露水平较高时,单个企业披露的边际信息价值显著下降,其对资本市场定价效率的影响亦随之减弱。然而,不同企业所处的信息环境、信息生产方式以及信息使用主体的特征存在显著差异,即便在相同的整体披露水平阶段,数据资产信息披露的经济后果也可能表现出明显的异质性。因此,本文在前述“整体披露水平高和低”分组的基础上,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异质性分析。文本信息相似度指标能够衡量企业在不同年度间信息披露内容的差异程度,其有助于分析当企业特征信息供给存在差异时,数据资产相关披露所产生的差异化影响;区域数字化发展程度则决定了信息传播的速度与广度,可以从信息传播视角展示数据资产披露内容在不同地域下的扩散效果;而机构投资者分心程度指标则量化了专业投资者关注焦点的集中程度,其反映了在信息融入机制过程中,数据资产相关信息是否能够被机构投资者充分捕捉,并最终通过市场交易行为反映到股票价格中。通过这三项异质性分析,可有效辨析数据资产信息

披露在不同情境下对资本市场定价效率影响的差异化表现，进一步验证和丰富前文的结论。

6.1 文本信息相似度

管理层讨论与分析(MD&A)是企业年报中最重要、信息含量最多的一部分之一(王雄元 et al., 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有效满足信息使用者对相关性、可靠性和前瞻性信息的需求(Clarkson et al., 1999)。文本相似度是企业年报的重要特征之一，相邻年度 MD&A 文本的余弦相似度可间接反映企业当期特质信息含量的高低(Brown and Tucker, 2011)。现有研究认为，数据资产信息披露可以通过补充公司特质信息、缓解信息不对称进而提升资本市场定价效率(李世刚 et al., 2025)。然而根据前文分析，当企业普遍披露数据资产时，单个企业的披露更难被注意到。此时，即使企业的 MD&A 文本相对于上期的相似度较高，数据资产信息披露补充特质信息的功能或许也会减弱。因此我们预期，当整体数据资产信息披露水平较低时，如果企业的 MD&A 文本与上期的相似度较高，数据资产信息披露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显著的作用，但当整体数据资产信息披露水平较高时，不论企业的 MD&A 文本与上期的相似度较高还是较低，这种作用都不会再显著。

故为检验不同的整体披露水平下 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(王雄元 et al., 2018)的做法，利用企业年报中 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把样本分为文本相似度高组和低组，并在不同整体披露水平下分别进行回归分析，结果如表 12 所示。可以看出，只有在整体数据资产信息披露水平较低且企业 MD&A 文本相对上期相似度较高时，回归系数才具有统计显著性。该结果表明，在市场整体数据资产信息披露水平较低时，数据资产信息披露确实能够提供更多的特质信息，从而更有效地降低股价同步性，提高资本市场定价效率；然而当市场整体数据资产信息披露水平较高时，这种作用随之减弱。

表 12 文本信息相似度异质性分析

面板 A. 以 $LNDA$ 作被解释变量

	$\overline{LNDA}_t \leq 0.8402$		$\overline{LNDA}_t > 0.8402$	
	(1)	(2)	(3)	(4)
	文本相似 度高	文本相似 度低	文本相似 度高	文本相似 度低
$LNDA$	-0.036* (-1.84)	-0.050 (-1.34)	-0.046 (-1.31)	-0.046 (-0.34)
控制变量	是	是	是	是
三项固定效应	是	是	是	是
观测值	8866	3320	5904	494
调整后的 R ²	0.391	0.374	0.460	0.594

面板 B. 以 $PERDA$ 作被解释变量

	$\overline{PERDA}_t \leq 0.0046$		$\overline{PERDA}_t > 0.0046$	
	(1)	(2)	(3)	(4)
	文本相似 度高	文本相似 度低	文本相似 度高	文本相似 度低
$PERDA$	-4.521* (-1.81)	-1.411 (-0.58)	2.288 (0.33)	2.365 (0.34)
控制变量	是	是	是	是

三项固定效应	是	是	是	是
观测值	10,078	9,963	3,047	3,185
调整后的 R ²	0.568	0.558	0.741	0.730

6.2 地区信息化水平

地区信息化水平是影响信息传播效率与投资者信息获取能力的关键外部环境因素。投资者能否及时获取并处理股市相关信息，对其投资决策具有重要影响。从信息渠道视角看，计算机与互联网等信息技术的普及不仅降低了交易成本，还显著减少了投资者获取公司特质信息的障碍，从而提升其市场参与度与决策质量(Goldstein et al., 2023; 郭士祺 and 梁平汉, 2014)。现有研究认为，数据资产信息披露可通过吸引新闻媒体等信息中介关注，将特质信息扩散至更广泛的投资者群体，进而提升资本市场定价效率；而地区信息化水平作为信息传播基础设施的表征，直接决定了媒体所传播的信息能否高效触达目标投资者(李世刚 et al., 2025)。然而根据前文分析，当企业普遍披露数据资产时，投资者的注意力会被分散，即使在信息化水平较高的地区，单个企业披露的信息获取也可能更难被有效识别与吸收。因此我们预期，当整体数据资产信息披露水平较低时，如果地区信息化水平较高，数据资产信息披露确能通过高效的信息扩散渠道显著降低股价同步性；而当整体披露水平较高时，不论地区信息化水平高低，这种作用都不会再显著。

故为检验不同的整体披露水平下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把样本企业划分为地区信息化水平高组和低组，并在不同的整体披露水平下分别进行回归分析^⑥，结果如表 13 所示。可以看出，只有在整体数据资产信息披露水平较低且地区信息化水平较高时，回归系数才具有统计显著性。该结果表明，在市场整体数据资产信息披露水平较低时，数据资产信息披露确能通过高效的信息扩散渠道显著降低股价同步性，提高资本市场定价效率；然而当市场整体数据资产信息披露水平较高时，这种作用随之减弱。

表 13 地区信息化水平异质性影响

	面板 A. 以 $LNDA$ 作被解释变量			
	$LNDA_t \leq 0.8402$		$LNDA_t > 0.8402$	
	(1)	(2)	(3)	(4)
	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
$LNDA$	-0.043** (-2.27)	-0.007 (-0.37)	-0.003 (-0.05)	0.025 (0.53)
控制变量	是	是	是	是
三项固定效应	是	是	是	是
观测值	10,078	9,963	3,047	3,185
调整后的 R ²	0.568	0.558	0.741	0.730
面板 B. 以 $PERDA$ 作被解释变量				
	$PERDA_t \leq 0.0046$		$PERDA_t > 0.0046$	

^⑥ 在以 $PERDA$ 作解释变量时，在市场整体数据资产信息披露水平较高的年份中（2020 和 2022 年），地区信息化水平较低的样本量不足以获取回归结果。故笔者在此处采用“折中方法”，将 2021 年的样本也加入回归。

	(1)	(2)	(3)	(4)
	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>PERDA</i>	-3.173* (-1.74)	-5.567 (-1.27)	-1.182 (-0.15)	-3.169 (-0.78)
控制变量	是	是	是	是
三项固定效应	是	是	是	是
观测值	18376	4077	1998	8669
调整后的 R ²	0.432	0.369	0.467	0.499

6.3 机构投资者分心程度

机构投资者作为资本市场的重要参与者，在信息融入过程中发挥着关键作用。相较于个人投资者，机构投资者凭借其专业的信息挖掘能力与大规模交易优势，能够深入识别企业特质信息并迅速将其反映在股价中，从而提升资本市场定价效率(侯宇 and 叶冬艳, 2008)。现有研究认为，数据资产信息披露可通过吸引机构投资者关注，进而加快特质信息融入股价(李世刚 et al., 2025)。然而根据前文分析，机构投资者的认知资源同样存在约束(Kwan et al., 2026)，当市场整体数据资产信息披露水平较高时，机构投资者有限的注意力也会被分散(袁玉 et al., 2025)，即使其分心程度较低，也可能会难以有效识别单个企业的数据资产特质信息。因此我们预期，当整体披露水平较低时，如果机构投资者分心程度较低，数据资产信息披露能通过获得机构投资者的关注显著降低股价同步性；而在整体披露水平较高时，这种作用也会随之减弱。

故为检验不同整体披露水平下机构投资者分心程度的异质性影响，本文借鉴袁玉等(袁玉 et al., 2025)的做法构建企业层面的机构投资者分心指标，并依据行业—年度中位数将样本划分为机构投资者分心程度高组与低组，并分别在不同的整体披露水平下分别进行回归分析，结果如表 14 所示。可以看出，只有在整体数据资产信息披露水平较低且机构投资者分心程度低时，回归系数才具有统计显著性。该结果表明，在市场整体数据资产信息披露水平较低时，数据资产信息披露确能通过吸引机构投资者关注显著降低股价同步性，提高资本市场定价效率；然而当市场整体数据资产信息披露水平较高时，这种作用随之减弱。

表 14 机构投资者注意力异质性分析

面板 A. 以 <i>LNDA</i> 作被解释变量				
	$\overline{LNDA}_t \leq 0.8402$		$\overline{LNDA}_t > 0.8402$	
	(1)	(2)	(3)	(4)
	高	低	高	低
<i>LNDA</i>	-0.004 (-0.25)	-0.046** (-2.24)	-0.050 (-1.22)	-0.026 (-0.53)
控制变量	是	是	是	是
三项固定效应	是	是	是	是
观测值	10,134	9,970	3,621	3,350
调整后的 R ²	0.552	0.537	0.717	0.717
面板 B. 以 <i>PERDA</i> 作被解释变量				
	$\overline{PERDA}_t \leq 0.0046$		$\overline{PERDA}_t > 0.0046$	
	(1)	(2)	(3)	(4)

	高	低	高	低
PERDA	-1.386 (-0.60)	-5.322* (-1.96)	-6.153 (-1.04)	-10.222 (-1.36)
控制变量	是	是	是	是
三项固定效应	是	是	是	是
观测值	10,134	9,970	3,621	3,350
调整后的 R ²	0.552	0.537	0.717	0.717

7 结论

本文立足于有限注意力理论与动态信息环境视角,利用2007-2022年中国A股上市公司数据,重新审视并挑战了关于数据资产信息披露经济后果的现有认知。研究发现,数据资产信息披露对资本市场定价效率的提升作用并非一成不变,而是随着市场整体披露水平的提升而减弱。具体而言,在市场整体披露水平较低的时期,数据资产信息披露确实能够通过提高信息供给质量、增强媒体传播效应以及加速机构投资者信息融入,有效降低股价同步性;然而,随着上市公司普遍开始披露数据资产,同质化的信息分散了投资者们有限的注意力,导致上述传导机制失效,最终呈现出“都披露等于都没披露”的现象。进一步的异质性分析揭示,仅在市场整体数据资产信息披露水平较低,且企业披露的特质信息相对较少、信息传播环境优越或机构投资者注意力未被过度分散的情境下,披露行为才能维持其正向效应。上述发现不仅拓展了有限注意力理论在信息环境维度的边界,为其他信息披露相关研究提供了参考,也为政策实践提供了深刻启示:监管部门在推动数据要素入表与披露时,不应单纯追求披露数量的增长,而应致力于建立差异化、实质性的披露规范,抑制“蹭热点”式的同质化披露;同时,企业应摒弃模板化的披露模式,着力挖掘并呈现具有独特价值的数据资产信息,以在高度拥挤的信息市场中通过实质性差异化重获投资者关注,从而真正实现数据要素价值的有效传递与资本市场的效率提升。

参考文献

- Aboody, D., Lehavy, R., Trueman, B., 2010. Limited attention and the earnings announcement returns of past stock market winners. *Rev Account Stud* 15, 317–344.
<https://doi.org/10.1007/s11142-009-9104-9>
- Bai, X., Dong, Y., Hu, N., 2019. Financial report readability and stock return synchronicity. *Applied Economics* 51, 346–363. <https://doi.org/10.1080/00036846.2018.1495824>
- Bartov, E., Faurel, L., Mohanram, P.S., 2018. Can Twitter Help Predict Firm-Level Earnings and Stock Returns? *The Accounting Review* 93, 25–57. <https://doi.org/10.2308/accr-51865>
- Blankespoor, E., deHaan, E., Marinovic, I., 2020. Disclosure processing costs, investors' information choice, and equity market outcomes: A review. *Journal of Accounting and Economics* 70, 101344. <https://doi.org/10.1016/j.jacceco.2020.101344>
- Blankespoor, E., Miller, G.S., White, H.D., 2014. The Role of Dissemination in Market Liquidity: Evidence from Firms' Use of Twitter™. *The Accounting Review* 89, 79–112.
<https://doi.org/10.2308/accr-50576>
- Bochkay, K., Brown, S.V., Leone, A.J., Tucker, J.W., 2023. Textual Analysis in Accounting:

- What's Next? *Contemporary Accounting Research* 40, 765–805.
<https://doi.org/10.1111/1911-3846.12825>
- Brown, S.V., Tucker, J.W., 2011. Large-Sample Evidence on Firms' Year-over-Year MD&A Modifications. *J of Accounting Research* 49, 309–346.
<https://doi.org/10.1111/j.1475-679X.2010.00396.x>
- Chan, K., Hameed, A., 2006. Stock price synchronicity and analyst coverage in emerging markets. *Journal of Financial Economics* 80, 115–147.
<https://doi.org/10.1016/j.jfineco.2005.03.010>
- Clarkson, P.M., Kao, J.L., Richardson, G.D., 1999. Evidence That Management Discussion and Analysis (MD&A) is a Part of a Firm's Overall Disclosure Package. *Contemporary Accounting Research* 16, 111–134. <https://doi.org/10.1111/j.1911-3846.1999.tb00576.x>
- Dellavigna, S., Pollet, J.M., 2009. Investor Inattention and Friday Earnings Announcements. *The Journal of Finance* 64, 709–749. <https://doi.org/10.1111/j.1540-6261.2009.01447.x>
- Dhaliwal, D.S., Li, O.Z., Tsang, A., Yang, Y.G., 2011. Voluntary Nonfinancial Disclosure and the Cost of Equity Capital: The Initiation of Corporate Social Responsibility Reporting. *The Accounting Review* 86, 59–100. <https://doi.org/10.2308/accr-00000005>
- Durnev, A., Morck, R., Yeung, B., Zarowin, P., 2003. Does Greater Firm-Specific Return Variation Mean More or Less Informed Stock Pricing? *Journal of Accounting Research* 41, 797–836. <https://doi.org/10.1046/j.1475-679X.2003.00124.x>
- Egeth, H., Kahneman, D., 1975. Attention and Effort. *The American Journal of Psychology* 88, 339. <https://doi.org/10.2307/1421603>
- Fama, E.F., MacBeth, J.D., 1973. Risk, Return, and Equilibrium: Empirical Tests. *Journal of Political Economy* 81, 607–636. <https://doi.org/10.1086/260061>
- Ferracuti, E., Lind, G., 2025. Macroeconomic Information Acquisition Around Earnings Clusters. *J of Accounting Research* 1475–679X.70020. <https://doi.org/10.1111/1475-679X.70020>
- Goh, B.W., Lee, J., Li, D., Wang, M., 2024. Informativeness of Key Audit Matters: Evidence from China. *AUDITING: A Journal of Practice & Theory* 43, 139–164.
<https://doi.org/10.2308/AJPT-2020-099>
- Goldstein, I., Yang, S., Zuo, L., 2023. The Real Effects of Modern Information Technologies: Evidence from the EDGAR Implementation. *Journal of Accounting Research* 61, 1699–1733. <https://doi.org/10.1111/1475-679X.12496>
- Guo, F., Lyu, B., Lyu, X., Zheng, J., 2025. Social Media Networks and Stock Price Synchronicity: Evidence from a Chinese Stock Forum. *Abacus* 61, 419–461.
<https://doi.org/10.1111/abac.12341>
- Hansen, B.E., 1999. Threshold effects in non-dynamic panels: Estimation, testing, and inference. *Journal of Econometrics* 93, 345–368. [https://doi.org/10.1016/S0304-4076\(99\)00025-1](https://doi.org/10.1016/S0304-4076(99)00025-1)
- Hirshleifer, D., Lim, S.S., Teoh, S.H., 2011. Limited Investor Attention and Stock Market Misreactions to Accounting Information. *Rev Asset Pric Stud* 1, 35–73.
<https://doi.org/10.1093/rapstu/rar002>
- Hirshleifer, D., Lim, S.S., Teoh, S.H., 2009. Driven to Distraction: Extraneous Events and Underreaction to Earnings News. *The Journal of Finance* 64, 2289–2325.
<https://doi.org/10.1111/j.1540-6261.2009.01501.x>
- Hirshleifer, D., Teoh, S.H., 2003. Limited attention, information disclosure, and financial reporting. *J. Account. Econ.* 36, 337–386. <https://doi.org/10.1016/j.jacceco.2003.10.002>

- Jin, L., Myers, S.C., 2006. *R2* around the world: New theory and new tests. *Journal of Financial Economics* 79, 257–292. <https://doi.org/10.1016/j.jfineco.2004.11.003>
- Jones, C.I., Tonetti, C., 2020. Nonrivalry and the Economics of Data. *American Economic Review* 110, 2819–2858. <https://doi.org/10.1257/aer.20191330>
- Kwan, A., Liu, Y., Matthies, B., 2026. Institutional Investor Attention. *The Journal of Finance* jofi.70009. <https://doi.org/10.1111/jofi.70009>
- Liu, D., Li, Mingzhu, Li, Mingsheng, Shi, J., 2025. Market reaction to announcement of accounting treatment of data assets: evidence from China. *Journal of Accounting Literature*. <https://doi.org/10.1108/JAL-10-2024-0290>
- Liu, J.-Y., Jia, F.-X., Zhang, Y.-N., Liu, Z.-Y. (Ralph), Tse, Y.K. (Mike), 2026. Does data asset disclosure mitigate stock mispricing? A signalling perspective. *Enterprise Information Systems* 2607361. <https://doi.org/10.1080/17517575.2025.2607361>
- Liu, L.X., Lu, R., Sherman, A.E., Zhang, Y., 2023. IPO underpricing and limited attention: Theory and evidence. *Journal of Banking & Finance* 154, 106932. <https://doi.org/10.1016/j.jbankfin.2023.106932>
- Liu, L.X., Sherman, A.E., Zhang, Y., 2014. The Long-Run Role of the Media: Evidence from Initial Public Offerings. *Management Science* 60, 1945–1964. <https://doi.org/10.1287/mnsc.2013.1851>
- Lu, J., 2022. Limited Attention: Implications for Financial Reporting. *Journal of Accounting Research* 60, 1991–2027. <https://doi.org/10.1111/1475-679X.12432>
- McCAHERY, J.A., Sautner, Z., Starks, L.T., 2016. Behind the Scenes: The Corporate Governance Preferences of Institutional Investors. *The Journal of Finance* 71, 2905–2932. <https://doi.org/10.1111/jofi.12393>
- Morck, R., Yeung, B., Yu, W., 2000. The information content of stock markets: why do emerging markets have synchronous stock price movements? *Journal of Financial Economics*, Special Issue on International Corporate Governance 58, 215–260. [https://doi.org/10.1016/S0304-405X\(00\)00071-4](https://doi.org/10.1016/S0304-405X(00)00071-4)
- Odean, T., 1999. Do Investors Trade Too Much? *American Economic Review* 89, 1279–1298. <https://doi.org/10.1257/aer.89.5.1279>
- Peng, L., Xiong, W., 2006. Investor attention, overconfidence and category learning. *Journal of Financial Economics* 80, 563–602. <https://doi.org/10.1016/j.jfineco.2005.05.003>
- Piotroski, J.D., Roulstone, D.T., 2004. The Influence of Analysts, Institutional Investors, and Insiders on the Incorporation of Market, Industry, and Firm-Specific Information into Stock Prices. *The Accounting Review* 79, 1119–1151. <https://doi.org/10.2308/accr2004.79.4.1119>
- Yu, Z., Li, L., Tian, G., Zhang, H., 2013. Aggressive reporting, investor protection and stock price informativeness: Evidence from Chinese firms. *Journal of International Accounting, Auditing and Taxation* 22, 71–85. <https://doi.org/10.1016/j.intacaudtax.2013.07.004>
- 伊志宏, 杨圣之, 陈钦源, 2019. 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. *中国工业经济* 156–173. <https://doi.org/10.19581/j.cnki.ciejournal.2019.01.009>
- 何瑛, 陈丽丽, 杜亚光, 2024. 数据资产化能否缓解“专精特新”中小企业融资约束. *中国工业经济* 154–173. <https://doi.org/10.19581/j.cnki.ciejournal.2024.08.008>
- 侯宇, 叶冬艳, 2008. 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.

- 金融研究 131–145.
- 危雁麟, 张俊瑞, 汪方军, 程茂勇, 2022. 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. 管理工程学报 36, 130–141.
<https://doi.org/10.13587/j.cnki.jieem.2022.05.011>
- 吴武清, 赵越, 闫嘉文, 汪寿阳, 2020. 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. 管理科学学报 23, 108–126.
- 宋昕倍, 陈莹, 逯东, 程杰, 2024. 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. 南开管理评论 27, 30–39.
- 张娆, 张雅婷, 2025. 数据资产信息披露与分析师乐观偏差——“光环效应”还是“信息风险”? 中央财经大学学报 74–90. <https://doi.org/10.19681/j.cnki.jcufe.2025.11.003>
- 彭红星, 毛新述, 2017. 政府创新补贴、公司高管背景与研发投入——来自我国高科技行业的经验证据. 财贸经济 38, 147–161.
- 方军雄, 伍琼, 傅颀, 2018. 有限注意力、竞争性信息与分析师评级报告市场反应. 金融研究 193–206.
- 朱红军, 何贤杰, 陶林, 2007. 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. 金融研究 110–121.
- 李世刚, 邵宏彬, 方芳, 卢福财, 2025. 企业数据资产信息披露与资本市场定价效率. 中国工业经济 138–155. <https://doi.org/10.19581/j.cnki.ciejournal.2025.07.010>
- 江艇, 2022. 因果推断经验研究中的中介效应与调节效应. 中国工业经济 100–120.
<https://doi.org/10.19581/j.cnki.ciejournal.2022.05.005>
- 牛彪, 于翔, 2024. 数据资产获得投资者偏好了吗? ——基于权益资本成本视角. 证券市场导报 68–79.
- 王亚平, 刘慧龙, 吴联生, 2009. 信息透明度、机构投资者与股价同步性. 金融研究 162–174.
- 王木之, 李丹, 2019. 新审计报告和股价同步性. 会计研究 86–92.
- 王艳, 杨达, 2024. 中国式管理会计体系变革: 从数据要素到数据资产. 管理世界 40, 171–189. <https://doi.org/10.19744/j.cnki.11-1235/f.2024.0111>
- 王艳艳, 于李胜, 安然, 2014. 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. 金融研究 178–191.
- 王英允, 彭正银, 高敬忠, 2019. 投资者注意力与管理层业绩预告择机——积极监督抑或过度压力. 经济管理 41, 139–155. <https://doi.org/10.19616/j.cnki.bmj.2019.02.009>
- 王雄元, 高曦, 何捷, 2018. 年报风险信息披露与审计费用——基于文本余弦相似度视角. 审计研究 98–104.
- 胡军, 王甄, 2015. 微博、特质性信息披露与股价同步性. 金融研究 190–206.
- 苑泽明, 于翔, 李萌, 2022. 数据资产信息披露、机构投资者异质性与企业价值. 现代财经(天津财经大学学报) 42, 32–47. <https://doi.org/10.19559/j.cnki.12-1387.2022.11.003>
- 蒋丹凌, 彭珣, 朱宏泉, 2023. A 股关联公司的股票动量溢出效应研究——基于分析师共同覆盖的视角. 系统工程理论与实践 43, 1891–1909.
- 袁玉, 吴战篪, 廖佳, 2025. 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. 南开管理评论 28, 123–134.
- 许年行, 于上尧, 伊志宏, 2013. 机构投资者羊群行为与股价崩盘风险. 管理世界 31–43.
<https://doi.org/10.19744/j.cnki.11-1235/f.2013.07.004>
- 贺佳, 屈波, 郭俊汝, 2025. 基金经理媒体报道与个体投资者行为——来自基金申赎微观大数据的证据. 金融研究 188–206.
- 赵纳晖, 刘瑾, 2024. 内部控制评价报告文本特征与资本市场定价效率. 审计研究 148–160.

- 郭士祺, 梁平汉, 2014. 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. 经济研究 49, 116–131.
- 陈雨恬, 杨子晖, 温雪莲, 2024. 预期引导、经济韧性与宏观经济治理. 管理世界 40, 66–88.
<https://doi.org/10.19744/j.cnki.11-1235/f.2024.0124>
- 马勇, 喻逸伟, 2023. 众人拾柴火焰高:不完全信息交流社会网络与市场质量. 系统工程理论与实践 43, 3424–3440.
- 黄俊, 郭照蕊, 2014. 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. 管理世界 121–130. <https://doi.org/10.19744/j.cnki.11-1235/f.2014.05.010>
- 黄蓉, 陶长风, 韩若琦, 2025. 金融服务型数据平台企业数据资产会计处理——基于数库科技的案例分析. 管理世界 41, 181–199.
<https://doi.org/10.19744/j.cnki.11-1235/f.2025.0119>