

公共数据开放何以改善资本市场定价效率*

孙 凡¹ 孙泽宇²

(1.山西财经大学会计学院,山西 太原 030031;
2.北京师范大学湾区国际商学院,广东 珠海 519087)



内容提要:公共数据是我国占比最大的数据资源,评估其开放共享对资本市场定价效率的影响,是探寻数据要素何以成为经济增长新动能的基础性问题。本文以我国各省份逐步上线公共数据开放平台为准自然实验,检验了公共数据开放对资本市场定价效率所产生的实际影响及其内在机理。研究发现,公共数据开放显著改善了资本市场定价效率,表现为公司股价同步性显著降低。促进投资者私有信息搜集和提升上市公司信息披露质量是上述关系的两条影响路径。公共数据开放对资本市场定价效率的积极影响在数据开放质量较高、企业数字化转型程度较低、散户规模较大以及地方政府信息透明度较高中更为显著。本文不仅为新时代我国数据要素驱动资本市场高质量发展提供了基于定价效率层面的经验证据,还对不断完善和优化我国地区层面公共数据开放以及逐步完善资本市场价格形成机制具有有益启示。

关键词:数据要素 公共数据 资本市场定价效率 股价同步性

中图分类号:F275 **文献标志码:**A **文章编号:**1002—5766(2025)01—0166—24

一、引 言

当前我国公共事业领域的相关数据资源80%以上由各级政府所掌握(方锦程等,2023)^[1],在确保国家安全的大前提下,逐步稳妥向各类社会主体开放公共数据资源,是激发数据要素活力、更好地赋能经济社会高质量发展的必由之路。作为全球最大的数据生产国,数据要素市场是当前及今后一个时期内驱动经济发展新动能的重要领域。2023年12月31日,国家数据局等十七部门联合印发了《“数据要素×”三年行动计划(2024—2026)》,以推动数据要素高水平应用,发挥数据要素乘数效应。为切实鼓励各类社会主体积极开发利用政府数据资源以释放公共数据潜力与价值,自2012年以来,我国便开始逐步推行公共数据开放,各地区纷纷建设公共数据开放平台并以此作为释放公共数据价值的政策着力点。有鉴于此,深入探讨公共数据开放会否及如何驱动资本市场高质量发展具有一定的现实意义。

国务院于2024年4月印发的《关于加强监管防范风险推动资本市场高质量发展的若干意见》(即“新国九条”)把促进资本市场高质量发展摆在更加突出的位置。切实改善资本市场定价效率不仅事关当前及今后一个时期我国资本市场高质量发展的底色与成色(巫岑等,2022)^[2],也是学界、业界与各级政府的主要关切所在,更涉及广大市场投资者的切身利益(孙泽宇和孙凡,2024)^[3],因此,系统探寻新时代背景下有效优化我国资本市场价格形成机制的可行之策具有重

收稿日期:2024-07-30

* 基金项目:国家自然科学基金面上项目“股票发行注册制改革、信息披露质量与资源分配效率”(72172116)。

作者简介:孙凡,男,教授,硕士生导师,管理学博士,研究领域为企业信息披露与会计电算化,电子邮箱:707350026@qq.com;孙泽宇,男,特聘副研究员,硕士生导师,管理学博士,研究领域为宏观经济金融政策、中观资本市场制度安排与微观企业行为决策间互动关系,电子邮箱:zeyusun666@bnu.edu.cn。通讯作者:孙泽宇。

要且紧迫的现实意义。为此,本文试图以公共数据开放这一近年来各级政府为推动数据要素充分流动所实施的重要决策作为研究切入点,系统考察公共数据开放会否及如何改善资本市场定价效率。本研究主要关注数字经济背景下公共数据开放所激发的数据要素动能对资本市场定价效率的外部治理效应,试图揭示其背后机制,并进一步探寻影响这一治理效应的相关差异化情境。

从理论层面而言,一方面,公共数据开放为广大市场投资者提供了可信度较高的增量信息(Ma 等,2024)^[4],有助于降低投资者以往获取这类信息的成本,并促使其将更多精力投入到针对目标公司私有信息搜集,这类增量公共数据所蕴藏的有关行业及地区层面的信息能为投资者私有信息搜集指明方向,这有助于促进投资者私有信息搜集倾向与效率,通过市场交易行为将其及时反映到股价中以提升定价效率。另一方面,公共数据开放在客观上削弱了以管理层和大股东为代表的内部人相较于以投资者为代表的外部人具有的信息优势(黄俊威,2020)^[5],既加剧了公司选择持有而非披露相关信息所面临的潜在成本,同时也增加了公司选择披露而非持有相关信息所面临的潜在收益,有助于倒逼企业改善信息披露质量,这有助于优化资本市场定价效率。从方法层面来看,各省份渐次上线公共数据开放平台为本研究提供了较理想的准自然情境(彭远怀,2023)^[6],有助于通过构建多时点双重差分模型来考察公共数据开放对资本市场定价效率的影响,以期缓解以往研究所面临的内生性。

基于前述考量,本文以 2007—2022 年我国沪深 A 股非金融类上市公司为研究样本,利用各省份分时分点逐步上线公共数据开放平台这一外生冲击,构造多时点双重差分模型,系统考察了公共数据开放对资本市场定价效率的具体影响及内在机理,以期揭示数据要素赋能资本市场高质量发展的潜在路径与深层逻辑。研究发现,企业所在地区公共数据开放平台上线后,其股价同步性显著降低,证实了公共数据开放对改善资本市场定价效率所发挥的关键作用。机制分析发现,公共数据开放所引致的市场投资者私有信息搜集质量以及上市公司信息披露质量提升是其改善资本市场定价效率的两个内在逻辑。异质性分析表明,公共数据开放对资本市场定价效率的积极影响在数据开放质量较高、企业数字化转型程度较低、散户规模较大以及地方政府透明度较高的样本中更显著。

本文的研究贡献主要体现在以下三方面:第一,立足于近年来各地区公共数据开放平台建设实际,本文提供了数据要素赋能资本市场高质量发展的微观证据,并有效丰富了数字经济优化资本市场价格形成过程的文献。现有关于公共数据开放在资本市场发展层面所引致经济后果的文献尚处早期,少量研究仅考察了其对公司股价崩盘风险的影响(Ma 等,2024)^[4],而对以公共数据开放为代表的要素流动会否及如何有效改善资本市场定价效率仍缺乏理论与经验证据。本文借助我国各省份公共数据开放平台逐步上线这一准自然实验,发现公共数据这一关键数据要素在改善资本市场定价效率层面所产生的积极影响,不仅为数字经济发展与资本市场定价效率等研究领域提供了有力补充,还为数据要素推动资本市场高质量发展形成启发。第二,基于定价效率这一事关中小投资者切身利益以及资本市场健康发展的重要视角,本研究从公共数据开放视角丰富了有关资本市场定价效率的研究文献,并为改善资本市场定价效率提供了新的思路。尽管大量文献从外部政策制度及内部公司治理等多维视角系统考察了资本市场定价效率的诸多影响因素(Boubaker 等^[7],2014;Neifar 和 Ajili,2019^[8];Zhai 等,2021^[9];Zheng 等,2023^[10]),但作为信息的重要载体,数据要素会否及如何影响资本市场定价效率这一重要命题尚未得到学界有效关注。第三,基于新时代背景下“数字中国”这一重要国家战略背景,本研究对逐步推动并完善公共数据平台立法和实践具有有益启示。推动公共数据共享利用是“数字中国”建设中完善数据资源体系这一基础工作的重要任务(陈艳利和蒋琪,2024)^[11],本文探讨了公共数据开放对资本市场定价效率在不

同情境下的差异化影响,这些研究结论揭示了不同平台、企业、投资者以及地区在数据要素赋能股价形成过程中所面临的潜在缺陷,为当前及今后一个时期逐步从上述视角来完善制度设计形成启发。

二、制度背景与文献回顾

1. 制度背景

随着数字经济的发展,数据要素对于促进经济社会各领域高效运行的作用不断凸显。公共数据是我国占比最大的数据资源(方锦程等,2023)^[1],促进公共数据开放共享和高效利用,不仅是新时代政府治理转型的关键路径,更是推动经济高质量发展的重要环节。公共数据开放已逐步成为各国政府部门数据治理的新方向,截至目前,英国、美国、加拿大等国都建立了数据开放平台以推动公共数据的开放共享与充分利用。近年来,我国政府逐步有序推动公共数据开放,2015年8月31号,国务院发布了《促进大数据发展行动纲要》,强调“大力推动政府信息系统和公共数据互联开放共享”,提出“形成公共数据资源合理开放的法规制度和政策体系”的总体目标,至此,我国公共数据开放被正式提上日程。2017年2月,中央全面深化改革领导小组会议审议通过《关于推进公共信息资源开放的若干意见》,进一步推进和规范了公共信息资源开放,明确了公共信息资源所具有的经济价值和社会效应。随后,中共中央在《“十四五”数字经济发展规划》《关于加强数字政府建设的指导意见》《关于构建数据基础制度更好发挥数据要素作用的意见》《数字中国建设整体布局规划》等多项文件中,数次着重强调了促进公共数据资源开放共享的重要性,并明确提出构建公共数据开放平台的总体要求。2023年12月31日,国家数据局等十七部门联合印发了《“数据要素×”三年行动计划(2024-2026)》,明确把更好发挥政府作用,扩大公共数据资源供给置于数据要素驱动经济发展的突出位置。

在国家大力部署与积极推动下,各地公共信息资源开放工作稳步推进。复旦大学数字与移动治理实验室发布的《中国地方公共数据开放利用报告》显示,截至2023年8月,我国已有226个省级(不含港澳台地区)和市级地方政府上线了数据开放平台,其中省级平台26个,市级平台200个。表1列示了全国已有省级公共数据开放平台的上线时间分布情况。从开放数据的特点来看,以2019年8月出台的《上海市公共数据开放暂行办法》为例,其中提到,上海市数据开放遵循“需求导向、安全可控、分级分类、统一标准、便捷高效”的基本原则,其他省份的立法也基本强调了类似的数据质量和开放及时性等关键点。从开放数据的内容来看,各地公共数据开放平台提供的数据大致可分为生产资料、产品市场以及政务服务三类。涉及生产资料的信息包括地理、能源等方面;产品市场信息包括人口、消费等方面;政务服务信息则包括信用、司法等方面(彭远怀,2023)^[6]。从开放数据的服务对象来看,社会公众、政府管理机构及企事业单位、商业机构是三大主要的数据利用主体(周文泓等,2022)^[12]。

表1 省级公共数据开放平台上线时间

开放时间	公共数据开放平台上线的省份
2012年	北京市、上海市
2015年	浙江省
2016年	广东省、贵州省
2018年	河南省、江西省、宁夏回族自治区、山东省、陕西省
2019年	天津市、福建省、海南省、江苏省、四川省、新疆维吾尔自治区

续表 1

开放时间	公共数据开放平台上线的省份
2020 年	广西壮族自治区、湖北省、湖南省、青海省
2021 年	重庆市、安徽省、甘肃省、河北省
2022 年	辽宁省
2023 年	山西省

2. 文献回顾

本研究主要与公共数据开放以及资本市场定价效率两方面文献有关,以下将分别进行文献回顾。

(1)公共数据开放相关研究。现有关于公共数据开放的研究大多集中在公共管理领域,研究了公共数据开放的理论内涵、制度基础以及现实挑战(姚志奋和王保民,2023^[13];范佳佳,2024^[14];赵需要等,2024^[15])。总体来看,通过经验研究来分析公共数据开放经济效应的文献尚处早期发展阶段,主要集中在企业价值创造、资本市场发展与宏观经济运行三个层面。

第一,就公共数据开放对企业价值创造这一领域的研究而言,现有文献主要考察了其在提升企业全要素生产率(彭远怀,2023^[6];杨秀云和韩奇,2023^[16];吴武清等,2024^[17])、促进企业投资与创新(陈艳利和蒋琪,2024^[11];于文超等,2020^[18];蔡菲莹和黄秀霁,2022^[19];刘志伟和张秋生,2024^[20];张吉昌和龙静,2024^[21])、强化企业劳动雇佣(刘达等,2024)^[22]以及提升企业数字化转型水平(韩奇和杨秀云,2024)^[23]四个方面的效应。

第二,还有一支文献从资本市场发展视角研究发现,公共数据开放所引致的经济后果体现在公共数据开放通过促进投资者信息发现而降低了上市公司股价崩盘风险(Ma等,2024)^[4],同时,公共数据开放也会对资本市场信息中介产生影响,表现为降低审计师信息搜集与处理成本,进而减少审计费用(潘俊等,2023)^[24]。

第三,另一支文献则集中探讨了公共数据开放对宏观经济运行的影响。具体地,方锦程等(2023)^[1]立足于区域协调发展这一宏观视角,发现公共数据开放通过破除区域信息壁垒以及弥合区域资源禀赋差距,有助于实现区域协调发展。类似地,张晨等(2023)^[25]利用大样本跨国数据研究发现,公共数据开放能够显著促进国家层面的经济发展。同时,黄先海和虞柳明(2023)^[26]发现,政府数据开放通过提高城市信息数据产业的创业活力和人才集聚程度来促进城市创新水平的提升。张莉和林安然(2023)^[27]借助 76 个地方政府数据开放平台上的开放数据数量和质量信息,考察了其对地区经济高质量发展的影响,发现公共数据开放主要通过降低信息不对称程度、减少企业创新成本等机制,促进以全要素生产率为衡量标准的经济发展质量的提高。进一步地,欧阳伊玲等(2024)^[28]则利用 2003—2019 年城投债发行数据研究发现,公共数据开放显著降低了当地城投债的发行利差,且上述效应还具有一定的经济意义。

(2)资本市场定价效率相关研究。现有研究主要使用股价同步性来衡量资本市场定价效率。所谓股价同步性,即指上市公司股价相较于其同行企业的同涨同跌程度(许年行等,2011^[29];陈冬华和姚振晔,2018^[30])。现有研究一般认为,股价同步性与资本市场定价效率间呈负向关系,亦即股价同步性越高,则意味着股价中所包含的公司特质信息越少而噪音交易较多,则资本市场定价效率较低(Gassen等,2020^[31];Qiu等,2020^[32])。相反地,若股价同步性越低,则表明公司股价中所包含的自身特质信息较多而噪音交易或与其他公司共有信息较少,则资本市场定价效率较高。Roll(1988)^[33]首次明确提出“股价同步性”这一概念以衡量个股“同涨同跌”程度。已有相关文献主

要从外部政策制度以及内部公司治理两个层面对企业股价同步性的影响因素进行探讨。就外部政策制度层面而言,Yoon(2021)^[34]发现中国资本市场开放政策通过强化私有信息披露的方式,显著降低了试点公司股价同步性,进而提升了资本市场运行效率;Kim等(2021)^[35]则利用美国各州渐次采用不可避免披露原则所形成的准自然情境,利用双重差分模型考察了其对上市公司股价同步性的影响。就内部公司治理层面而言,邓伟等(2020)^[36]则发现上市公司设置反收购条款越多则股价同步性越高;Barka等(2022)^[37]研究发现家族控制会通过强化公司信息披露的途径降低股价同步性。

(3)文献述评。有关公共数据开放的实证研究近两年开始显现,但总体数量仍处发展之势,虽然部分研究发现公共数据开放在资本市场发展层面对上市公司股价崩盘风险的影响,但鲜有文献从资本市场定价效率这一独特视角来探讨公共数据开放所能发挥的潜在外部治理效应。同时,现有关于资本市场定价效率的研究主要集中在外部政策制度以及内部公司治理两个层面,而数据要素作为实现“数字中国”战略的基础支撑,其与资本市场定价效率间的具体关系尚不明晰。定价效率是我国资本市场健康发展的核心衡量指标,探究数据要素对资本市场定价效率的实际影响在客观上有助于更好地挖掘经济增长新动能并切实完善资本市场价格形成机制。本研究试图从公共数据开放视角来探究其对资本市场定价效率的具体影响,不仅有效补充了公共数据开放所引致经济后果及资本市场定价效率各自领域的相关研究视角,具体研究结论还能够对切实促进我国数字经济发展以及更好地促进资本市场高质量发展具有一定借鉴,因此,在客观上兼具一定的理论及现实意义。

三、理论分析与研究假设

作为资本市场高质量发展的核心衡量指标之一,资本市场定价效率不仅反映出市场中相关投资者搜集信息、处理信息并将其运用于交易行为的能力与程度,同时也在一定程度上映射出市场内相关上市公司的信息披露意愿及水平。如何有效提升资本市场定价效率进而促使其价格形成过程更加符合公司基本面及投资者预期,一直是我国政府及资本市场监管机构工作的重中之重。从理论上讲,资本市场定价效率不仅取决于上市公司信息披露质量,同时也受到投资者信息获取能力的影响(黄俊和郭照蕊,2014)^[38]。换言之,作为信息披露主体的上市公司披露更多有用信息,且市场投资者能够更好地获取增量信息并将其用于市场交易行为,则资本市场定价效率显然会有明显提升。本文认为,公共数据开放不仅在客观上为市场投资者特别是散户投资者提供了有关上市公司所处地区及行业的增量有用信息,降低投资者针对这类宏观信息的获取成本,进而促进其将更多精力投入到针对公司微观信息的搜集获取,以更好地将前述宏观信息通过股票交易行为反映到公司股价中,还能通过降低公司内外部人之间的信息壁垒来促使企业切实提升信息披露质量,因此,上市公司能够向市场提供更多有用信息,这些信息同样会通过投资者交易行为适时体现在公司股价中,最终有效改善资本市场定价效率。

一方面,公共数据开放为广大市场投资者提供了有关地区与行业层面的海量公共事业数据(彭远怀,2023)^[6],有效降低了投资者获取这类信息的综合成本,进而有助于投资者将其有限的精力与注意力转移到对其他公司私有信息的搜集处理上,从而促使这些信息更加及时有效地反映到公司股价中,最终达到改善资本市场定价效率的积极效果。市场投资者所用以指导投资决策的信息按照其来源大致可划分为宏观及微观两大类,宏观决策信息主要是指宏观经济以及目标公司所在地区、所处行业甚至是所参与各类联盟的一类信息,这些信息虽然并不直接与公司发展经营有关,但却是公司所处各类环境的重要呈现指标,有助于投资者从宏观层面了解公司发展现状及前景,进而帮助其更好地做出股票持有或卖出的相关决策。而另一类信息则是与公司经营活动直接

相关的微观信息,这类信息往往获取难度比较大且时效性比较强,对于散户投资者而言更是如此。根据有限注意力理论,任何个体或组织的精力都是有限的(Hirshleifer 和 Teoh, 2003^[39]; Kempf 等, 2017^[40]; 彭情和唐雪松, 2019^[41]),因此,分配给不同决策的精力是有差异的,如果个体或组织聚焦于某一决策并为其付出大量精力,则势必意味着为其他决策所能够付出的精力会更少。反之,若对某一决策所付出的精力变少,则能够为其他决策所付出的精力将有效增加。有鉴于此,本文认为,公共数据开放通过普惠且公平的方式为公司外部利益相关者特别是散户投资者提供了真实准确且具有一定时效性的海量公共事业数据,这类数据可信度高且广泛涉及地区经济社会发展、公共事业服务以及重点行业等相关信息,有效降低了散户投资者获取宏观信息的成本,有助于其将有限注意力及精力更多地转移到对微观信息(例如公司层面的私有信息)的搜集与处理决策中,进而更好地通过交易行为将宏观信息一并有效反映至公司股价中,从而通过降低股价同步性进而改善资本市场定价效率,最终实现数据要素优化资本市场价格形成过程这一治理效应。需要强调的是,公共数据开放平台所提供的相关数据对使用者而言获取成本极低,只需投资者关注相关平台并实名注册便可批量下载用于指导其投资决策(Ma 等, 2024)^[4],因此,在一定程度上确实有助于节省宏观信息获取成本。从另一个角度讲,公共数据开放所带来的有用信息不仅直接节省了投资者之前为获取这类宏观数据所需付出的可观成本,还能为投资者进行微观信息搜集与处理指明方向且发挥适度的引导作用,进而有利于提升投资者微观信息搜集效率,这同样对改善资本市场定价效率具有一定的促进作用。进一步地,市场投资者关注公共数据是此条逻辑成立的一个基本前提,本研究通过手工搜集以股吧为代表的广大中小投资者讨论平台数据信息发现,有关公共数据相关主题的讨论帖子呈逐渐上升之势,从一个侧面反映出我国散户投资者确实会在日常投资决策中关注并探讨相关公共数据。同时,本研究搜索券商研报发现,其也有提示投资者关注“公共数据”的相关表述,例如中信建投研报曾明确表示“近期,建议关注公共数据以及数据资产评估与入表等相关投资机遇”,这同样为投资者关注并使用公共数据提供了间接证据。

另一方面,公共数据开放在一定程度上弱化了上市公司内部人相较于以市场投资者为代表的外部人的信息优势,进而促使管理层主动向市场提供更多有用信息,这些信息能够通过市场交易行为有效反映到公司股价中,从而改善资本市场定价效率。公司进行信息披露实质上是其权衡信息披露成本与收益后的综合决策,同时受限于信息披露潜在成本及潜在收益这两个方面(Li 和 Zhang, 2015^[42]; Huang, 2022^[43]; Bertomeu, 2023^[44])。公共数据开放平台的建设运营通过向公司外部利益相关者特别是市场投资者提供海量信息的方式,在一定程度上削弱了上市公司内部人所具有的信息优势,此时,公司进行信息披露的成本会有所下降,而其收益则会有所增加,从而促使管理层进行信息披露。具体地,当外部利益相关者通过公共数据开放这一政策获取了更多有关地区与行业甚至是直接与公司有关的交易或其他类型数据时,显然,公司内外部间信息壁垒会有所减弱,亦即管理层或大股东相较于外部投资者的信息优势会被削弱,此时公司进行信息披露所需背负的潜在成本会有所减少。公共数据为包括市场投资者以及潜在竞争对手的外部利益相关者提供了增量有用信息,因此,这些外部利益相关者可能会据此掌握更多有关公司经营发展及行业前景的状况,公司由于担心信息披露带来的诸如信息泄露等构成的潜在风险成本便会有所降低。同时,公司内外部人之间的信息壁垒降低,在一定程度上意味着公司保有而非披露信息的价值意义会有所降低,亦即管理层或大股东等内部人将难以利用原有信息优势进行信息操纵或从事减持行为等来满足私利(Gao 等, 2014)^[45],因此,其可能倾向于增加信息披露强度或改善信息披露质量。当外部投资者的信息搜集与处理精力及能力因公共数据开放而显著改善时,上市公司选择披露信息这一决策所具有的潜在收益也会有所增加,这是因为外部投资者有更多时间精力关注上市公司信息披露,故企业通过信息披露来获取投资者关注信任并促使其买入或继续持有公司股票的一系

列可能性都会有所增加(胡志勇等,2020)^[46]。在这一情形下,市场投资者会通过交易将公司所披露信息及时反映到股价中,进而有助于改善资本市场定价效率。总之,公共数据开放不仅在一定程度上降低了企业信息披露所面临的潜在成本,同时也适度增加了企业选择信息披露所能获得的潜在收益,进而促使理性的管理层通过增加信息披露的途径来予以有效应对,这对切实改善资本市场定价效率具有积极意义。

综上所述,本文认为,公共数据开放不仅能够通过降低宏观信息获取成本的方式促进投资者微观信息搜集处理,同时也可通过缓解公司内外部人之间的信息壁垒来促使管理层改善信息披露质量,进而有效提升资本市场定价效率,最终实现数据要素对资本市场价格形成过程的外部治理效应。因此,本文提出以下假设:

H₁:设定其他条件不变,公共数据开放能够显著改善资本市场定价效率。

四、研究设计

1. 样本选择与数据来源

本文以2007—2022年沪深A股非金融类上市公司为研究样本,以2007年为研究样本起点是考虑到新会计准则的实施对相关财务指标的影响。按照以往研究惯例,本文进一步对前述初始样本做如下剔除:(1)删除相关变量存在缺失值的样本;(2)删除样本期内曾被PT、ST以及ST*的样本;(3)删除上市公司所在地不明确或为缺失值的样本。最终得到36430个公司一年份观测值。为缓解变量异常值对回归结果造成的可能影响,本文对所有回归中涉及的连续变量均进行了上下1%的缩尾处理。进一步地,为缓解异方差及自相关问题对研究结论的干扰,本文亦对所有回归结果进行了公司层面聚类处理。

2. 模型设计与变量定义

为检验公共数据开放这一准外生政策对资本市场定价效率的影响,本文借鉴Chen等(2021)^[47]的做法设定以下多时点双重模型:

$$Syn1_{i,t}/Syn2_{i,t} = \alpha_0 + \beta_1 PostList_{i,t} + \sum Controls_{i,t} + \sum Firm + \sum Year + \sum Province + \varepsilon_{i,t} \quad (1)$$

其中,被解释变量Syn1与Syn2分别为用以衡量资本市场定价效率的两个上市公司年度股价同步性指标。具体地,本文参考许年行等(2011)^[29]的做法,先构造如下模型计算初始个股年收益率的拟合优度指标Syn:

$$r_{i,t} = \chi_i + \delta_{i,1} r_{m,t} + \delta_{i,2} r_{l,t} + \phi_{i,t} \quad (2)$$

其中, $r_{i,t}$ 为股票*i*在第*t*周考虑现金红利再投资的回报率,而 $r_{m,t}$ 则为市场指数第*t*周收益率, $r_{l,t}$ 为行业*l*第*t*周收益率。分别使用每只股票每年数据对模型(2)进行回归,所得到的拟合优度R²即为衡量上市公司年度股价同步性的初始指标Syn。需要强调的是,模型(2)中所使用的市场以及行业层面周收益率采用综合市场等权平均方法进行计算时得到对应的指标Syn1,而使用综合市场总市值平均法计算时得到另一项对应指标Syn2。考虑到前述两个模型拟合优度指标Syn1与Syn2取值位于0~1之间,不便于运用最小二乘法进行线性回归,因此,借鉴已有研究的常用做法,分别对其进行ln((1-Syn1)/Syn1)以及ln((1-Syn2)/Syn2)处理,以得到用以衡量样本公司股价同步性的最终指标Synch1与Synch2。显然,这两个指标取值越大,则意味着上市公司股价同步性越低,亦即资本市场定价效率越高。

在解释变量方面,考虑到省级公共数据开放平台建设运营这一外生冲击实质上呈现“分批多次”这一客观特点,因此,本文主要设置了核心虚拟变量PostList来对其进行刻画。具体地,若上市公司所在省份建设并运营了公共数据开放平台,则在公共数据开放平台运营首年及样本期内以后

各年 *PostList* 均取 1, 否则为 0。实质上, *PostList* 是用来区分上市公司是否受到公共数据开放这一冲击影响的状态变量。

其余影响资本市场定价效率亦即公司股价同步性的相关控制变量参照已有文献设置(王木之和李丹, 2019^[48]; 赵天骄等, 2020^[49]), 同时, 模型(1)中还进一步控制了公司、年份以及省份三个层面的固定效应, 并对回归结果进行了公司层面聚类处理。

变量的具体定义如表 2 所示。

表 2 主要变量定义

变量类型	变量符号	变量名称	变量定义
被解释变量	<i>Synch1</i>	股价同步性 1	具体定义如前文所述
	<i>Synch2</i>	股价同步性 2	具体定义如前文所述
解释变量	<i>PostList</i>	政策效应变量	样本期内上市公司所在省份开通公共数据开放平台的当年及以后各年取 1, 否则为 0
控制变量	<i>Size</i>	公司规模	期末总资产的自然对数
	<i>Lev</i>	资产负债率	负债总额/期末总资产
	<i>Roa</i>	盈利能力	净利润/期末总资产
	<i>Soe</i>	产权性质	若上市公司为国有则取 1, 否则为 0。以上市公司实际控制人性质来判断产权性质
	<i>OwnCon</i>	股权集中度	前十大股东持股数量/公司总股数
	<i>Hhi</i>	行业竞争程度	以公司主营业务收入的赫芬达尔指数计算
	<i>Age</i>	公司年龄	年末减去公司上市日期除以 365 后取对数
	<i>Board</i>	董事会规模	董事会人数加 1 后取自然对数
	<i>Indb</i>	独董比例	董事会中独立董事人数/董事会总人数
	<i>MShare</i>	管理层持股	管理层持股数量/公司总股数
	<i>MSalary</i>	管理层薪酬	高管薪酬总额的自然对数
<i>AnaCover</i>	分析师关注	分析师关注数量加 1 后取自然对数	

五、实证结果与分析

1. 主要变量描述性统计

表 3 列示了模型(1)中所涉及的主要变量的描述性统计结果。两项用以衡量资本市场定价效率的股价同步性指标 *Synch1* 和 *Synch2* 的均值分别为 0.284 以及 0.455, 且其标准差分别为 0.948 与 1.067, 这一结果表明, 样本期内, 我国资本市场定价效率仍偏低, 且不同上市公司间股价同步性存在一定程度的差异, 因此深入探讨改善我国资本市场定价效率的应对之策具有一定的现实意义。在解释变量方面, *PostList* 项的均值为 0.491, 表明样本中约有近一半的观测值受到了本文所关注的公共数据开放这一外生冲击的影响, 在一定程度上体现了公共数据开放的普适性与广泛性, 进而为本研究奠定了良好的研究基础。在控制变量方面, *Soe* 与 *MShare* 这两个指标的均值分别为 0.385 以及 0.069, 表明样本中约有 38.5% 的国有上市公司且管理层持股平均为 6.9%, 其余控制变量的描述性统计结果与已有研究基本保持一致。总体而言, 表 3 的结果证实了本研究所选取的样本区间基本合理。

表 3 主要变量描述性统计结果

变量	样本量	均值	标准差	最小值	25分位数	中位数	75分位数	最大值
<i>Synch1</i>	36430	0.284	0.948	-1.610	-0.378	0.191	0.841	3.219
<i>Synch2</i>	36430	0.455	1.067	-1.737	-0.273	0.340	1.047	3.867
<i>PostList</i>	36430	0.491	0.500	0	0	0	1	1
<i>Size</i>	36430	22.143	1.283	19.738	21.221	21.961	22.871	26.105
<i>Lev</i>	36430	0.431	0.206	0.056	0.266	0.424	0.584	0.924
<i>Roa</i>	36430	0.037	0.064	-0.258	0.014	0.037	0.068	0.204
<i>Soe</i>	36430	0.385	0.487	0	0	0	1	1
<i>OwnCon</i>	36430	0.581	0.152	0.233	0.471	0.588	0.697	0.901
<i>Hhi</i>	36430	0.049	0.083	0.004	0.007	0.010	0.054	0.394
<i>Age</i>	36430	2.018	0.922	-0.363	1.417	2.229	2.763	3.331
<i>Board</i>	36430	2.130	0.199	1.609	1.946	2.197	2.197	2.708
<i>Indb</i>	36430	0.375	0.053	0.333	0.333	0.353	0.429	0.571
<i>MShare</i>	36430	0.069	0.136	0	0	0.001	0.060	0.598
<i>MSalary</i>	36430	14.895	0.833	12.730	14.371	14.900	15.427	17.036
<i>AnaCover</i>	36430	1.384	1.191	0	0	1.386	2.398	3.761

2. 基准回归结果

为检验前述假设所关注的公共数据开放与资本市场定价效率间的具体关系,本文以OLS方法对模型(1)进行回归,具体结果如表4所示。列(1)、(2)中,主要解释变量*PostList*的回归系数均在5%的水平上显著为正,从而验证了前述假设H₁,即公共数据开放显著提升了资本市场定价效率,具体表现为上市公司股价同步性显著降低。同时,表4的相关回归结果还具有一定的经济意义。在列(1)中,*PostList*的回归系数为0.038,这意味着公共数据开放所在地区的上市公司股票定价效率较其他地区的上市公司平均提升约13.4%(0.038/0.284)。同理,在列(2)中,*PostList*的回归系数为0.041,这表明公共数据开放所在地区的上市公司股票定价效率较其他地区的上市公司平均提升约9.0%(0.041/0.455)。就控制变量的显著性水平而言,除行业竞争程度(*Hhi*)、公司年龄(*Age*)以及董事会规模(*Board*)这三项未呈现出显著性以外,其余所有控制变量均具有一定程度的显著性水平。

表 4 公共数据开放与资本市场定价效率间关系回归结果

变量	(1)	(2)
	<i>Synch1</i>	<i>Synch2</i>
<i>PostList</i>	0.038** (0.023)	0.041** (0.017)
<i>Size</i>	-0.158*** (0.000)	-0.271*** (0.000)
<i>Lev</i>	0.533*** (0.000)	0.632*** (0.000)

续表 4

变量	(1)	(2)
	<i>Synch1</i>	<i>Synch2</i>
<i>Roa</i>	0.696*** (0.000)	0.379*** (0.000)
<i>Soe</i>	-0.080** (0.013)	-0.011 (0.749)
<i>OwnCon</i>	0.702*** (0.000)	0.666*** (0.000)
<i>Hhi</i>	0.300** (0.043)	0.627*** (0.000)
<i>Age</i>	0.024 (0.164)	-0.021 (0.242)
<i>Board</i>	-0.052 (0.320)	-0.049 (0.379)
<i>Indb</i>	-0.097 (0.559)	-0.181 (0.291)
<i>MShare</i>	-0.321*** (0.000)	-0.327*** (0.000)
<i>MSalary</i>	-0.004 (0.712)	-0.029** (0.021)
<i>AnaCover</i>	0.086*** (0.000)	0.008 (0.254)
常数项	3.186*** (0.000)	6.381*** (0.000)
公司/年份/省份固定效应	是	是
样本量	36430	36430
R ²	0.460	0.537

注：括号内为经公司层面聚类调整的 *p* 值；***、**、* 分别代表 1%、5%、10% 的显著性水平，下同

3. 稳健性检验

为确保前文所发现的公共数据开放与资本市场定价效率间关系具有一定稳健性且确为因果关系所致，本文主要从以下八个方面进行了稳健性检验。

(1) 考虑遗漏变量的可能影响。前述所发现的公共数据开放与资本市场定价效率间关系在一定程度上会受到遗漏变量的影响，即存在某些变量会同时影响公共数据开放效果与资本市场定价效率进而对研究结论造成干扰。具体地，本研究主要从企业与地区两个层面考虑遗漏变量，并分别将其作为控制变量补充到基准回归模型中进行检验。就企业层面而言，企业对人工智能以及大数据等前沿技术的感知与应用程度，不仅会影响其对公共数据开放政策的响应力度，同时也会对其股票定价效率产生一定的积极影响(雷光勇等, 2022)^[50]。为此，本研究借鉴吴非等(2021)^[51]的做法，首先计算了“人工智能技术”有关词汇在公司年报中的总词频来近似衡量企业对人工智能的感知及应用程度，并据此生成变量 *AIT*；同理，亦计算了“大数据技术”有关词汇在公司年报中的总词频来近似衡量企业对大数据的感知及应用程度，并据此生成变量 *BDT*，将其作为补充控制变量

一并带入模型(1)中进行回归,结果如表5中Panel A的列(1)、(2)所示,*PostList*的回归系数仍在5%的水平上显著为正,*AIT*的回归系数在1%的水平上显著为正,而*BDT*的回归系数则未有显著性,表明在考虑人工智能及大数据技术的可能影响后,本文所发现的主效应仍成立,同时,相较于大数据技术,企业对人工智能技术的感知及应用程度可能更会对其股票定价效率产生积极影响。进一步地,本文还综合计算了“人工智能技术”“大数据技术”“云计算技术”以及“区块链技术”这四类有关词汇在公司年报中的总词频,据以定义企业整体数字化转型程度(*DigiTran*),并将其作为控制变量带入模型(1)中进行回归,结果如表5中Panel A的中间列(3)、(4)所示,不难发现,*PostList*的回归系数仍具有显著性且*DigiTran*的回归系数也具有显著性,表明在充分考虑多种前沿技术背景下的企业数字化转型程度后,本文的研究结论并未有实质性改变。就地区层面而言,本研究认为,省份层面经济发展水平以及市场化程度既会影响公共数据开放政策的实施意愿及效果,同时也会对相关企业的股票定价效率产生影响(蓝发钦等,2024^[52];Hasan等,2014^[53]),因此,可能会对本文主要研究结论造成影响。本研究分别选取省份层面人均地区生产总值年度增长率(*GdpGrowth*)与其市场化程度指数(*MarketIndex*)作为补充控制变量进行检验,相关结果如表5中Panel A的列(5)、(6)所示,*PostList*的回归系数在1%的水平上显著,而*GdpGrowth*与*MarketIndex*的回归系数亦均在10%的水平上显著,表明在充分考虑地区层面经济发展及市场化程度相关特征后,本文研究结论仍成立。

表5 稳健性检验结果

Panel A: 考虑遗漏变量的可能影响								
变量	(1)	(2)	(3)	(4)	(5)	(6)		
	<i>Synch1</i>	<i>Synch2</i>	<i>Synch1</i>	<i>Synch2</i>	<i>Synch1</i>	<i>Synch2</i>		
<i>PostList</i>	0.037** (0.027)	0.039** (0.025)	0.036** (0.034)	0.038** (0.032)	0.031* (0.065)	0.034* (0.054)		
<i>AIT</i>	0.036*** (0.002)	0.044*** (0.000)						
<i>BDT</i>	-0.002 (0.877)	0.009 (0.427)						
<i>DigiTran</i>			0.022*** (0.001)	0.032*** (0.000)				
<i>GdpGrowth</i>					0.002* (0.058)	0.003*** (0.006)		
<i>MarketIndex</i>					0.024* (0.053)	0.023* (0.067)		
控制变量	控制	控制	控制	控制	控制	控制		
公司/年份/省份固定效应	是	是	是	是	是	是		
样本量	35417	35417	35417	35417	36430	36430		
R ²	0.464	0.540	0.464	0.540	0.460	0.538		
Panel B: 考虑样本期内其他相关政策的影响								
变量	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	<i>Synch1</i>	<i>Synch2</i>	<i>Synch1</i>	<i>Synch2</i>	<i>Synch1</i>	<i>Synch2</i>	<i>Synch1</i>	<i>Synch2</i>
<i>PostList</i>	0.046* (0.056)	0.047* (0.056)	0.038** (0.023)	0.041** (0.017)	0.047* (0.056)	0.049* (0.055)	0.038** (0.024)	0.040** (0.023)

续表 5

Panel B: 考虑样本期内其他相关政策的影响

变量	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	<i>Synch1</i>	<i>Synch2</i>	<i>Synch1</i>	<i>Synch2</i>	<i>Synch1</i>	<i>Synch2</i>	<i>Synch1</i>	<i>Synch2</i>
<i>BDPolicy</i>			-0.012 (0.579)	-0.004 (0.849)				
<i>AIPolicy</i>							0.000 (0.984)	-0.018 (0.417)
控制变量	控制	控制	控制	控制	控制	控制	控制	控制
公司/年份/省份固定效应	是	是	是	是	是	是	是	是
样本量	21340	21340	36430	36430	19800	19800	36430	36430
R ²	0.462	0.532	0.460	0.537	0.462	0.538	0.460	0.537

Panel C: 工具变量处理

变量	(1)	(2)	(3)
	<i>PostList</i>	<i>Synch1</i>	<i>Synch2</i>
<i>PostList_IV</i>		1.282** (0.040)	1.855** (0.010)
<i>IV</i>	0.218*** (0.000)		
<i>PostList</i>			
控制变量	控制	控制	控制
公司/年份/省份固定效应	是	是	是
Cragg-Donald Wald F	44.039		
Kleibergen-Paap rk LM	14.29		
样本量	36430	36430	36430
R ²	0.734		

(2)考虑样本期内其他相关政策的影响。在本文的研究样本期间即 2007—2022 年,正是我国大力提倡并发展大数据与人工智能技术的关键时期,因此与本研究所关注的公共数据开放这一外生冲击相关的其他制度安排同样可能会对本文研究结论造成一定程度的干扰。有鉴于此,本研究借鉴已有文献的做法(王海等,2024)^[54],分别选取国家级大数据试验区以及国家新一代人工智能创新发展试验区两项样本期内与大数据及人工智能息息相关的政策进行稳健性检验。具体地,国家级大数据试验区旨在打造一批大数据先进产品,培育一批大数据骨干企业,推动经济转型升级,且其主要以省份为区域进行分批试点;而国家新一代人工智能创新发展试验区是依托地方开展人工智能技术示范、政策试验和社会实验,在推动人工智能创新发展方面先行先试、发挥带动作用的区域,其主要以城市为区域进行分批试点,本文通过搜索整理相关官方文件获取了详细的前述两项政策分批试点的相关区域及时间节点数据用以支撑该检验。为有效考虑这两种政策对本文研究结论的可能影响,本文主要采用了两种方式:首先,分别从样本中剔除了前述两项政策的试点区域并利用剩余样本对基准回归模型进行检验,相关结果如表 5 的 Panel B 的第(1)、(2)、(5)、(6)所

示,其中,第(1)和(2)列为剔除国家级大数据试验区试点省份的回归结果,而第(5)和(6)列则为剔除国家新一代人工智能创新发展试验区试点城市的回归结果,不难发现,在前述四列中 *PostList* 的回归系数仍均在 10% 的水平上显著。其次,本研究依据所获取的前述两项政策的详细试点区域及时间节点数据分别定义了类似于 *PostList* 的用以衡量特定公司一年份数据是否受政策影响的两个虚拟变量 *BDPolicy* 与 *AIPolicy*。具体地,若上市公司所在地处于国家级大数据试验区试点省份,且其观测值位于试点年份及以后年份,则 *BDPolicy* 取 1,否则为 0;若上市公司所在地处于国家新一代人工智能创新发展试验区试点城市,且其观测值位于试点年份及以后年份,则 *AIPolicy* 取 1,否则为 0。进一步地,分别将前述两个变量带入模型(1)进行回归,相关结果如表 5 的 Panel B 的第(3)、(4)、(7)、(8)所示,不难发现,这四列中 *PostList* 的回归系数仍均在 5% 的水平上显著为正,而 *BDPolicy* 与 *AIPolicy* 的回归系数则未有显著性,表明在充分考虑样本期内其他政策影响后,本文研究结论仍具有稳健性。

(3)工具变量处理。为缓解诸如区域内随时间变化的资源禀赋因素或长期生产力冲击等对资本市场定价效率的系统性影响而引致的内生性困扰,本文进一步采取工具变量法对此进行处理。借鉴相关文献的做法(方锦程等,2023^[1];王海等,2024^[54]),本研究将样本内所涉及各省份从改革开放以来到首批省份公共数据平台上线的前一年(1979—2011年)期间历任省委书记籍贯所在城市到海岸线距离平均值的倒数与全国当期即各年份互联网宽带接入端口数进行交互以形成省级公共数据开放平台上线的工具变量 *IV*。从工具变量相关性要求来看,省委书记对省份经济社会发展负有直接且主要的责任,其籍贯所在城市到海岸线的距离往往与其在对外开放以及经济发展决策中所采取的思路与对策具有一定关联性,也能体现省委书记对开放型政策的包容与接受程度(余华义,2015)^[55],因此可能与公共数据开放决策有关。同时,由于特定省份自改革开放以来的历任省委书记籍贯城市到海岸线距离的均值的倒数实际并不具有时间趋势,因此进一步将全国同期互联网宽带接入端口数量作为乘数进行交乘,而较多互联网宽带接入端口代表着数字基础设施建设水平较高,应当能为公共数据开放政策提供一定支撑。进一步地,从工具变量外生性要求来看,1979—2011年有关历史数据应当与样本期内相关企业经营状况关联较弱且全国互联网宽带接入端口数量的变化相对企业而言外生,应当不会对其股票定价效率产生直接影响。表 5 的 Panel C 的第(1)列呈现了第一阶段的回归结果,*IV* 的回归系数为正且在 1% 的水平上显著,同时,在工具变量识别不足检验中,Kleibergen-Paaprk 的 LM 统计量为 14.29, *p* 值为 0.0002,显著拒绝原假设,而在工具变量弱识别检验中,Cragg-Donald Wald F 的 Wald F 统计量为 44.039,大于 Stock-Yogo 在 10% 水平上的临界值 16.38。进一步地,在表 5 的 Panel C 的第(2)和(3)列中, *PostList_IV* 的回归系数在 5% 的水平上显著为正,表明公共数据开放有助于改善资本市场定价效率。

(4)替换公司股价同步性衡量指标。主要变量的差异化衡量方式同样会对研究结论的可靠性造成一定程度的干扰,为此,本文借鉴已有文献的做法(黄新飞等,2023^[56];刘向强等,2023^[57]),对模型(1)中所涉及的被解释变量即公司股价同步性程度这一指标进行替代衡量。具体而言,本文对模型(2)分别使用分市场等权平均法以及分市场总市值平均法进行计算得到不同拟合优度并进行相应对数处理,以得到另外两个新的用以衡量公司股价同步性的指标 *Synch3* 与 *Synch4*,显然,这些指标取值越大,则意味着资本市场定价效率越高。本文进一步将前述两项新的被解释变量分别带入模型(1)进行回归,其余解释变量以及控制变量保持不变,相关结果表明^①,解释变量 *PostList* 的回归系数仍为正且其显著性水平平均不低于 10%,表明在考虑核心变量多种衡量方式后本文研究

① 因篇幅所限,相关内容正文略去。详见本刊网站登载扩展资料中的附录,下文同。

结论仍成立。

(5) 改变聚类处理维度。在前述模型(1)中,本文仅对所有回归结果进行了公司层面的单一聚类处理,考虑到随不同维度变化的相关因素同样会对研究结果造成一定影响,因此进一步对表4中列示的所有回归结果进行了“省份”层面、“省份×年份”层面以及“公司和年份”层面补充聚类处理。相关结果表明, *PostList* 的回归系数均至少在 10% 的水平上显著为正,表明差异化聚类处理并未对主要研究结论造成实质性影响,进而在一定程度上增强了本文研究结论可靠性。

(6) 平行趋势及时间效应检验。在实证研究中,采用双重差分模型进行政策评估的一个重要前提在于,由政策所自然区分出的处理组与控制组样本应在政策实施前在结果变量方面不存在显著差异,亦即需符合平行趋势假定,唯有如此,方可在一定程度上证实处理组与控制组样本在政策实施后的结果变量上呈现出的显著差异确由政策本身而非其他因素所导致。对本研究而言,需要确保公共数据开放政策实施前,处理组与控制组上市公司在股价同步性这一用以衡量资本市场定价效率的指标间不存在显著差异,而这一差异主要出现在政策实施以后各期间。为此,借鉴已有文献的做法,本文以政策实施当年作为基期,并分别设置了三个政策实施前 ($PostList^{-3} / PostList^{-2} / PostList^{-1}$) 以及另外三个政策实施后 ($PostList^{+1} / PostList^{+2} / PostList^{+3}$) 的时间指示变量来考察平行趋势及时间效应。具体地,若上市公司属于处理组样本且其位于政策实施前三年时则 $PostList^{-3}$ 取 1, 否则为 0; 若上市公司属于处理组样本且其位于政策实施前两年时则 $PostList^{-2}$ 取值为 1, 否则为 0; 若上市公司属于处理组样本且其位于政策实施前一年时则 $PostList^{-1}$ 取值为 1, 否则为 0; 若上市公司属于处理组样本且其位于政策实施后一年时则 $PostList^{+1}$ 取值为 1, 否则为 0; 若上市公司属于处理组样本且其位于政策实施后两年时则 $PostList^{+2}$ 取值为 1, 否则为 0; 若上市公司属于处理组样本且其位于政策实施后三年时则 $PostList^{+3}$ 取值为 1, 否则为 0。本文进一步将上述六个时间指示变量代替 *PostList* 代入模型(1)进行回归,其余变量保持不变。相关结果表明,前述三个指代政策实施前的时间指示变量 ($PostList^{-3} / PostList^{-2} / PostList^{-1}$) 均未有显著性,而其余三个指代政策实施后的时间指示变量 ($PostList^{+1} / PostList^{+2} / PostList^{+3}$) 则具有一定程度的显著性,这些结果不仅证实了平行趋势假定,同时也验证了本文所研究的公共数据开放对资本市场定价效率的改善作用呈现出一定程度的时间效应。

(7) 城市层面公共数据开放再检验。借鉴方锦程等(2023)^[1]的做法,本文进一步手工搜集了城市层面公共数据开放平台的上线时间,并据此对主要研究结论进行了基于城市层面的再检验。具体地,与模型(1)中解释变量类似,本文定义了新的基于城市公共数据开放平台上线情况的虚拟变量 *CPostList*。具体地,若上市公司所在城市在样本期内上线了公共数据开放平台且样本年份位于该项政策实施当年及以后各年时则 *CPostList* 取 1, 否则为 0。进一步地,本文将前述虚拟变量代替模型(1)中原有的解释变量 *PostList* 进行回归,其余变量保持不变,且将省份固定效应替换为与变量衡量维度一致的城市固定效应。相关结果表明, *CPostList* 的回归系数仍均至少在 5% 的水平上显著为正,且其回归结果仍具有相当水平的经济意义(15.49%/8.57%)。这一结果表明,公共数据开放对资本市场定价效率的改善效应同时存在于省份与城市层面,进而强化了研究结论的可信度。

(8) 安慰剂检验。为进一步排除公共数据开放与资本市场定价效率间关系可能受到的人为设定或遗漏变量问题,本文参考周泽将等(2019)^[58]的做法,对主要回归结果进行了安慰剂检验。具体地,根据主要解释变量 *PostList* 中取值为 1 占其样本量的固定比例,随机为新变量 *NPostList* 赋值(仅赋 1 或 0),确保这一新生成的随机变量仍为虚拟变量,且其中取值为 1 的比例与 *PostList* 完全一致但分布情况不同。将上述过程重复 1000 次并分别将新生成的两个指标 *NPostList* 放入模

型(1)中,以替代原有 *PostList* 进行回归,并计算 *NPostList* 对应的具体 T 值。这种分配方法既保证了新定义的处理组公司在全样本中所占比例与实际一致,又充分确保了上述过程的随机性。这一检验的总体思路在于,若人为随机改变处理组公司在全样本中分布情况后,上市公司股价同步性程度并未发生显著变化(增加或减少),则强化了公共数据开放与资本市场定价效率间因果关系推断。

本文将 *NPostList* 对应的 1000 个 T 值的具体分布情况按照两个不同的被解释变量分别列示成两张图,可以发现,这些随机生成项的回归系数对应 T 值基本在 0 附近呈正态分布。本研究还进一步检验了上述 T 值的均值是否等于 0,结果显示,不能拒绝其均值为 0 的原假设,表明这些 T 值基本在 0 附近。这一检验结果表明,不仅前述随机生成项的回归系数基本是不显著的,同时,其实际 T 值亦小于实际 *PostList* 在模型(1)中回归所得 T 值,这在一定程度上表明,本文所发现的公共数据开放对资本市场定价效率影响是因果关系而非随机结果。

4. 影响机制检验

依据前述分析,公共数据开放主要通过促进投资者私有信息搜集以及提升上市公司信息披露质量这两条途径来改善资本市场定价效率。一方面,公共数据开放为市场投资者提供了有关地区、行业以及公共交易等层面的增量有用信息,不仅有效节省了市场投资者信息搜集成本,进而有助于促使其将有限精力转向其他私有信息搜集,同时,也通过切实强化公共信息供给的途径为投资者私有信息搜集指明方向,这都有助于通过促进投资者私有信息搜集来有效改善资本市场定价效率,进而呈现外部治理效应。另一方面,公共数据开放不仅弱化了公司内部人相较于外部人所固有的信息优势,进而降低了公司从事信息披露所面临的可能成本,同时,也在一定程度上增加了公司选择进行信息披露的潜在收益,从而有助于提升上市公司信息披露质量,这同样会更好地提升资本市场定价效率。

为有效检验前述两条影响机制,有鉴于现有文献对传统三步法来进行影响机制检验所存在的一定质疑(江艇,2022)^[59],本文选取一步法来进行影响机制检验,即只选取相关中介变量并以解释变量对其进行回归这一做法。

首先,参考孙泽宇和孙凡(2024)^[3]的做法,以散户投资者在股吧针对上市公司及其股票的提问行为来近似刻画投资者私有信息搜集程度。具体地,本文构造了三个用以衡量投资者私有信息搜集程度的指标 *InforSearch1*、*InforSearch2* 以及 *InforSearch3*。其中,*InforSearch1* 为东方财富网股吧中特定年度针对特定公司的发帖数量加 1 后的自然对数;*InforSearch2* 为东方财富网股吧中特定年度针对特定公司的所有帖子的阅读数量加 1 后的自然对数;*InforSearch3* 为东方财富网股吧中特定年度针对特定公司的所有帖子的评论数量加 1 后的自然对数。显然,上述三个指标取值越大,则意味着投资者私有信息搜集程度越高。本文将上述三个中介指标作为被解释变量代入模型(1)中进行回归,具体结果如表 6 的 Panel A 所示。不难发现,其列(1)~(3)中 *PostList* 的回归系数均在 1% 的水平上显著为正,这意味着公共数据开放显著促进了投资者私有信息搜集,进而验证了第一条影响机制。

其次,为有效刻画上市公司信息披露质量,本文分别从盈余质量、年报可读性以及管理层业绩预告三个层面来进行综合衡量。一方面,参考柳光强和王迪(2021)^[60]的做法,以修正琼斯模型所得残差的绝对值来衡量公司在特定年度的盈余质量(*EarnMana*),该指标取值越小,则意味着企业盈余质量越高,即公司信息披露质量相对较高;另一方面,借鉴 Bonsall 等(2017)^[61]的做法,以 Bog 指数来度量年报可读性(*ReadAbility*),具体对全部年报进行分词,并统计每个词的词频,用词频的倒数对年报中的每个词进行加权,从而计算每份年报的可读性。该指标取值越大,则意味着年报可读性越高,即上市公司信息披露质量较高。同时,鉴于前述两种指标均属

于强制性信息披露的范畴,为进一步有效地从自愿性信息披露视角来全面刻画上市公司信息披露质量,本文参考李晓溪等(2019)^[62]的做法,通过选取管理层自愿性业绩预告发布频率来衡量上市公司信息披露质量(*ManaFore*),该指标为上市公司在特定年度自愿发布业绩预告的具体次数,若从未发布相关业绩预告则取0。显然,该指标取值越大,则意味着管理层发布自愿性业绩预告的频率越高,即公司信息披露质量相对较好。本文将前述三种衡量上市公司信息披露质量的综合指标分别代入模型(1)替代被解释变量进行回归,相关结果如表6的Panel B所示。不难发现,各列中 *PostList* 回归系数均至少在5%的水平上显著,且其正负方向与前述预期一致,这表明公共数据开放显著提升了相关地区上市公司的信息披露质量,从而验证了前述第二条影响机制。

表 6 影响机制检验结果

Panel A: 促进投资者私有信息搜集			
变量	(1)	(2)	(3)
	<i>InforSearch1</i>	<i>InforSearch2</i>	<i>InforSearch3</i>
<i>PostList</i>	0.055*** (0.000)	0.056*** (0.000)	0.069*** (0.000)
控制变量	控制	控制	控制
公司/年份/省份固定效应	是	是	是
样本量	35374	35374	35374
R ²	0.671	0.774	0.645

Panel B: 提升上市公司信息披露质量			
变量	(1)	(2)	(3)
	<i>EarnMana</i>	<i>ReadAbility</i>	<i>ManaFore</i>
<i>PostList</i>	-0.007*** (0.004)	0.106** (0.014)	0.042** (0.042)
控制变量	控制	控制	控制
公司/年份/省份固定效应	是	是	是
样本量	29168	36104	36430
R ²	0.246	0.673	0.566

5. 异质性分析

为进一步考察公共数据开放与资本市场定价效率间关系在不同治理情境下存在的显著差异,本文分别选取公共数据开放质量(信息发布方)、企业数字化转型程度(信息接收方)、散户规模(信息使用方)以及地方政府透明度(信息支撑方)这四个指标作为具体调节变量,使用分组回归方式进行考察。

(1)公共数据开放质量的调节作用。从理论上讲,公共数据开放质量应当会对前述所关注的主效应具有调节作用。一方面,以开放数据集数量与深度为代表的公共数据开放质量应当有助于更好地促进投资者信息搜集行为,进而更好地降低股价同步性以改善资本市场定价效率;另一方

面,更高质量的省份层面公共数据开放同样能够更好地倒逼地区内企业主动且积极地改善其信息披露质量,从而更好地改善资本市场定价效率。为此,本文借鉴已有文献的做法(彭远怀,2023)^[6],采用复旦大学数字与移动治理实验室构建的中国开放数据指数中所包含的“数据层指数”来衡量省份层面开放数据质量(*OpenDataQuality*)。需要强调的是,该指标最早可获取年份为2017年,为尽可能避免样本损失,本研究将各省份2017—2022年的数据层指数进行平均化处理,并按照其中位数将全样本划分为公共数据开放质量“高”和“低”两组,并分别将其样本带入基准回归模型中进行检验,相关结果如表7所示。在奇数列中,*PostList*的回归系数均在10%的水平上显著,而在其余偶数列中,其回归系数则未有显著性,验证了前述理论预期,即公共数据开放对资本市场定价效率的积极影响在公共数据开放质量更高的样本中更显著。

表7 公共数据开放质量调节作用的回归结果

变量	(1)	(2)	(3)	(4)
	<i>Synch1</i>	<i>Synch1</i>	<i>Synch2</i>	<i>Synch2</i>
	<i>High</i>	<i>Low</i>	<i>High</i>	<i>Low</i>
<i>PostList</i>	0.046* (0.097)	0.006 (0.867)	0.055* (0.052)	0.005 (0.893)
控制变量	控制	控制	控制	控制
公司/年份/省份固定效应	是	是	是	是
样本量	18641	10227	18641	10227
R ²	0.456	0.464	0.539	0.522

(2)企业数字化转型程度的调节作用。前文的影响机制检验表明,公共数据开放促使企业主动积极改善其信息披露质量而优化资本市场定价效率,但这一机制的一个重要前提在于,企业能够及时感知所在地区公共数据开放政策并对此做出有效反应。有鉴于此,本研究认为,企业数字化转型程度将会对前述主效应产生调节作用。一方面,对于数字化转型程度较高的企业而言,其获取并利用所处地区公共数据与信息的能力应当天然强于那些数字化转型程度较低的企业,因此其对公共数据开放政策的反应可能相对较弱。换言之,在这种情境下,公共数据开放对资本市场定价效率的积极影响可能会在企业数字化转型程度较低的企业中更为显著。然而,另一方面,企业数字化转型程度较高则意味着其对所开放的地区层面公共数据更为敏感与重视,且能够运用公司数字化平台与手段进行有效分析并加以利用,这应当能更好地促进其改善信息披露质量,进而优化资本市场定价效率。从这个层面而言,公共数据开放对企业股价同步性的积极影响将会在企业数字化转型程度较高的样本中更显著。换言之,有关企业数字化转型的调节作用有待进一步实证检验。为检验企业数字化转型对前述主效应的调节作用,本文借鉴已有文献的做法(吴非等,2021)^[51],以企业年报中有关“大数据技术”“人工智能技术”“区块链技术”以及“云计算技术”等关键词的总词频作为衡量企业数字化转型程度的代理指标,并依据其中位数,将全样本划分为数字化转型程度较高与较低两组(*High/Low*),分别将其样本代入模型(1)回归,具体结果如表8所示。其奇数列中*PostList*的回归系数均为正且至少在5%的水平上显著,而偶数列中对应的回归系数则未有显著性,这些结果验证了前述理论预期,即公共数据开放对资本市场定价效率的改善作用在企业数字化转型程度较低的样本中更显著,揭示了企业数字化转型对主效应的实际调节作用。

表 8 企业数字化转型调节作用的回归结果

变量	(1)	(2)	(3)	(4)
	<i>Synch1</i>	<i>Synch1</i>	<i>Synch2</i>	<i>Synch2</i>
	<i>Low</i>	<i>High</i>	<i>Low</i>	<i>High</i>
<i>PostList</i>	0.052** (0.023)	0.021 (0.444)	0.063*** (0.009)	0.003 (0.904)
控制变量	控制	控制	控制	控制
公司/年份/省份固定效应	是	是	是	是
样本量	20459	14215	20459	14215
R ²	0.485	0.497	0.559	0.570

(3)散户规模的调节作用。尽管市场投资者作为典型的公司外部人,与以管理层抑或是大股东为代表的公司内部人相比存在天然的信息劣势,但散户与机构投资者被认为同样具有较为显著的信息搜集处理分析差异(谢玲玲,2020)^[63]。具体地,机构投资者由于规模效应等因素的影响,其在针对上市公司及其股票的信息获取、处理以及分析各个过程中都相较于散户投资者具有明显优势,大量文献对此进行了深入研究并得到了相对一致的结论(Aggarwal等,2011^[64];Dyck等,2019^[65])。因此,一个有趣的问题在于,究竟是哪种市场投资者会更加依赖公共信息开放来指导其投资决策进而影响资本市场定价效率?从直觉上看,散户投资者往往缺乏有效的信息来源,公共数据开放为其提供了大量准确且相对及时的信息,这类信息应当能够在促进其私有信息搜集以及指导具体交易决策过程中发挥更大作用,而机构投资者本身就具有相对丰富的信息来源且其与目标公司管理层往往具有较为紧密的非正式联系,同时,机构投资者亦可通过实地调研抑或是电话会议等形式来直接了解公司实际经营情况,因此公共数据开放为其信息获取方面所带来的边际效益可能相对有限。但从另一个角度而言,机构投资者显然相较散户投资者拥有更强的信息分析处理能力,因此在面对公共数据开放所带来的海量相关数据时,机构投资者可能更有能力来有效分析并利用这些信息为其投资决策进行服务,譬如其可利用机器学习抑或大数据模型将这类公共信息加以有效利用,进而优化其投资决策。显然,散户投资者缺乏这样的分析处理能力。总之,究竟何种投资者更能有效利用公共数据来从事投资交易尚不明晰,有待进一步实证检验。

本文参考马连福等(2022)^[66]的做法,以特定年度区间内上市公司股票的散户资金流入流出量之和除以其与机构资金流入流出量之和这一相对指标来近似衡量上市公司在特定年度所拥有的散户规模,依照万德数据库划分,将指定区间内挂单成交额小于4万元的资金划为散户资金而将挂单成交额大于100万元的资金划为机构资金用以计算上述散户规模指标。进一步地,本文依据前述指标中位数将全样本划分为大规模与小规模散户投资者两组(*Large/Small*),分别将其样本代入模型(1)进行检验,相关结果如表9所示。值得注意的是,表9中各奇数列中*PostList*的回归系数均为正且在10%的水平上显著,而各偶数列中对应的回归系数虽为正但却未有显著性,这意味着公共数据开放对资本市场定价效率的积极影响在散户规模较大的样本中更显著,表明散户投资者而非机构投资者更能享有公共数据开放所带来的制度红利,进而将其反映到股价中以实质性影响资本市场定价效率。

表 9 散户规模调节作用的回归结果

变量	(1)	(2)	(3)	(4)
	<i>Synch1</i>	<i>Synch1</i>	<i>Synch2</i>	<i>Synch2</i>
	<i>Large</i>	<i>Small</i>	<i>Large</i>	<i>Small</i>
<i>PostList</i>	0.043* (0.068)	0.023 (0.350)	0.046* (0.066)	0.025 (0.325)
控制变量	控制	控制	控制	控制
公司/年份/省份固定效应	是	是	是	是
样本量	18079	18037	18079	18037
R ²	0.496	0.464	0.556	0.544

(4)地方政府透明度的调节作用。作为公共数据开放平台的建设者、运营者与管理者,地方政府特质在决定公共数据开放所引致经济后果方面具有一定解释力。有鉴于此,本文选取地方政府透明度这一特征来考察其对主效应的调节作用。从理论上讲,一方面,地方政府透明度较高意味着社会公众更加相信地方政府所披露的公共数据(姜艳,2023)^[67],进而更好地发挥其经济效益,因此主效应可能会在地方政府透明度较高的样本中更显著;另一方面,公共数据开放本身即可构成提高地方政府透明度的一个重要因素(赵敬丹和王欢,2021)^[68]。换言之,有效充分且及时地进行公共数据开放,显然有助于提升地方政府透明度,因此,不难预期公共数据开放对资本市场定价效率的积极影响可能会在地方政府透明度较低的样本中更显著。总之,地方政府透明度在影响主效应方面所发挥的具体调节作用有待进一步实证检验。参考已有文献的做法(董新宇和鞠逸飞,2023)^[69],本文以中国人民大学国家发展与战略研究院自2017年起公开对外发布的《中国城市政商关系(排行榜)评价报告》中城市层面“政府透明”这一指标来对地方政府透明程度进行有效量化,显然,该指标取值越大,则意味着地方政府透明度越高。特别地,由于该指数目前只包括2017、2018、2020、2021、2022这五年,因此,按照最近替代原则,将2017年以前的样本期间按照2017年数据进行赋值,而对2019年的数据则以2018年与2020年相关指标进行均值替代。本文进一步根据上述地方政府透明度指标的中位数将全样本划分为高与低透明度两组(*High/Low*),分别将其样本代入模型(1)进行回归,结果如表10所示。不难发现,表10各奇数列中*PostList*的回归系数均为正且至少在5%的水平上显著,而各偶数列中对应的回归系数均未呈现出显著性水平,表明公共数据开放对资本市场定价效率的改善作用在地方政府透明度较高的样本中更显著,从而证实了地方政府特质对主效应的实际调节作用。这一检验同样直接强调了建设高透明度政府对于切实释放公共数据开放所蕴含的政策红利所能发挥的重要作用。换言之,公共数据开放的经济效益很大程度上取决于社会公众是否信任政府。

表 10 地方政府透明度调节作用的回归结果

变量	(1)	(2)	(3)	(4)
	<i>Synch1</i>	<i>Synch1</i>	<i>Synch2</i>	<i>Synch2</i>
	<i>High</i>	<i>Low</i>	<i>High</i>	<i>Low</i>
<i>PostList</i>	0.064** (0.012)	0.018 (0.514)	0.080*** (0.002)	-0.002 (0.945)

续表 10

变量	(1)	(2)	(3)	(4)
	<i>Synch1</i>	<i>Synch1</i>	<i>Synch2</i>	<i>Synch2</i>
	<i>High</i>	<i>Low</i>	<i>High</i>	<i>Low</i>
控制变量	控制	控制	控制	控制
公司/年份/省份固定效应	是	是	是	是
样本量	16614	16338	16614	16338
R ²	0.469	0.456	0.544	0.532

六、研究结论与政策启示

1. 研究结论

切实改善资本市场定价效率不仅对促进资本市场健康发展至关重要,同时也能够在很大程度上实现更好地保护中小投资者切身利益的重要目标,更对行业、地区乃至宏观经济高质量发展具有显著意义。有鉴于此,本文借助省份公共数据开放平台渐次上线这一准外生政策,通过构造多时点双重差分模型,系统考察了以公共数据开放为代表的要素流动对资本市场定价效率的影响及其内在机理。研究不仅证实了公共数据开放有助于切实改善资本市场定价效率,主要体现在有效降低上市公司股价同步性水平,同时,也以理论与实证相结合的方式深刻揭示了促进投资者私有信息搜集与提升上市公司信息披露质量这两条蕴藏在公共数据开放与资本市场定价效率间关系的可能影响机制。进一步地,上述相关结论在经过包括考虑遗漏变量、考虑样本期内其他政策、工具变量处理、替换关键变量衡量方式、改变聚类维度、平行趋势及时间效应检验、城市层面公共数据开放再检验以及安慰剂检验等在内的一系列稳健性检验后依然成立。异质性分析表明,公共数据开放对资本市场定价效率的积极影响在公共数据开放质量较高、企业数字化转型程度较低、散户规模较大以及地方政府透明度较高的样本中更为显著。

2. 政策建议

第一,本研究发现,地方政府信息透明度是公共数据开放改善资本市场定价效率方面的重要保障,因此各地尤其是中西部与欠发达地区政府应着力强化其信息透明度建设,以期更好地助力数据要素赋能经济社会高质量发展。公共数据开放实际是促进全社会数据要素充分流动的第一步,而外部利益相关者是否相信并有意愿在其行为决策中考虑这类数据则很大程度上取决于地方政府透明度。换言之,较高的地方政府透明度不仅使社会公众能够注意到公共数据,还能提高其对公共数据真实性和可靠性的信任程度,因此,只有不断提高地方政府透明度,才能有效释放公共数据开放的制度红利,进而更好地促进包括市场投资者在内的广大社会公众更加积极主动地挖掘和分析利用相关公共数据,从而不仅对其个人决策同样也对整体经济社会高质量发展产生有效地促进作用。

第二,本研究发现公共数据开放质量在很大程度上直接决定了其在改善资本市场定价效率方面所发挥的积极作用,这一结论表明各地政府在建立起各自公共数据开放平台后还应紧盯后续运营,持续发力推动公共数据开放平台的运营质量实现有序且有效地提升,唯有此方可用好公共数据开放平台这一政策工具以期更好促进全社会数据要素充分流动。具体地,建立起公共数据开放平台仅仅是政策的起步阶段,而面对纷繁复杂且日新月异的各类海量公共数据,各地区的公共数据开放平台如何在确保国家数据安全这一大前提下,通过不断丰富所开放数据的维度、深度与广度来有效提升平台运营质量是各地政府应着力重视并予以解决的。事实上,现阶段各省份公共数

据开放平台的相关数据还面临不少时效性、可靠性以及有用性方面的缺陷,这在一定程度上不利于发挥其应有的积极促进数据要素充分流动的应有之义,因此各地政府应该将更多精力集中到公共数据开放平台的不断优化完善之中,进而以持续改善的平台数据质量来更好促进经济社会发展。

第三,本研究发现公共数据开放对资本市场定价效率的积极影响在散户持股比例较高的企业中更为显著,这充分揭示了公共数据开放这一政策对我国广大中小投资者的积极信息效应,进而为我国资本市场中广大散户投资者重视并充分运用目标公司所在地区开放的公共数据提供了有益启发。在我国资本市场中,散户占比相对较高,由此带来的“追涨杀跌”以及“盲目跟风”等非理性短期投资行为对资本市场高质量发展造成不利影响,本研究事实上为以优化资本市场定价效率来有效促进其高质量发展提供了基于政府主导下促进数据要素充分流动的全新治理思路,特别对广大散户投资者而言,应当在其投资决策中适当注意并相对重视公共数据这类独特信息,这对于其全面评估目标公司所处外部环境、行业及地区的综合情况应当具有一定帮助,亦有助于其理性长期决策。

3. 研究局限与展望

本研究虽然得到了有关数据要素流动何以改善资本市场定价效率的初步研究结果,但尚存在一些有待进一步完善之处。首先,囿于数据可得性,本文的研究对象为上市公司,未来研究可结合问卷调查等形式来将广大散户投资者作为研究主体,进而通过其对相关公共数据的感知及应用情况来推测数据要素流动对股票价格形成过程的新影响。其次,现阶段我国各省份公共数据开放平台还处于不断建设完善时期,后续研究可结合具体的公共数据类型来考察其对资本市场定价效率的差异化影响,以期更好揭开数据要素充分流动何以助推资本市场高质量发展这一时代课题。

参考文献

- [1]方锦程,刘颖,高昊宇,等.公共数据开放能否促进区域协调发展?——来自政府数据平台上线的准自然实验[J].北京:管理世界,2023,(9):124-142.
- [2]巫岑,饶品贵,岳衡.注册制的溢出效应:基于股价同步性的研究[J].北京:管理世界,2022,(12):177-202.
- [3]孙泽宇,孙凡.投股中心持股与上市公司股价同步性——基于一项准自然实验[J].北京:管理评论,2024,(2):194-209.
- [4]Ma, R., F. Guo, and D. Li. Can Public Data Availability Affect Stock Price Crash Risk? Evidence from China[J]. *International Review of Financial Analysis*, 2024, 94: 1-33.
- [5]黄俊威.融资融券制度与公司内部人减持——一种市场化治理机制的探索[J].北京:管理世界,2020,(11):143-167.
- [6]彭远怀.政府数据开放的价值创造作用:企业全要素生产率视角[J].北京:数量经济技术经济研究,2023,(9):50-70.
- [7]Boubaker, S., H. Mansali, and H. Rjiba. Large Controlling Shareholders and Stock Price Synchronicity[J]. *Journal of Banking & Finance*, 2014, 40: 80-96.
- [8]Neifar, S., and H. Ajili. CEO Characteristics, Accounting Opacity and Stock Price Synchronicity: Empirical Evidence from German Listed Firms[J]. *Journal of Corporate Accounting & Finance*, 2019, 30, (2): 29-43.
- [9]Zhai, H., M. Lu, Y. Shan, et al. Key Audit Matters and Stock Price Synchronicity: Evidence from a Quasi-natural Experiment in China[J]. *International Review of Financial Analysis*, 2021, 75: 1-11.
- [10]Zheng, S., X. Zhang, and H. Wang. Green Credit Policy and the Stock Price Synchronicity of Heavily Polluting Enterprises[J]. *Economic Analysis and Policy*, 2023, 77: 251-264.
- [11]陈艳利,蒋琪.数据生产要素视角下开放公共数据与企业创新——基于建立公共数据开放平台的准自然实验[J].北京:经济管理,2024,(1):25-46.
- [12]周文泓,张晓宇,田欣,等.我国地方政府开放数据的社会利用现状与促进对策[J].广州:图书馆论坛,2022,(4):76-85.
- [13]姚志奋,王保民.政府数据开放的公共安全悖论及其法治策应[J].北京:中国科技论坛,2023,(8):139-149.
- [14]范佳佳.公共数据进入数据要素市场模式研究[J].武汉:信息资源管理学报,2024,(2):68-81.

- [15]赵需要,姬祥飞,樊振佳.政府数据开放到公共数据开放的嬗变[J].北京:情报理论与实践,2024,(4):50-58.
- [16]杨秀云,韩奇.公共数据开放能提升企业全要素生产率吗[J].深圳:证券市场导报,2023,(12):18-30.
- [17]吴武清,李祁恒,章柳漪,等.公共数据资源与企业全要素生产率——基于地方政府数据开放的准自然实验[J].北京:系统工程理论与实践,2024,(6):1815-1833.
- [18]于文超,梁平汉,高楠.公开能带来效率吗?——政府信息公开影响企业投资效率的经验研究[J].北京:经济学(季刊),2020,(3):1041-1058.
- [19]蔡菲莹,黄秀霁.政府数据开放、企业数字化转型与企业创新[J].武汉:统计与决策,2022,(23):175-179.
- [20]刘志伟,张秋生.数据要素赋能企业投资效率研究——基于公共数据开放的准自然实验[J].成都:经济学家,2024,(9):42-52.
- [21]张吉昌,龙静.政府数据开放何以赋能企业创新?[J].南京:现代经济探讨,2024,(4):29-43.
- [22]刘达,王晓丹,石玉堂,等.公共数据开放的稳就业效应——来自政府数据平台开通的经验证据[J].成都:软科学,2024,(1):1-16.
- [23]韩奇,杨秀云.公共数据开放能否促进企业数字化转型?[J].南京:现代经济探讨,2024,(4):44-59.
- [24]潘俊,卞子咏,赵洵,等.政府公共数据开放与审计师定价决策——基于构建数据开放平台的准自然实验[J].北京:审计研究,2023,(6):136-148.
- [25]张晨,万相昱,姜智超,等.开放政府数据的经济增长效应研究[J].北京:中国软科学,2023,(2):1-11.
- [26]黄先海,虞柳明.政府数据开放能提升城市创新水平吗——来自中国地级市的证据[J].杭州:浙江学刊,2023,(4):88-98.
- [27]张莉,林安然.中国地方政府数据开放与经济发展质量[J].广州:公共行政评论,2023,(2):120-137.
- [28]欧阳伊玲,王愉靖,李平,等.数据要素与城投债定价:基于公共数据开放的准自然实验[J].北京:世界经济,2024,(2):174-203.
- [29]许年行,洪涛,吴世农,等.信息传递模式、投资者心理偏差与股价“同涨同跌”现象[J].北京:经济研究,2011,(4):135-146.
- [30]陈冬华,姚振晔.政府行为必然会提高股价同步性吗?——基于我国产业政策的实证研究[J].北京:经济研究,2018,(12):112-128.
- [31] Gassen, J., H. A. Skaife, and D. Veenman. Illiquidity and the Measurement of Stock Price Synchronicity [J]. *Contemporary Accounting Research*, 2020, 37, (1): 419-456.
- [32] Qiu, B., J. Yu, and K. Zhang. Trust and Stock Price Synchronicity: Evidence from China [J]. *Journal of Business Ethics*, 2020, 167: 97-109.
- [33] Roll, R. Presidential Address: R^2 [J]. *Journal of Finance*, 1988, 43, (2): 540-566.
- [34] Yoon, A. S. The Role of Private Disclosures in Markets with Weak Institutions: Evidence from Market Liberalization in China [J]. *The Accounting Review*, 2021, 96, (4): 433-455.
- [35] Kim, Y., L. Su, Z. Wang, et al. The effect of Trade Secrets Law on Stock Price Synchronicity: Evidence from the Inevitable Disclosure Doctrine [J]. *The Accounting Review*, 2021, 96, (1): 325-348.
- [36] 邓伟,王涛,成园园.反收购条款、股价同步性与投资效率[J].北京:管理评论,2020,(11):33-47.
- [37] Barka, Z., R. Benkraiem, T. Hamza, et al. How Does Family Control Affect Stock Price Synchronicity? [J]. *Finance Research Letters*, 2022, 49: 1-7.
- [38] 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].北京:管理世界,2014,(5):121-130.
- [39] Hirshleifer, D., and S.H. Teoh. Limited Attention, Information Disclosure, and Financial Reporting [J]. *Journal of Accounting and Economics*, 2003, 36, (1-3): 337-386.
- [40] Kempf, E., A. Manconi, and O. Spalt. Distracted Shareholders and Corporate Actions [J]. *The Review of Financial Studies*, 2017, 30, (5): 1660-1695.
- [41] 彭情,唐雪松.流言招来的“是非”:股市传闻与盈余价值相关性[J].北京:管理世界,2019,(3):186-204.
- [42] Li, Y., and L. Zhang. Short Selling Pressure, Stock Price Behavior, and Management Forecast Precision: Evidence From a Natural Experiment [J]. *Journal of Accounting Research*, 2015, 53, (1): 79-117.
- [43] Huang, Y. Government Subsidies and Corporate Disclosure [J]. *Journal of Accounting and Economics*, 2022, 74, (1): 1-28.
- [44] Bertomeu, J. Managers' Choice of Disclosure Complexity [J]. *Journal of Accounting and Economics*, 2023, 76, (2-3): 1-8.
- [45] Gao, F., L. L. Lisic, and I. X. Zhang. Commitment to Social Good and Insider Trading [J]. *Journal of Accounting and Economics*,

2014, 57, (2-3): 149-175.

[46]胡志勇,徐建挺,黄琮宇,等.信任与盈余价值相关性[J].北京:会计研究,2020,(5):48-61.

[47]Chen, W., H.Wu, and L.Zhang.Terrorist Attacks, Managerial Sentiment, and Corporate Disclosures[J].The Accounting Review, 2021, 96, (3): 165-190.

[48]王木之,李丹.新审计报告和股价同步性[J].北京:会计研究,2019,(1):86-92.

[49]赵天骄,李成,张冰石.母子公司产业网络关系与资本市场信息效率——基于股价同步性的经验证据[J].北京:会计研究,2020,(10):136-149.

[50]雷光勇,买瑞东,左静静.数字化转型与资本市场效率——基于股价同步性视角[J].深圳:证券市场导报,2022,(8):48-59.

[51]吴非,胡慧芷,林慧妍,等.企业数字化转型与资本市场表现——来自股票流动性的经验证据[J].北京:管理世界,2021,(7):130-144.

[52]蓝发钦,胡晓敏,徐卓琳.公共数据开放能否拓展资本跨区域流动距离——基于异地并购视角[J].北京:中国工业经济,2024,(9):156-174.

[53]Hasan, I., L.Song, and P.Wachtel.Institutional Development and Stock Price Synchronicity: Evidence from China[J].Journal of Comparative Economics, 2014, 42, (1): 92-108.

[54]王海,叶帅,尹俊雅.公共数据开放如何提振企业有效投资——基于产能利用视角[J].北京:中国工业经济,2024,(8):137-153.

[55]余华义.城市化、大城市化与中国地方政府规模的变动[J].北京:经济研究,2015,(10):104-118.

[56]黄新飞,叶梓南,王升泉.反倾销与中国资本市场信息效率——基于股价同步性的研究[J].北京:经济学(季刊),2023,(5):1954-1972.

[57]刘向强,杨晴晴,胡珺.ESG评级分歧与股价同步性[J].北京:中国软科学,2023,(8):108-120.

[58]周泽将,罗进辉,李雪.民营企业身份认同与风险承担水平[J].北京:管理世界,2019,(11):193-208.

[59]江艇.因果推断经验研究中的中介效应与调节效应[J].北京:中国工业经济,2022,(5):100-120.

[60]柳光强,王迪.政府会计监督如何影响盈余管理——基于财政部会计信息质量随机检查的准自然实验[J].北京:管理世界,2021,(5):157-169.

[61] Bonsall, IV. S. B., A. J. Leone, B. P. Miller, et al. A Plain English Measure of Financial Reporting Readability [J]. Journal of Accounting and Economics, 2017, 63, (2-3): 329-357.

[62]李晓溪,饶品贵,岳衡.年报问询函与管理层业绩预告[J].北京:管理世界,2019,(8):173-188.

[63]谢玲玲.散户与机构投资者差异及分析师报告策略——来自中国市场的经验证据[J].上海:上海金融,2020,(10):58-68.

[64] Aggarwal, R., I. Erel, M. Ferreira, et al. Does Governance Travel Around the World? Evidence from Institutional Investors [J]. Journal of Financial Economics, 2011, 100, (1): 154-181.

[65] Dyck, A., K. V. Lins, L. Roth, et al. Do Institutional Investors Drive Corporate Social Responsibility? International Evidence [J]. Journal of Financial Economics, 2019, 131, (3): 693-714.

[66]马连福,王博,宋婧楠.散户更偏爱数字化吗?——基于投资者情绪异质性的研究[J].北京:经济与管理研究,2022,(9):32-54.

[67]姜艳.公众使用社交媒体、政府透明度与政府信任——基于网民社会意识调查的实证研究[J].桂林:社会科学家,2023,(7):127-133.

[68]赵敬丹,王欢.基于数据开放的透明政府建设创新路径探析[J].北京:理论视野,2021,(9):65-70.

[69]董新宇,鞠逸飞.政府透明能否促进政府服务的提升:基于中国城市面板数据的门槛回归分析——以政府的企业服务为例[J].北京:公共管理与政策评论,2023,(5):80-98.

How does Public Data Openness Improve Capital Market Pricing Efficiency?

SUN Fan¹, SUN Ze-yu²

(1.School of Accounting, Shanxi University of Finance and Economics, Taiyuan, Shanxi, 030031, China;

2.Bay Area International Business School, Beijing Normal University, Zhuhai, Guangdong, 519087, China)

Abstract: Public data constitutes the largest proportion of data resources in China. Evaluating its openness and sharing in relation to the pricing efficiency of the capital market is a fundamental issue in exploring how data elements can become a new driving force for economic growth. Exploring new drivers of economic growth has been a key task in China's economic work in recent years. As the world's largest producer of data, the data element market is the most likely field in China to nurture new drivers. The Central Economic Work Conference held in December 2023 systematically deployed the economic work for 2024, emphasizing the important role of the digital economy and data elements. On December 31, 2023, the National Data Bureau and seventeen other departments jointly issued the "Three-Year Action Plan for 'Data Elements ×' (2024–2026)" to promote the high-level application of data elements and leverage the multiplier effect of data elements. Currently, more than 80% of the relevant information and data resources in the public domain in China are controlled by various government departments. Therefore, gradually opening public data to society under the premise of ensuring national security should be the basic prerequisite for achieving data element-driven economic development. In view of this, to effectively encourage all social entities to fully develop and utilize government data resources to release the potential value contained in public data, China has actively promoted the gradual opening of public data in recent years. In this context, various regional departments are gradually building their own public data open platforms and using them as key tools to unlock the value of public data. Against this backdrop, whether and how the opening of public data will drive the high-quality development of the capital market has gradually become an important topic of common concern in society.

Based on the aforementioned considerations, this study uses non-financial listed companies on the Shanghai and Shenzhen A-shares from 2007 to 2022 as the research sample. It leverages the exogenous shock of the phased launch of public data open platforms across various provinces, and systematically examines the specific impact and internal mechanisms of public data openness on capital market pricing efficiency through the construction of a multi-period difference-in-differences model, aiming to reveal the potential pathways and underlying logic of data elements enabling high-quality development of the capital market. The study finds that after the launch of the public data open platform in the enterprise's region, stock price synchronicity significantly decreases, confirming the critical role of public data in improving capital market pricing efficiency. Mechanistic analysis reveals that the improvement in capital market pricing efficiency is attributed to market investors' private information collection and the enhancement of listed companies' information disclosure quality induced by public data openness. Heterogeneity analysis indicates that the positive impact of public data openness on capital market pricing efficiency is more significant in samples with higher data quality, lower degree of corporate digital transformation, larger individual investor base, and higher local government transparency.

The contributions of this study are mainly reflected in the following three aspects: First, based on the actual construction of public data open platforms in various regions in recent years, this paper provides micro-level evidence of how data elements empower the high-quality development of the capital market, effectively enriching the literature on how the digital economy optimizes the price formation process in the capital market. Second, from the important perspective of pricing efficiency, which concerns the vital interests of small and medium investors and the healthy development of the capital market, this study enriches the relevant research literature on capital market pricing efficiency from the unique perspective of public data openness. It provides a new governance approach based on promoting the flow of elements to effectively improve pricing efficiency and promote the healthy development of the capital market in the current and future periods. Third, based on the important national strategy of Digital China in the new era, this study offers beneficial insights and positive references for gradually promoting and improving the legislation and practice of public data platforms.

Key Words: data element; public data; pricing efficiency of capital market; stock price synchronicity

JEL Classification: G14, H44

DOI: 10.19616/j.cnki.bmj.2025.01.009

(责任编辑:刘建丽)