

# 公共数据开放与企业数字技术创新

郑志强<sup>1</sup>, 何佳俐<sup>2</sup>, 陈尧<sup>3</sup>

(1. 安徽财经大学 安徽经济社会发展研究院, 安徽 蚌埠 233030;  
2. 安徽农业大学 乡村振兴学院, 安徽 合肥 230036; 3. 华东师范大学 马克思主义学院, 上海 200241)

**摘要:** 数据是数字化、网络化、智能化的基础以及数字经济发展的关键生产要素, 推动公共数据开放是释放数据要素红利的重要举措。虽然已有文献充分探究了数字经济发展或数字技术应用的经济效应, 但是从数据要素视角的研究相对不足。文章以公共数据开放平台上线作为准自然实验, 使用双重差分法实证检验公共数据开放对企业数字技术创新的影响。研究发现, 公共数据开放显著促进了企业数字技术创新, 尤其是数字技术发明创新。机制检验表明, 强化数据要素供给、促进数字化投资以及降低政策不确定性感知是公共数据开放激励企业数字技术创新的主要渠道。进一步分析发现, 公共数据开放对企业数字技术创新的促进作用主要体现在非国有企业以及数字基础设施较为完善的地区; 此外, 公共数据开放质量越高, 企业数字技术创新水平也越高。文章的研究不仅丰富了关于数据生产要素经济价值的理论分析与经验研究, 对于加快发展数据生产力和释放数据要素经济效益也具有重要的政策启示。

**关键词:** 数据要素; 公共数据开放; 数字技术创新

**中图分类号:** F270 **文献标识码:** A **文章编号:** 1001-9952(2025)02-0019-15

**DOI:** 10.16538/j.cnki.jfe.20241115.301

## 一、引言

数字经济时代, 数据不仅仅是信息的载体, 更是与土地、劳动力、资本和技术并列的一种新型生产要素(杨俊等, 2022)。特别地, 政府在履行公职和提供公共服务过程中所创造、采集、获取的数据, 不仅种类丰富、量级庞大, 更是领域内重要的权威数据, 也被视为国家发展的战略性资源(Robinson 等, 2009; 彭远怀, 2023)。公共数据具有明显的公共品属性, 是社会公共利益的重要组成部分, 对不涉及国家安全和隐私保护的公共数据理应开放共享(郑磊, 2015)。2024年9月, 《中共中央办公厅 国务院办公厅关于加快公共数据资源开发利用的意见》发布, 指出要“以促进公共数据合规高效流通使用为主线, 以提高资源开发利用水平为目标, 破除公共数据流通使用的体制性障碍、机制性梗阻, 激发共享开放动力, 优化公共数据资源配置, 释放市场创新活力, 充分发挥数据要素放大、叠加、倍增效应, 为不断做强做优做大数字经济、构筑国家竞争新优势提供坚实支撑”。根据《中国数据产权制度蓝皮书(2022)》披露的数据, 2022年我国的数据产量为 8.1ZB, 占全球数据总量的 10.5%, 位居世界第二。而其中公共数据是占比最大的数据资源,

收稿日期: 2024-06-03

基金项目: 国家社会科学基金项目(23BJY194); 安徽省社会科学创新发展研究课题(2023CX031)

作者简介: 郑志强(1995—), 男, 安徽亳州人, 安徽财经大学安徽经济社会发展研究院助理研究员;

何佳俐(1996—)(通讯作者), 女, 江西九江人, 安徽农业大学乡村振兴学院讲师;

陈尧(1993—), 男, 福建漳州人, 华东师范大学马克思主义学院讲师。

在数据要素共享和开放中具有示范性、引领性和基础性作用(方锦程等, 2023)。在公共数据资源持续增长和不断积累的背景下, 评估公共数据开放的影响对于加快释放数据要素经济效益至关重要。

当前, 构筑自立自强的数字技术创新体系是强化数字中国关键能力的重要措施。以人工智能、大数据、云计算和区块链等为代表的数字技术创新已成为赋能传统企业数字化转型的重要路径, 也是加快培育新质生产力的关键力量。数字经济时代, 数据被视为数字技术创新的基础性资源(赵放等, 2024), 而随着政府部门在公共数据开放领域的持续推进, 公共数据作为新型生产要素汇入企业, 可为企业开展数字技术创新赋能。现实中, 诸多企业都在积极利用公共数据进行数字技术创新。例如, 北京集微科技有限公司将北京公共数据开放平台公开的海量 ICT 产业数据、国家政策、投融资数据、行业分析报告等用于模型训练, 研发出人工智能大模型产品 JiweiGPT, 用以提供文案创意与内容生成、文生图、文生视频等服务, 并成功申请了专利。<sup>①</sup>随着实践的深入, 从学理角度深入探究释放公共数据资源对数字技术创新的作用效果及其影响机制, 对于加快培育新质生产力和推动数字中国建设具有非常重要的理论价值与实践意义。

作为一种非稀缺性资源, 公共数据与劳动力、资本以及土地等传统要素有所不同, 其具有非耗费性、可重复使用性、非竞争性以及高度融合性等特征(黄先海等, 2023b), 从而使得公共数据开放过程能够为使用者的数字技术创新活动提供资源、技术和政策支撑。具体地, 本文借助公共数据开放平台上线的准自然实验场景, 使用双重差分法实证检验公共数据开放对企业数字技术创新的影响。研究发现, 公共数据开放显著促进了企业数字技术创新, 并且主要体现在激励企业数字技术的发明创新上。影响机制检验表明, 强化数据要素供给、促进数字化投资以及降低政策不确定性感知是公共数据开放促进企业数字技术创新的三个主要渠道。进一步分析发现, 公共数据开放对企业数字技术创新的促进作用主要体现在非国有企业以及数字基础设施较为完善的地区; 同时, 公共数据开放质量越高, 企业数字技术创新水平也越高。

本研究的边际贡献主要体现在以下三个方面: 第一, 有效补充了关于数据生产要素经济价值的理论分析与经验研究。数字技术的快速发展和广泛应用衍生出数字经济, 而数据是数字化、网络化、智能化的基础以及数字经济发展的关键生产要素。虽然已有文献充分探究了数字经济发展或数字技术应用的经济效应(赵涛等, 2020; 袁淳等, 2021), 但从数据要素视角的研究相对不足。本文借助公共数据开放平台上线的独特政策场景, 评估了数据生产要素的数字创新效应, 这不仅有效补充了关于数据生产要素经济价值的理论分析与经验研究, 也丰富了评估公共数据开放经济效应的相关文献。第二, 从构建自立自强的数字技术创新体系的战略高度出发, 深入研究数据要素的赋能作用, 将公共数据开放与数字技术创新纳入统一分析框架, 系统梳理并总结了公共数据开放驱动企业数字技术创新的传导机制。具体而言, 本文从资源、技术和政策三个维度入手, 理论分析和实证检验了公共数据开放在强化数据要素供给、促进数字化投资以及降低政策不确定性感知等方面对企业数字创新的激励效应, 进而为更深入理解二者之间的逻辑关系提供了详实的经验证据。第三, 本文研究具有重要的政策启示与现实意义。本文从数字创新视角验证了开放公共数据的价值创造效应, 这不仅为公共部门推动公共数据开放工作和私人部门利用公共数据提供了相关理论指导, 对于加快释放数据要素红利和推动数字经济高质量发展也具有重要的政策启示。

<sup>①</sup> 资料来源: <https://data.beijing.gov.cn/publish/bjdata/appyy/index1.htm>。

## 二、文献回顾、制度背景与理论分析

### （一）文献回顾

本文旨在探究公共数据开放对企业数字技术创新的影响，现有文献中有两支文献与我们的研究直接相关：

第一支文献主要考察数字技术创新的经济效应以及驱动因素。在数字技术创新的经济效应方面，已有研究发现，数字技术创新对于企业全要素生产率（黄勃等，2023）、企业市场价值（陶锋等，2023）、企业出口质量（黄先海等，2023a）以及企业并购绩效（周鹏等，2024）等均具有显著的促进作用。赵源（2023）的研究结果表明，数字技术创新所带来的租金分享更加偏向于技术要素，从而显著扩大了企业内部收入不平等。在数字技术创新的驱动因素上，有学者探究了税收激励（杨鹏等，2023）、市场一体化（胡增玺和马述忠，2023）和知识产权保护（韩先锋等，2024）等对企业数字技术创新的激励效应。

第二支文献聚焦于公共数据开放的影响。早期文献主要从理论分析和案例研究的视角关注公共数据开放的制度基础、决定因素以及挑战等（Safarov 等，2017；Hughes-Crowwick 和 Coronado，2019）。近年来，随着政府部门在开放数据方面的不断实践，评估公共数据开放的影响逐渐成为可能。已有文献主要借助公共数据平台上线的准自然实验，探究公共数据开放的经济价值，发现其对于缓解城市内区域发展差距（方锦程等，2023）、提升企业全要素生产率（彭远怀，2023）以及降低城投债利差（欧阳伊玲等，2024）等具有积极作用。Nagaraj（2022）发现，美国国家航空航天局卫星测绘数据开放使金矿发现率几乎翻了一番，并将新进入者的市场份额从 10% 提高到 25%。陈艳利和蒋琪（2024）考察了公共数据开放对企业创新的影响，发现数据生产要素的资源补给机制和创新意愿的动力刺激机制是公共数据开放促进企业创新的主要渠道。

综合上述文献来看，现有研究主要存在以下缺憾：一是，已有文献主要考察了数字技术创新对企业的影响，但是，探讨数字技术创新的驱动因素和实现路径的研究还相对较少。数字技术创新是数字经济高质量发展的微观基础，如何激励企业数字技术创新是需要进一步深入探究的重要问题。二是，从实证视角评估公共数据开放的微观经济效应的文献相对不足，进一步拓展这方面的研究有利于为充分释放数据要素红利和政府部门推进公共数据开放工作提供参考。鉴于此，本文考察公共数据开放对企业数字技术创新的影响。

### （二）制度背景

加快推进公共数据开放是释放数据要素红利的关键措施，也是推进国家治理体系和治理能力现代化的重要手段（郑磊，2015）。近年来，我国政府部门高度重视公共数据开放工作。2015 年 8 月，《促进大数据发展行动纲要》提出要稳步推动公共数据资源开放共享，加快建设政府数据统一开放平台。2020 年 4 月，《中共中央 国务院关于构建更加完善的要素市场化配置体制机制的意见》首次将数据列为生产要素，并提出建立促进公共数据开放共享和数据资源有效流动的的制度规范。2022 年 1 月，《“十四五”数字经济发展规划》指出要“提升公共数据开放水平，释放数据红利”。2023 年 2 月，《数字中国建设整体布局规划》强调，推动公共数据汇聚利用是畅通数据资源大循环的关键环节。2024 年 9 月，《中共中央办公厅 国务院办公厅关于加快公共数据资源开发利用的意见》发布，这是中央层面首次对公共数据资源开发利用进行系统部署。

公共数据开放平台既是开放获取公共数据的载体，也是连接数据供给侧和需求侧的桥梁，其用途在于公开国民经济和社会发展等各个方面的公共数据。公众可在公共数据开放平台上公开检索、浏览、获取和使用公共数据，并与政府部门进行良性互动。2012 年，上海率先探索推进

公共数据开放工作,上线“上海市政府数据服务网”,该网站是中国第一个省级公共数据开放平台。随后,各省市开始大力推进公共数据开放平台建设,截至 2024 年 7 月,已有 243 个省级和城市的地方政府上线了公共数据开放平台。

### (三)理论分析

根据熊彼特的创新理论,技术创新本质上是企业对现有生产要素的重新组合(约瑟夫·熊彼特, 2009),而数字技术创新实际上是以数字化思维为导向、数据为基础性资源、数字技术为工具而重组生产要素的过程(刘洋等, 2020)。在数字经济时代,数据已经成为与劳动力、资本、土地以及技术等并列的新型生产要素(杨俊等, 2022)。公共数据作为新型生产要素进入企业将变革原有生产函数,并产生“新组合”。一方面,与传统物质生产要素不同,公共数据具有非耗费性、可重复使用性以及非竞争性等特征(黄先海等, 2023b),这使得其在开放后可以在不损失原有价值的情况下被多次或多个生产环节同时使用和共享,从而能够为企业数字技术创新活动持续提供基础性资源投入。另一方面,公共数据具有高度融合性特征,其进入企业后可与劳动力、资本等传统生产要素相融合而共同参与生产过程。这种融合不仅有利于数字创新开发者做出更加科学化和数据驱动型的决策,也可帮助企业进一步全面地掌握市场动态和消费者需求,从而优化投资决策,提高资本利用效率,最终有助于企业开发出更具市场竞争力的数字产品(方锦程等, 2023)。例如,百度公司通过对北京市政府公开的交通流量、道路条件、交通事故记录等公共数据资源的深度学习和不断迭代,研发出了智能驾驶技术。综上,本文提出如下研究假说:

假说 H1: 公共数据开放有利于促进企业数字技术创新。

接下来,基于资源、技术和政策三个方面的支撑,本文分别从强化数据要素供给、促进数字化投资以及降低政策不确定性感知视角详细阐述公共数据开放如何促进企业数字技术创新。

第一,与传统技术创新相比,数字技术创新主要以数字化思维为导向实现对现有技术的改进和突破,其更加强调对数据的可得性和处理能力(刘洋等, 2020)。因此,在数字经济时代,数据被视为数字技术创新的基础性资源(赵放等, 2024)。例如,在机器学习领域,企业通过对大规模标注数据集的训练与学习,不仅可以提高模型的准确性和泛化能力,还能加速人工智能系统的迭代与升级,这种数据驱动的技术进步也能够推动企业在数字领域持续创新。因此,对数据的可得性是决定企业数字技术创新的重要因素。而政府开放的海量结构化和非结构化数据集为企业提供了一个可以低成本便捷获取数据要素的数据源。例如,济南市公共数据开放平台分主题列示了 19 个领域的的数据资源,包括市场监管、教育科技、文化休闲、经贸工商、资源环境、公共安全、交通出行以及气象服务等,数据总量达到 25.8 亿条。在过去,由于公共部门高垄断性、低开放性的缘故,这些数据通常是企业相对难以直接获取的。而如今,企业可在开放平台上查询、下载、使用以及分析这些数据(Hughes-Cromwick 和 Coronado, 2019),从而极大地增加了企业对数据要素的可得性。此外,开放的公共数据通常由各级政府部门提供,具备较高的权威性和可信度,有助于简化企业在数据收集与分析过程中辨别信息真伪的复杂程序,从而提升了数据使用效率(张莉和林安然, 2023)。因此,总的来说,开放具有权威性的公共数据极大地提高了数据资源的普惠化水平,强化了对企业的数据要素供给,进而能够为企业开展数字技术创新活动提供资源支撑。基于上述分析,本文提出如下研究假说:

假说 H2: 公共数据开放能通过强化数据要素供给而促进企业数字技术创新。

第二,在数字化时代,数据被视为一种新型生产要素,其价值不仅体现在数据本身的内容上,更在于对数据的整合、分析与应用能力上。通过有效整合多源数据,并运用先进的数据分析技术,企业能够挖掘出数据中潜在的商业价值,从而优化决策过程(李正辉等, 2024)。因此,在



公共数据开放平台上线后，为了将资源型数据转化为数据生产力，企业必须加大对人工智能、大数据、云计算等数字技术的投资与应用，以帮助其更有效地收集、存储、管理并深度分析这些体量庞大且种类丰富的数据资源(Einav 和 Levin, 2014)。例如，浙江永康农商银行通过使用浙江省医保局、药监局以及财政厅开放的信用、医疗和支付等领域的公共数据，并协同人脸识别、人工智能、云计算、互联网以及区块链等数字技术，开发了基于人脸识别的个人信用无感支付云平台。数字技术具有收敛性和自成长性特征，前者指数字技术对产品边界、部门边界以及组织边界具有模糊效应，从而能够促进不同主体间的深度融合(Nambisan 等, 2017)；后者指数字技术是动态更新的，可根据偏好和需求变化不断地进行迭代和更新(刘洋等, 2020)。因此，企业在数字技术上的投资不仅有利于增强对数据的多边处理和整合能力，进而实现创新资源在不同部门和组织间的有效配置，同时也能帮助企业持续积累数字技术资源和增长数字技术知识，进而为开展数字技术创新提供技术支撑。基于上述分析，本文提出如下研究假说：

假说 H3：公共数据开放能通过促进数字化投资而激励企业数字技术创新。

第三，无论是传统的技术创新还是数字经济时代下的数字技术创新，都是风险相对较高和投入相对较大的投资活动，这类投资活动易受外部政策不确定性因素的影响(Pástor 和 Veronesi, 2009)。特别地，政府在制定相关政策之后，为了应对突发情况或完成其他目标，可能会临时对现有政策做出调整或改变，从而造成政策的前后不一致性，而这种不一致性是制约企业数字技术创新的重要因素。因此，企业在进行日常经营决策尤其是研发决策前，不仅要了解消费者需求和市场竞争状况，还需要准确把握政策意图和导向(饶品贵和徐子慧, 2017)。而公共数据开放平台不仅开放政府部门在提供公共服务过程中所采集和收集的企业数据、行业数据以及市场数据等结构化数据，还开放包括政策制定、政策实施和执行过程等非结构化文本数据。这部分政策数据的公开提高了政府的开放性，降低了企业与政府间的信息不对称，由此相对缓解了企业部门的信息劣势(于文超等, 2020)。并且，政府公开的原始数据中详细地披露了数据的提供方和联系方式，从而极大地强化了问责机制。因此，公共数据开放可以通过降低企业部门的信息劣势和提升社会主体问责的实施效力，降低政策执行过程中的不一致性，从而营造稳定透明可预期的政策环境，最终降低企业政策不确定性感知。而政策不确定性感知是影响企业风险评估和投资决策的重要因素，当企业感知到外部政策不确定性下降时，其对未来具有较为稳定和积极的预期，从而促使企业对具有高风险和高收益特征的数字技术创新活动表现得更加积极(Gulen 和 Ion, 2016)，即促进企业数字技术创新。综上，本文提出如下研究假说：

假说 H4：公共数据开放能通过降低政策不确定性感知程度而促进数字技术创新。

### 三、研究设计

#### (一) 样本选择与数据来源

本文以上市公司作为研究对象，样本时间区间为 2010—2021 年，企业层面数据来源于国泰安数据库，宏观层面数据来源于《中国城市统计年鉴》，各地区公共数据开放平台上线时间根据复旦大学发布的《中国地方政府数据开放报告(城市)》和华东师范大学发布的《中国政府开放数据利用研究报告》以及百度搜索进行交叉验证。本文对数据做如下处理：剔除非正常上市的样本；删除变量取值缺失的样本；对连续变量进行前后 1% 的缩尾处理。

#### (二) 变量度量

被解释变量：数字技术创新(Diginnov)。专利的 IPC 信息详细列示了专利所属的技术领域，因此可根据专利的 IPC 分类号识别企业的数字技术专利，从而度量其数字技术创新。本文参考

陶锋等(2023)的研究,根据国家统计局发布的《数字经济及其核心产业统计分类(2021)》和《国际专利分类与国民经济行业分类参照关系表(2018)》,构建“国民经济行业代码—IPC 分类号—数字经济及其核心产业代码”的对照关系,识别出数字技术专利,并按年度在企业层面进行加总,最终以企业数字技术专利申请数加 1 的自然对数衡量其数字技术创新。

解释变量:公共数据开放(*Dataopen*)。公共数据开放是指政府部门将其在提供公共服务过程中所采集和获取的数据资源统一整合到数据开放平台上,并免费开放给社会主体下载和使用的过程。其中,公共数据开放平台是开放获取公共数据的载体,也是连接开放数据供给者和使用者之间的桥梁。用户可在公共数据开放平台上公开检索、浏览、获取和使用公共数据。因此,在公共数据平台上线的地区,其公共数据可得性将得到极大提升。具体地,本文构建公共数据开放政策的哑变量 *Dataopen*,企业所在城市上线公共数据开放平台的当年及以后, *Dataopen* 赋值为 1;其他情况 *Dataopen* 为 0。

控制变量。借鉴已有关于数字技术创新的相关研究,本文控制如下因素:企业规模(*Ass*)、企业年龄(*Age*)、所有权性质(*Soe*)、总资产净利润率(*Roa*)、现金资产比率(*Cas*)、无形资产比率(*Inv*)、资产负债率(*Lev*)、股权集中度(*Share*)、国有持股比例(*State*)、高管持股比例(*Exe*)、董事会独立性(*Indep*)、资本密集度(*Cap*)、各城市的人均 GDP(*PGDP*)、二产占比(*Second*)和财政收入占比(*Revenue*)。<sup>①</sup>

### (三)实证模型构建

本文以地方政府上线公共数据开放平台为准自然实验,使用双重差分法实证检验公共数据开放对企业数字技术创新的影响。公共数据开放平台上线由当地政府决定,对于企业来说是一个外生事件。具体地,本文将上线公共数据开放平台城市的企业作为实验组,将未上线公共数据开放平台城市的企业作为控制组,并构建如下基准回归模型:

$$Diginnov_{ijt} = \beta_0 + \beta_1 Dataopen_{jt} + \beta_2 Control_{ijt} + Year_t + Firm_i + \varepsilon_{ijt} \quad (1)$$

其中,被解释变量  $Diginnov_{ijt}$  表示  $j$  城市  $i$  企业在  $t$  年的数字技术创新水平,  $Dataopen_{jt}$  表示  $j$  城市在  $t$  年公共数据开放的政策哑变量,  $Control_{ijt}$  为控制变量,  $Year_t$  为年份固定效应,  $Firm_i$  为企业固定效应,  $\varepsilon_{ijt}$  表示残差项,标准误聚类到城市层面。

## 四、实证结果与分析

### (一)基准回归结果

表 1 报告了公共数据开放对企业数字技术创新影响的估计结果。其中,列(1)未加入控制变量,但控制企业和年份固定效应,列(2)在列(1)的基础上加入了企业和城市层面控制变量,上述估计结果均显示, *Dataopen* 的估计系数显著为正,表明公共数据开放和企业数字技术创新具有正相关关系。以列(2)为例, *Dataopen* 的回归系数为 0.0578,表明公共数据开放平台上线后,企业数字技术创新提升了 5.78%,相对于样本期间企业数字技术创新的均值(1.1552)提升了 5.0%。因此,无论在统计意义上还是经济意义上,公共数据开放确实促进了企业数字技术创新,研究假说 H1 成立。

上述结论表明,公共数据开放显著促进了数字技术创新,表现为企业数字技术专利申请量的增加。然而,数字技术专利可进一步细分为数字技术发明专利与数字技术实用新型专利,与后者相比,数字技术发明专利的申请需要具备实质性的技术突破,即更具创新性与新颖性,创新

<sup>①</sup> 限于篇幅,控制变量的详细定义和主要变量的描述性统计结果留存备索。

的质量更高。那么，公共数据开放对数字技术创新的提升作用究竟体现在何种数字技术专利上？为此，本文分别考察公共数据开放对企业数字技术发明创新(*Diginnov\_inv*)和数字技术实用新型创新(*Diginnov\_uti*)的影响。表1列(3)和列(4)的回归结果显示，*Dataopen*的估计系数仅在数字技术发明创新为被解释变量时显著为正，说明公共数据开放显著促进了企业数字技术发明创新，而对数字技术实用新型创新无显著影响，这也验证了公共数据开放对高质量数字技术创新的激励作用。

表1 基准回归结果

	(1) <i>Diginnov</i>	(2) <i>Diginnov</i>	(3) <i>Diginnov_inv</i>	(4) <i>Diginnov_uti</i>
<i>Dataopen</i>	0.0467** (0.0234)	0.0578** (0.0257)	0.0540** (0.0231)	0.0209 (0.0213)
控制变量	未控制	控制	控制	控制
企业和年份固定效应	控制	控制	控制	控制
<i>N</i>	17 123	17 123	17 123	17 123
<i>R</i> <sup>2</sup>	0.7742	0.7770	0.7806	0.7418

注：\*、\*\*、\*\*\*分别表示10%、5%和1%的显著性水平，括号内为聚类到城市层面的标准误。本表及以下各表略去常数项。下表同。

(二) 平行趋势检验

在本文的研究情境下，DID方法运用的前提是实验组和控制组满足平行趋势假设，即上线公共数据开放平台地区和未上线公共数据开放平台地区企业的数字技术创新变化趋势在公共数据开放平台上线前不存在显著差异。为此，本文使用事件研究法对平行趋势假定进行检验，并使用政策开始前一期作为基准期，即估计模型(2)：

$$Diginnov_{ijt} = \alpha + \sum_{K=-8, K \neq -1}^{K=8} \beta_K Treat_{ijt}^K + \gamma Control_{ijt} + Firm_i + Year_t + \varepsilon_{ijt} \quad (2)$$

其中， $Treat^K$ 为上线公共数据开放平台这一“事件”的虚拟变量， $K$ 为相对于事件的年份。具体而言， $Treat^{-2}$ 表示上线公共数据开放平台的前两年， $Treat^0$ 表示上线公共数据开放平台的当年， $Treat^1$ 表示上线公共数据开放平台后的第一年，以此类推。回归结果中 $Treat^k$ 的一系列估计系数反映了公共数据开放对数字技术创新影响的动态效应。图1描述了95%置信区间下 $Treat^k$ 的估计系数，可以看到，在公共数据开放平台上线前， $Treat^k$

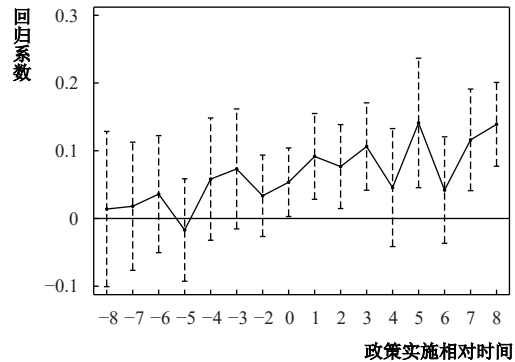


图1 平行趋势检验

的估计系数均不显著，这意味着数字技术创新随时间的变动趋势在实验组和控制组中无明显差异，由此平行趋势假设成立。而在公共数据开放平台上线后， $Treat^k$ 的系数显著为正，表明公共数据开放显著激励了企业数字技术创新。

(三) 稳健性检验<sup>①</sup>

1. 工具变量检验。在数字技术创新较为活跃的地区，当地的数字基础设施和政府对于数字经济发展的政策支撑可能较为完善，此时政府拥有的公共数据资源也相对更多，这会促使政府加

① 限于篇幅，稳健性检验的估计结果留存备案。

快推进公共数据开放,以释放更多的数据经济效益。因此,原文基准回归结果可能存在一定的内生性问题。对此,本文使用工具变量法进行检验。具体地,借鉴欧阳伊玲等(2024)的研究,选取 1984 年各城市邮电业务量作为公共数据开放的工具变量。城市过去使用的通信方式将从技术、基础设施和社会偏好等方面影响样本期内各地区的数字基础设施建设和数字经济政策,而政府越重视数字基础设施建设和数字经济发展就越可能搭建公共数据开放平台,从而满足工具变量选取的相关性要求。同时,历史邮电数据并不直接影响样本期的数字技术创新,因此也满足工具变量选取的外生性要求。此外,考虑到历史邮电业务量仅为截面数据,难以直接作为面板数据的工具变量,因此,本文使用 1984 年各城市人均邮电业务量和样本期间全国邮电业务增长率的交互项作为本文的工具变量(IV)。工具变量第一阶段的回归结果显示,IV 的估计系数在 1% 的水平上显著为正,说明工具变量与解释变量具有较好的相关性。工具变量第二阶段的回归结果显示,Dataopen 的回归系数显著为正,并且 Cragg-Donald Wald F 统计量的值为 383.236,即拒绝了弱工具变量检验的原假设,说明在使用历史邮电数据作为工具变量缓解内生性问题后,本文基准回归结果依然成立。

2. PSM-DID。实际上,公共数据开放平台上线并非严格的准自然实验,并且上线公共数据开放平台的地区和未上线公共数据开放平台的地区间存在诸多可观测和不可观测的特征差异,很可能是这些特征差异带来了数字技术创新的不同,从而影响基准回归结果的可靠性。为此,参考蔡运坤等(2024)的研究,本文使用倾向得分匹配法缓解上述可能存在的样本选择性偏差问题。先以基准模型中企业层面的相关变量为匹配变量对是否上线公共数据开放平台的政策变量进行 logit 回归,并通过 1:4 的近邻匹配法为位于上线公共数据开放平台地区的样本匹配具有共同支撑的控制组样本。再使用匹配后的样本进行回归,估计结果显示 Dataopen 的估计系数显著为正,进一步验证了公共数据开放对企业数字技术创新的促进作用。

3. 预期效应检验。为验证双重差分模型的有效性,本文进一步检验是否存在预期效应。在公共数据开放平台上线之前,当地企业可能会提前预期政府开放公共数据的政策,并据此调整其创新决策,这可能导致基准回归结果存在偏误。为控制企业对政策的预期效应,本文参考崔惠玉等(2023)的做法,将公共数据开放平台上线时间分别提前一年和两年,记为 Dataopen\_pre1 和 Dataopen\_pre2,并重新回归。此时,若 Dataopen\_pre1 和 Dataopen\_pre2 的估计系数显著异于 0,则说明企业在公共数据平台上线前就已形成调整创新决策的预期,进而影响基准回归结果的稳健性。预期效应的回归结果显示,Dataopen\_pre1 和 Dataopen\_pre2 的估计系数均不显著,说明企业在政策实施前并没有形成调整数字创新决策的预期。

4. 安慰剂检验。公共数据开放和数字技术创新的相关性可能存在伪回归问题,即上线公共数据开放平台地区的数字技术创新本身可能就比较,不受是否开放公共数据的影响。因此,本文进一步通过安慰剂检验方法对估计结果进行检验,以确保政策处理效应的有效性。首先,为每个地区随机选取一个年份作为虚拟的公共数据开放平台上线时间点,再随机选取与原实验组相同的城市数。其次,将得到的城市-年份数据与原样本匹配。其中,将随机选取的城市内的企业作为新实验组,其他企业则作为控制组。最后,按照基准回归模型估计政策效果。重复上述置换检验 500 次,得到 500 个随机政策的估计系数,并绘制随机估计系数的概率密度分布图。安慰剂检验的估计系数分布图显示,随机政策的估计系数分布在 0 附近,且绝大部分数值均显著小于真实的估计系数,这进一步验证了真实的公共数据开放政策的有效性。

5. 排除竞争性解释。现实中存在与公共数据开放相关,又可能影响数字技术创新的其他政策,从而会对前文的估计结果造成干扰。其中,省级公共数据开放平台的上线会显著影响城市



层面公共数据开放的作用效果，并且存在部分地级市未上线公共数据开放平台但省级上线的情况，此时企业仍可从省级平台下载和使用所在地级市的公共数据资源。为此，一方面，本文将省级层面的公共数据开放作为控制变量加入基准回归模型；另一方面，剔除地级市未上线平台但省级上线平台的地级市样本，重新对基准模型进行回归。进一步地，本文依次控制“宽带中国”示范城市政策、大数据试验区政策以及“智慧城市”试点政策。之后，本文同时控制上述所有竞争性政策。排除竞争性解释的估计结果显示，上述回归中 *Dataopen* 的估计系数均显著为正，再次验证了公共数据开放对企业数字技术创新的促进作用。

6. 其他稳健性检验。首先，鉴于公共数据平台的开放性特征，本地区企业亦可下载和使用其他地区公共数据平台上的数据资源，这导致其他城市公共数据开放也会影响本地区企业的数字技术创新，如果不控制上述溢出效应可能会导致基准回归结果有偏。为此，本文参考许文立和孙磊(2023)的研究，按照平台上线时点构建与该城市在行政区划上相邻的溢出效应变量<sup>①</sup>并将其作为控制变量加入基准回归中。其次，数字技术从研发到专利申请通常需要一定的时间周期，这会导致公共数据开放对企业数字技术创新的影响具有滞后性，为此我们将解释变量滞后一期重新按基准模型进行估计。最后，企业计数的非负变量通常高度右偏，且在0处有大量的值，这种分布特征对回归分析提出了挑战，使得线性回归效率低下。为了解决这一问题，本文使用面板tobit回归重新估计基准模型。上述回归结果显示，*Dataopen*的估计系数显著为正，进一步验证了基准回归结果是稳健的。

## 五、影响机制检验

接下来，结合理论分析，本文对强化数据要素供给、促进数字化投资以及降低政策不确定性感知这三个机制分别进行实证检验。

### （一）强化数据要素供给

如果公共数据开放能够通过强化数据要素供给而促进企业数字技术创新，那么公共数据开放对企业数字技术创新的促进效果应在公共数据可得性相对较低的样本中更为明显。公共数据主要是政府部门在履行公务和提供公共服务的过程中产生，并由政府收集和获取，是政府部门的垄断性资源。而当企业高管在政府部门具有任职经历时，其对公共数据的可得性可能相对较高；反之，则对公共数据的可得性相对较低。由此本文预期，公共数据开放对企业数字技术创新的促进效应将在高管无政府部门任职经历的样本中更为明显，如果这一预期成立，则可验证假说H2。本文参考陈艳利和蒋琪(2024)的研究，先构建分组虚拟变量 *Con1*，若公司高管（董事长或者总经理）曾任或现任政府官员，则 *Con1* 赋值为0，其他为1；然后，将交互项 *Dataopen* × *Con1* 和 *Con1* 加入基准模型中。表2列(1)的回归结果显示，*Dataopen* × *Con1* 的估计系数显著为正，表明公共数据开放对数字技术创新的促进效应在高管无政府部门任职经历的样本中更为明显，从而在一定程度上验证了公共数据开放能够通过强化数据要素供给而促进企业数字技术创新这一结论，假说H2得证。此外，本文进一步区分高管任职级别以验证上述机制。具体地，构建分组变量 *Con2*，当高管在政府部门无任职经历或任职科长时，*Con2* 赋值为1；而当高管任职处级及以上职务时，*Con2* 赋值为0。随后，将 *Con2* 和交互项 *Dataopen* × *Con2* 加入基准回归模型中，表2列(2)的回归结果显示，*Dataopen* × *Con2* 的估计系数显著为正，这进一步验证了假说H2。

<sup>①</sup> 例如，上海上线公共数据开放平台的当年，其相邻地区也视为上线了平台。

表 2 影响机制检验

	(1) <i>Diginnov</i>	(2) <i>Diginnov</i>	(3) <i>Diginnov</i>	(4) <i>Diginnov</i>	(5) <i>Diginnov</i>	(6) <i>Diginnov</i>
<i>Dataopen</i>	-0.0062 (0.0379)	-0.0031 (0.0418)	0.0249 (0.0348)	0.0292 (0.0265)	0.0114 (0.0389)	0.0134 (0.0362)
<i>Con1</i>	-0.0076 (0.0269)					
<i>Dataopen</i> × <i>Con1</i>	0.0888** (0.0404)					
<i>Con2</i>		0.0060 (0.0286)				
<i>Dataopen</i> × <i>Con2</i>		0.0816* (0.0453)				
<i>Dig1</i>			-0.0178 (0.0262)			
<i>Dataopen</i> × <i>Dig1</i>			0.0629** (0.0314)			
<i>Dig2</i>				0.0074 (0.0309)		
<i>Dataopen</i> × <i>Dig2</i>				0.0895** (0.0399)		
<i>FEPU1</i>					-0.0303 (0.0234)	
<i>Dataopen</i> × <i>FEPU1</i>					0.0658* (0.0384)	
<i>FEPU2</i>						-0.2347*** (0.0745)
<i>Dataopen</i> × <i>FEPU2</i>						0.0685* (0.0383)
<i>N</i>	17 123	17 123	13 734	17 011	13 924	13 924
<i>R</i> <sup>2</sup>	0.7772	0.7771	0.7761	0.7774	0.7971	0.7975

注：本表及以下各表均控制了企业和年份固定效应以及控制变量。

(二) 促进数字化投资

如果公共数据开放能够通过强化数字化投资而促进企业数字技术创新，那么公共数据开放对企业数字技术创新的促进效果应在数字化投资相对较低的企业中更为明显。首先，借鉴范合君等(2023)的做法，以上市公司年末无形资产明细中与数字技术相关的资产占比作为企业数字化投资的第一个替代指标。其次，鉴于企业对未来的展望及其战略特征多体现在其年报中，部分文献使用公司年报中与数字技术相关词汇的词频数来度量企业数字化投资，这一数值越大表示企业对数字化越重视，从而更可能进行数字化投资。具体地，借鉴何德旭等(2024)的研究，使用上市公司年报中数字化词汇词频数作为企业数字化投资的第二个替代指标。在检验部分，首先构建数字技术资产占比和数字化词频数的分组变量*Dig1*和*Dig2*，当其数值小于样本期间的中位数时，将样本划分为数字化投资水平相对较低的组别，*Dig1*和*Dig2*赋值为 1，否则为 0；然后，将*Dig1*和交互项*Dataopen* × *Dig1*以及*Dig2*和交互项*Dataopen* × *Dig2*分别加入基准回归模型中，

表2列(3)和列(4)的回归结果显示,  $Dataopen \times Dig1$ 和 $Dataopen \times Dig2$ 的估计系数均显著为正,表明公共数据开放对数字技术创新的促进效应在数字化投资水平较低的企业中更为明显,从而在一定程度上验证了公共数据开放能够通过强化企业数字化投资进而促进企业数字技术创新的结论。由此,研究假说H3得证。

### (三)降低政策不确定性感知

如果公共数据开放能够通过降低政策不确定性感知而促进企业数字技术创新,那么公共数据开放对企业数字技术创新的促进作用应在政策不确定性感知相对较高的样本中更为明显。本文参考聂辉华等(2020)的做法,使用“词表法”筛选上市公司年报中政策不确定性相关内容文本来构建企业政策不确定性感知指数。具体地,以年报中不确定句子数占总句子数比例和不确定词语数占总词语数的比重作为企业政策不确定性感知程度的两个代理指标,指标的数值越大,意味着企业的政策不确定性感知程度越高。具体地,先构建政策不确定性感知的分组变量 $FEPU1$ 和 $FEPU2$ ,当不确定性句子数占比和词语数占比分别大于样本中位数时,将样本划分为政策不确定性感知程度较高的组别,此时 $FEPU1$ 和 $FEPU2$ 赋值为1;否则为0。然后,将 $FEPU1$ 和交互项 $Dataopen \times FEPU1$ 、 $FEPU2$ 和交互项 $Dataopen \times FEPU2$ 分别纳入基准回归模型中,表2列(5)和列(6)的回归结果显示,  $Dataopen \times FEPU1$ 和 $Dataopen \times FEPU2$ 的估计系数均显著为正,表明公共数据开放对数字技术创新的促进作用主要体现在政策不确定性感知程度较高的样本中,从而在一定程度上验证了公共数据开放能够通过降低政策不确定性感知程度进而激励企业数字技术创新这一结论。由此,研究假说H4得证。

## 六、进一步分析

不同产权性质的企业,其高管在政府部门的任职经历通常存在差异,从而导致对公共数据的可得性显著不同,最终影响公共数据开放对数字技术创新的作用效果。而数字基础设施是政府建设公共数据开放平台和用户访问平台的重要基础,因此数字基础设施的完善程度也会影响本文的基准回归结果。与此同时,不同公共数据开放平台,其公共数据开放质量参差不齐,进而影响公共数据开放的数字创新效应。由此,本文主要从产权、数字基础设施以及公共数据开放质量三个视角开展进一步分析。

### (一)产权的影响

本文根据最终控制权的不同将样本划分为国有企业和非国有企业,以考察公共数据开放对数字技术创新的影响在不同产权性质企业中的异质性。表3列(1)和列(2)的回归结果显示,  $Dataopen$ 的估计系数在非国有企业中显著为正,而在国有企业中并不显著,这说明公共数据开放主要提升了非国有企业的数字技术创新水平,而对国有企业的数字技术创新并没有显著影响。可能的原因在于:一方面,与非国有企业相比,国有企业高管在政府部门的任职经历相对较多,其本身就更易获取公共数据资源(陈艳利和蒋琪,2024)。因此,国有企业对公共数据开放相对不敏感,从而导致公共数据开放对其数字技术创新并没有显著的促进作用。另一方面,国有企业多处于垄断性行业,且更多地承担着稳就业和保障员工福利等社会性责任,其创新动力和意愿不强;相比之下,非国有企业的市场化程度更高,经济效益和市场绩效是非国有企业追求的首要目标,其会抓住公共数据开放的契机将更多的资源投入研发创新活动中,从而更有利于自身数字技术创新水平的提升。

### (二)数字基础设施的影响

公共数据开放平台是下载和获取公共数据的载体,是开放数据供给侧和需求侧沟通互动的节点和桥梁。但是,公共数据开放平台的建设以及企业对公共数据的下载和获取一定程度上依

赖于其所在地区的数字基础设施建设。一方面,良好的数字基础设施建设能够帮助政府部门在提供公共服务过程中获取更多的公共数据,从而有利于推进和完善公共数据开放平台建设。另一方面,当数字基础设施建设较差时,即使政府部门开放大量公共数据,用户对公共数据的可得性也相对较低。因此,公共数据开放对企业数字技术创新的促进效果应在数字基础设施建设较为完善的地区更为明显。为检验上述猜想,本文使用各城市每万人互联网宽带接入端口数作为数字基础设施建设的代理指标,并根据其数值是否大于样本中位数将样本划分为数字基础设施建设较好的组别和较差的组别。表 3 列(3)和列(4)的分组回归结果显示, *Dataopen* 的估计系数仅在数字基础设施建设相对较好的地区中显著为正,这表明公共数据开放主要促进了数字基础设施建设相对完善地区的企业数字技术创新。

表 3 进一步分析

	(1) 国有企业	(2) 非国有企业	(3) 数字基础设施好	(4) 数字基础设施差	(5) <i>Diginov</i>	(6) <i>Diginov</i>
<i>Dataopen</i>	0.0530 (0.0410)	0.0758** (0.0333)	0.0961*** (0.0354)	0.0148 (0.0411)		
<i>Index</i>					0.0032* (0.0019)	
<i>Dataset</i>						0.0387** (0.0174)
<i>N</i>	5782	11282	8142	8446	2965	3364
<i>R</i> <sup>2</sup>	0.8260	0.7508	0.8101	0.8121	0.8664	0.8658

(三)公共数据开放质量的影响

公共数据开放的实践形式是政府将公共数据资源统一整合到开放平台上以供需求者下载和使用。目前,尽管已有大部分地区都上线了公共数据开放平台并提供各种类型的公共数据资源,但各平台运营状况和数据开放质量参差不齐。当公共数据开放平台运营不善或开放数据的范围和数量不足时,企业在获取和利用数据资源的过程中可能面临阻碍,进而削弱公共数据开放对企业数字技术创新的激励作用。基于此,本文使用《中国地方政府数据开放报告(城市)》中披露的中国开放林指数和有效数据集总数作为公共数据开放质量的代理变量,对公共数据开放质量的数字技术创新效应进行检验。其中,开放林指数是对公共数据开放平台建设和运营情况的综合评价,其二级指标包括平台层指数、数据层指数和准备度指数等。而有效数据集总数则是对公共数据平台开放数据量的客观评价。由此,本文主要使用开放林综合指数(*Index*)和有效数据集总量(*Dataset*)进行检验。需要说明的是,《中国地方政府数据开放报告(城市)》中开放林指数和有效数据集总数自 2017 年开始公布,因此匹配数据后样本量会大幅减少,但这可在一定程度验证本文想要考察的问题。表 3 列(5)和列(6)的回归结果显示, *Index*和 *Dataset* 的估计系数均显著为正,表明公共数据开放质量越高,企业数字技术创新水平也越高。

七、结论与启示

本文以各城市公共数据开放平台上线作为准自然实验,使用双重差分法实证检验公共数据开放对企业数字技术创新的影响。得到以下主要结论:(1)公共数据开放显著促进了企业数字技术创新,且该结论通过了一系列稳健性检验。(2)强化数据要素供给、促进数字化投资以及降低政策不确定性感知程度是公共数据开放促进企业数字技术创新的三个主要渠道。(3)公共数



据开放对企业数字技术创新的促进作用主要体现在非国有企业中和数字基础设施建设相对较好的地区。(4)公共数据开放质量越高,企业数字技术创新水平也越高。

数字技术创新是数字经济高质量发展的核心源泉,而公共数据开放对数字技术创新具有促进作用,这充分验证了公共数据要素的创新驱动效应,对于加快发展数据生产力和推动数字经济高质量发展具有重要意义。基于上述研究结论,本文提出如下政策启示:第一,加快推进公共数据开放工作。对于目前仍未上线公共数据开放平台的地区而言,可与已成功运营数据开放平台的地区进行对接,借鉴成功经验,以减少平台建设中可能遇到的障碍。同时,可设立专项资金,增加技术和人员等方面的投入,以争取尽早推动平台落地。对于已经上线平台的地区而言,一方面可通过广泛征集社会公众对不同类型公共数据的需求,优先开放需求量大、商业价值高的公共数据,不断扩大数据开放范围;另一方面,应逐步提高数据更新频率,对数据资源进行分行业、分主题列示,同时提供分类检索导航和高级搜索功能,以满足用户的多样化与个性化需求。第二,公共部门应充分重视公共数据的资源属性,加大海量数据的开放与共享,而企业应积极将高价值数据应用于模型训练和软件开发,从而为开展数字技术创新提供资源供给。同时,企业在下载和使用经济效益较高的数据资源的过程中,也要加强对政策性文本数据的分析和理解,不断优化内部应对机制,主动适应政策变化,尽可能降低政策不确定性对创新决策的影响。此外,企业还应将数字化投资作为战略重点,加快对大数据、人工智能和物联网等数字技术的研发与应用,从而为充分激励数字技术创新提供技术支撑。第三,非国有企业要持续挖掘公共数据资源中潜藏的商业价值,不断开发面向公众的新产品和服务。而国有企业不仅要加大对公共数据的开发与利用,还要坚持市场化改革和强化市场化经营。政府部门要加强数字基础设施建设,尤其是5G、物联网、工业互联网以及千兆光网等新型网络基础设施的规模化部署,从而为企业开展数字技术创新提供底层技术支撑。

#### 参考文献:

- [1]蔡运坤,周京奎,袁旺平.数据要素共享与城市创业活力——来自公共数据开放的经验证据[J].数量经济技术经济研究,2024,(8):5-25.
- [2]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].经济管理,2024,(1):25-46.
- [3]崔惠玉,王宝珠,徐颖.绿色金融创新、金融资源配置与企业污染减排[J].中国工业经济,2023,(10):118-136.
- [4]范合君,吴婷,何思锦.企业数字化的产业链联动效应研究[J].中国工业经济,2023,(3):115-132.
- [5]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].管理世界,2023,(9):124-142.
- [6]韩先锋,肖坚,董明放.创新驱动、产权保护与数字技术创新“量质齐升”——基于多重政策组合新视角[J].产业经济研究,2024,(1):29-41.
- [7]何德旭,申程程,徐子尧.企业数字化、ESG表现与高质量发展[J].经济学动态,2024,(7):21-37.
- [8]胡增玺,马述忠.市场一体化对企业数字创新的影响——兼论数字创新衡量方法[J].经济研究,2023,(6):155-172.
- [9]黄勃,李海彤,刘俊岐,等.数字技术创新与中国企业高质量发展——来自企业数字专利的证据[J].经济研究,2023,(3):97-115.
- [10]黄先海,王瀚迪,孙涌铭,等.数字技术与企业出口质量升级——来自专利文本机器学习的证据[J].数量经济技术经济研究,2023a,(12):69-89.

- [11]黄先海,虞柳明,戴岭.政府数据开放与创新驱动:内涵、机制及实践路径[J].东南学术,2023b,(2):102-113.
- [12]李正辉,许燕婷,陆思婷.数据价值链研究进展[J].经济学动态,2024,(2):128-144.
- [13]刘洋,董久钰,魏江.数字创新管理:理论框架与未来研究[J].管理世界,2020,(7):198-217.
- [14]聂辉华,阮睿,沈吉.企业不确定性感知、投资决策和金融资产配置[J].世界经济,2020,(6):77-98.
- [15]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].世界经济,2024,(2):174-203.
- [16]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].数量经济技术经济研究,2023,(9):50-70.
- [17]饶品贵,徐子慧.经济政策不确定性影响了企业高管变更吗?[J].管理世界,2017,(1):145-157.
- [18]陶锋,朱盼,邱楚芝,等.数字技术创新对企业市场价值的影响研究[J].数量经济技术经济研究,2023,(5):68-91.
- [19]许文立,孙磊.市场激励型环境规制与能源消费结构转型——来自中国碳排放权交易试点的经验证据[J].数量经济技术经济研究,2023,(7):133-155.
- [20]杨俊,李小明,黄守军.大数据、技术进步与经济增长——大数据作为生产要素的一个内生增长理论[J].经济研究,2022,(4):103-119.
- [21]杨鹏,刘航,张柳钦,等.税收激励与企业数字技术创新——来自固定资产加速折旧政策的经验证据[J].财政研究,2023,(9):86-102.
- [22]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].经济学(季刊),2020,(3):1041-1058.
- [23]袁淳,肖土盛,耿春晓,等.数字化转型与企业分工:专业化还是纵向一体化[J].中国工业经济,2021,(9):137-155.
- [24]约瑟夫·熊彼特.资本主义、社会主义与民主[M].吴良健译.北京:商务印书馆,2009.
- [25]张莉,林安然.中国地方政府数据开放与经济发展质量[J].公共行政评论,2023,(2):120-137.
- [26]赵放,蒋国梁,马婉莹.数据要素市场赋能数字产业创新——来自准自然实验的证据[J].经济评论,2024,(3):109-125.
- [27]赵涛,张智,梁上坤.数字经济、创业活跃度与高质量发展——来自中国城市的经验证据[J].管理世界,2020,(10):65-75.
- [28]赵源.数字技术创新、租金分享与企业内部收入不平等[J].南开经济研究,2023,(5):74-92.
- [29]郑磊.开放政府数据研究:概念辨析、关键因素及其互动关系[J].中国行政管理,2015,(11):13-18.
- [30]周鹏,王卓,谭常春,等.数字技术创新的价值——基于并购视角和机器学习方法的分析[J].中国工业经济,2024,(2):137-154.
- [31]Einav L, Levin J. Economics in the age of big data[J]. Science, 2014, 346(6210): 1243089.
- [32]Gulen H, Ion M. Policy uncertainty and corporate investment[J]. The Review of Financial Studies, 2016, 29(3): 523-664.
- [33]Hughes-Cromwick E, Coronado J. The value of US government data to US business decisions[J]. Journal of Economic Perspectives, 2019, 33(1): 131-146.
- [34]Nagaraj A. The private impact of public data: Landsat satellite maps increased gold discoveries and encouraged entry[J]. Management Science, 2022, 68(1): 564-582.
- [35]Nambisan S, Lyytinen K, Majchrzak A, et al. Digital innovation management: Reinventing innovation management research in a digital world[J]. MIS Quarterly, 2017, 41(1): 223-238.
- [36]Pástor L, Veronesi P. Technological revolutions and stock prices[J]. American Economic Review, 2009, 99(4): 1451-1483.

[37]Robinson D, Yu H, Zeller W P, et al. Government data and the invisible hand[J]. *Yale Journal of Law and Technology*, 2009, 11: 160–175.

[38]Safarov I, Meijer A, Grimmelikhuijsen S. Utilization of open government data: A systematic literature review of types, conditions, effects and users[J]. *Information Polity*, 2017, 22(1): 1–24.

## Public Data Openness and Corporate Digital Technology Innovation

Zheng Zhiqiang<sup>1</sup>, He Jiali<sup>2</sup>, Chen Yao<sup>3</sup>

(1. *Institute of Anhui's Economic and Social Development Research, Anhui University of Finance & Economics, Bengbu 233030, China*; 2. *School of Rural Revitalization, Anhui Agricultural University, Hefei 230036, China*; 3. *School of Marxism, East China Normal University, Shanghai 200241, China*)

**Summary:** In the era of the digital economy, data serves not only as a carrier of information but also as a new factor of production. Particularly, data created, collected, and obtained by the government in the course of performing official duties and providing public services are rich in variety, large in scale, and important and authoritative within their fields, and they are also considered strategic resources for national development. With the continuous growth and accumulation of public data resources, assessing the impact of open public data has become a significant topic of interest in both theoretical and practical circles.

This paper leverages the launch of the public data open platform as a quasi-natural experiment and employs the DID method to empirically test the impact of public data openness on corporate digital technology innovation. The study reveals that public data openness significantly promotes digital technology innovation, particularly in the realm of digital technology invention innovation. Mechanism testing indicates that strengthening the supply of data elements, promoting digital investment, and reducing the perception of policy uncertainty are the primary channels through which public data openness stimulates digital technology innovation. Further analysis shows that the promoting effect of public data openness on digital technology innovation is mainly evident in non-state-owned enterprises and areas with relatively comprehensive digital infrastructure. Additionally, the higher the quality of public data openness, the higher the level of digital technology innovation.

The rapid development and widespread application of digital technology have given rise to the digital economy, with data as the foundation of digitization, networking, and intelligence, and a key production factor for the development of the digital economy. While existing literature has thoroughly explored the economic effect of digital economic development or digital technology application, there is a relative dearth of research from the perspective of data elements. This paper utilizes the unique policy scenario of the public data open platform to evaluate the digital innovation effect of data elements, which not only effectively supplements the theoretical analysis and empirical research on the economic value of data as a production factor, but also enriches the knowledge base for evaluating the economic effect of public data openness. It also holds significant practical importance and policy implications for accelerating the development of data productivity and promoting the high-quality development of the digital economy.

**Key words:** data elements; public data openness; digital technology innovation

(责任编辑 石 慧)