

【数字经济】

公共数据开放与城市创新：影响效应与作用机制

郑国强 吕航

摘 要：公共数据作为数字经济时代重要的战略性资源，为政府引导城市创新和生产力变革提供了新动能。以各地政府陆续设立公共数据开放平台作为准自然实验，基于2010—2021年中国287个地级市的面板数据，采用渐进式双重差分模型考察了公共数据开放对城市创新的影响及作用机制。研究发现，公共数据开放能显著促进城市创新，该结论在经过一系列稳健性检验后依然成立；机制检验显示，公共数据开放主要通过降低政策不确定性、改善要素配置扭曲和优化数字发展环境有效驱动城市创新；异质性检验结果表明，在政府数字关注度、市场化程度以及区域一体化程度较高的城市，公共数据开放的创新驱动效应更为显著。进一步研究发现，除增加创新数量外，公共数据开放的创新驱动效应还体现在对创新效率的提升效应以及对周边城市的创新溢出效应。研究结论为数字经济背景下政府培育新质生产力提供了经验证据和政策方向。

关键词：公共数据开放；城市创新；政策不确定性；要素配置；数字经济

作者简介：郑国强，经济学博士，吉林大学东北亚研究中心研究员（长春130015）；吕航，吉林大学东北亚研究中心博士研究生（长春130015）

基金项目：国家社科基金青年项目“数据要素市场化配置的效应评估与政策优化研究”（23CJL021）；吉林省社科基金一般项目“吉林省推进数字产业化和产业数字化研究”（2025B39）

DOI 编码：10.19941/j.cnki.CN31-1957/F.2025.04.001

创新能够催生新产业、新模式、新动能，是培育和发展新质生产力的关键。随着世界百年变局加速演进，经济全球化遭遇逆流，发达国家持续对华实施科技“脱钩断链”与封锁限制。在此背景下，中国亟须提升自主创新能力，真正掌握科技竞争和创新发展的主动权。与此同时，随着新一轮数字革命的深入推进，数据已成为推动创新发展的关键生产要素。熊彼特创新理论将创新定义为通过生产要素和生产方式的不断重组，创造出新技术和新产品的过程^①。数据要素凭借其非竞争性、非排他性以及规模报酬递增等特性，通过推动生产要素的创新性配置，重塑了创新模式与创新体系^②。中国政府高度重视数据要素市场的培育发展，先后出台了《促进大数据发展行动纲要》《关于构建数据基

① 钟腾、罗吉罡、汪昌云：《地方政府人才引进政策促进了区域创新吗？——来自准自然实验的证据》，《金融研究》2021年第5期，第135—152页。

② Baumol W J, “Productivity Growth, Convergence, and Welfare: What the Long-Run Data Show”, in *American Economic Review*, 1986, Vol.76, No.5, pp.1072—1085.

础制度更好发挥数据要素作用的意见》《“数据要素×”三年行动计划（2024—2026年）》等政策文件，旨在通过公共数据开放和数据要素市场化配置，充分释放数据要素的价值与潜能，推动科技自立自强和经济高质量发展。

公共数据资源是政府履行行政职能与提供公共服务过程中产生的战略性基础资源，涵盖经济社会发展诸多领域，构成了国家数据要素体系的核心组成部分^①。相较于其他数据，公共数据具有更强的权威性、通用性和基础性，兼具“生产要素”与“治理要素”双重属性，其开放为创新提供了要素供给和制度支撑，是驱动区域创新发展的“助燃剂”^②。实践中，各地政府持续探索推进公共数据的授权开放与运营。《2023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个省级和地级以上政府上线了公共数据开放平台，各地平台有效数据集增至近25万个，无条件开放的可下载数据集容量超480亿条。由此引发思考：随着政府公共数据开放程度的不断深化，其创新驱动效应是否得到有效发挥？公共数据主要通过哪些途径推动区域创新？

学术界已就数据要素的创新效应展开深入探讨。从创新主体视角来看，韩璐等指出，数据要素能为各行业提供精准信息，企业等创新主体通过深度挖掘和分析海量数据，可发现新的市场机会与潜在需求，进而积极开展创新活动^③。从创新要素视角分析，蔡跃洲和马文君认为，数据要素通过与传统要素的有机融合，能促进资源跨区域、跨界流动，驱动创新性生产要素实现高效配置^④。同时，数据要素的价值唯有在流动中才能体现，部分文献因此探讨了数据交易市场化对创新的影响。一方面，谢地等认为，只有实现数据的市场化配置、降低数据获取成本，才能充分发挥数据在共享增值和科技创新中的重要作用^⑤。另一方面，王海杰和王开阳则认为，由于数据要素在利用过程中会产生价值和产权，因此需建立合理的市场交易机制、明晰数据产权，方能激发数据的创新潜力^⑥。郑国强等的研究发现，数据要素市场化可通过风险承担效应、融资约束效应以及环境信息披露效应，推动企业开展绿色创新活动^⑦。除数据交易外，数据的开放与共享也是促进数据流通、充分释放数据要素价值的重要途径。李长英等认为，政府数据开放能缓解企业与公共制度及市场主体间的摩擦，减少对生产资源的挤占，从而促使创新主体将更多资源

① 张杰、陈志远、杨连星等：《中国创新补贴政策的绩效评估：理论与证据》，《经济研究》2015年第10期，第4—17+33页。

② 梁若冰：《财政分权下的晋升激励、部门利益与土地违法》，《经济学（季刊）》2010年第1期，第283—306页。

③ 韩璐、陈志远、梁玲玲：《数字经济、创新环境与城市创新能力》，《科研管理》2021年第4期，第35—45页。

④ 蔡跃洲、马文君：《数据要素对高质量发展影响与数据流动制约》，《数量经济技术经济研究》2021年第3期，第64—83页。

⑤ 谢地、王荣基、贺城：《数据要素市场化配置赋能企业新质生产力发展》，《经济学动态》2025年第5期，第19—37页。

⑥ 王海杰、王开阳：《数据要素驱动新质生产力发展的机制、挑战与应对措施》，《中国流通经济》2025年第1期，第3—13页。

⑦ 郑国强、张馨元、赵新宇：《数据要素市场化能否促进企业绿色创新？——基于城市数据交易平台设立的准自然实验》，《上海财经大学学报》2024年第3期，第33—48页。

投入高质量创新^①。黄先海和虞柳明认为，政府数据开放有助于创新主体获取数据要素，有效缓解信息不对称问题，降低创新风险，进而提升城市创新水平^②。陈艳利和蒋琪实证考察了政府数据开放对创新的影响^③，但其研究视角主要聚焦于省级政府公共数据开放对企业创新的作用。

相较于既有文献，本研究的边际贡献可能体现在：第一，本研究通过手工搜集各地级市政府公共数据开放平台的设立情况，将公共数据开放与创新的研究视角细化至城市层面，城市作为创新要素创造、集聚和配置的核心空间载体，其数据要素的供给与集聚情况对创新行为的影响更为直接；第二，本研究不仅分析了公共数据开放对城市创新的直接影响效应，还从政策不确定性缓解、要素配置扭曲改善以及数字发展环境优化等维度，探究了公共数据开放促进城市创新的内在机制；第三，本研究还进一步考察了公共数据开放对创新效率的影响及其空间溢出效应，丰富和拓展了公共数据开放创新效应的相关研究。

一、理论分析与研究假设

（一）公共数据开放对城市创新的影响效应

首先，公共数据具有高准确性、高价值性和高应用性，其开放可为城市创新提供高质量的数据要素供给^④。公共数据平台的设立实现了公共数据的线上共享与开放，打破了数据获取的时空壁垒，为数据和知识流通创造了良好平台，有利于降低创新主体的数据获取成本和验证成本，促进其充分挖掘公共数据的潜在价值并开展创新应用^⑤。其次，公共数据覆盖生产制造、财政税收、市场监管等诸多领域和场景，能较好满足不同场景的公共数据需求，使创新主体可从公共数据中提炼更多有效信息。各类创新主体通过分析城市经济发展中的客观数据和因社会需求形成的真实数据，能够识别出数据中隐藏的规律和趋势，科学预测市场对未来创新产品的需求，进而通过需求引致技术研发的迭代升级^⑥。最后，地区的公共数据开放水平越高，越有利于该地区形成开放合作的创新文化理念和创新氛围，进而激发城市的创新活力。若数据要素无法实现开放共享，各类创新主

① 李长英、王曼、张帅：《公共数据开放对民营企业创新的影响》，《广东财经大学学报》2025年第2期，第4—21页。

② 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》，《浙江学刊》2023年第4期，第88—98页。

③ 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》，《经济管理》2024年第1期，第25—46页。

④ 杨新铭、熊华明、刘洪槐：《政策保障、制度完善与市场主体发展——以四川省为例》，《上海商学院学报》2024年第2期，第89—102页。

⑤ 任保平、迟克涵：《数字经济支持我国实体经济高质量发展的机制与路径》，《上海商学院学报》2022年第1期，第3—14页。

⑥ 郑国强、张馨元、赵新宇：《数据要素市场化如何驱动企业数字化转型？》，《产业经济研究》2023年第2期，第56—68页。

体开展的封闭式创新可能是他人已证实无结果的创新项目，进而造成城市创新资源的配置损失^①。公共数据开放平台的设立为数据要素市场的规范发展树立了标杆，可激励各创新主体进行数据共享，通过以数增信、以数搭桥、以数促融等方式实现合作创新与协同创新^②。基于上述分析，本研究提出如下假设。

假设 1：公共数据开放能有效促进城市创新。

（二）公共数据开放对城市创新的作用机制

1. 公共数据开放、政策不确定性与城市创新

政策环境决定创新的风险和机遇，是城市创新过程中的关键变量。政府政策颁布的时滞性、非连续性及政策内容的非一致性，会加剧创新主体面临的政策预期不确定性，进而对其创新决策产生重要影响。经济政策不确定性是指政府在制定和执行宏观经济政策时，因信息不对称等原因，导致创新主体难以预测政策变化和决策结果的情况。这种不确定性会影响创新决策，因为企业需要考虑政策变化对其未来收益的影响。实物期权理论认为，创新投资具有长周期和不可逆特征，政策不确定性会加剧当前创新投资的风险，理性主体会延迟投资期权，以便在未来确定性环境下开展更高回报的创新投资，进而导致城市各类创新主体选择延迟创新活动^③。风险规避理论则认为，在信息不对称和政策不确定的条件下，创新活动的未来前景难以准确掌控，这会诱发各类创新主体的风险厌恶偏好，使其更倾向于投资稳健项目而非高风险的创新活动，从而降低城市整体创新水平^④。

政府公共数据开放可降低各类创新主体面临的“事前”和“事后”政策不确定性，进而有效促进城市创新活动的开展。一方面，公共数据开放能赋能政府治理，基于数据的政府决策可提升政策的针对性和实施效果，降低各项政策内容的非一致性^⑤。同时，公共数据开放可有效缓解政府与创新主体间的信息不对称，提高政府部门的信息透明度。创新主体借助数字技术对公共数据进行分析，可预测未来一定时期的经济走向和政策方向，减少“事前”政策不确定性对城市创新的不利影响^⑥。另一方面，公共数据记录着政府部门的行为轨迹，其开放有利于公众对政策执行情况和实施效果进行监督和问责，规

① 徐霞、吴福象、王兵：《政府大数据赋能城市创新：理论机制与经验证据》，《经济理论与经济管理》2021 年第 12 期，第 42—56 页。

② 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》，《管理世界》2023 年第 9 期，第 124—142 页。

③ 彭远怀、胡军：《政府数据开放与资本区际流动：企业异地投资视角》，《数量经济技术经济研究》2024 年第 10 期，第 89—110 页。

④ 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》，《浙江学刊》2023 年第 4 期，第 88—98 页。

⑤ 王春蕊、李恩极、彭兰凌等：《公共数据开放的就业效应》，《人口研究》2024 年第 6 期，第 3—17 页。

⑥ 陈晓红、李杨扬、宋丽洁等：《数字经济理论体系与研究展望》，《管理世界》2022 年第 2 期，第 208—224 页。

范政府的权力边界，从而降低“事后”政策不确定性对城市创新的不利影响^①。基于上述分析，本研究提出如下假设。

假设 2a：公共数据开放可以通过缓解政策不确定性促进城市创新。

2. 公共数据开放、要素配置扭曲与城市创新

在生产要素增量有限的约束下，劳动和资本等要素的配置方式与效率将对创新活动产生重要影响，要素配置扭曲不利于激发产学研等主体的创新活力，进而阻碍城市创新水平的提升^②。一方面，市场信息不对称导致创新要素配置效率偏离最优水平。在市场机制失灵时，要素资源难以按正常市场需求有效配置，各类创新主体无法根据市场信息准确判断要素的相对丰裕程度和投入结构配比，为获取必要创新要素付出的交易成本将大幅提高，进而对创新研发投入造成转移效应和挤出效应^③。另一方面，在政府掌控部分关键要素供给与配置的情况下，企业为获取稀缺或廉价要素往往会主动与政府官员建立寻租关系，这种以要素资源为纽带的政企联盟会导致要素资源集中于少数企业，抑制多数缺乏政治关联的创新主体的创新积极性^④。

公共数据开放将从要素资源的增量补充、降低市场信息不对称以及减少政府不当干预等方面缓解要素配置扭曲，进而助推城市创新水平提升。首先，公共数据开放有助于各类创新主体利用政府数据挖掘市场潜在需求，激发社会闲置要素活力，通过要素的增量补充和存量优化为创新活动开展奠定基础^⑤。其次，信号传递理论认为，公共数据开放有利于打破要素市场的信息壁垒，实现创新要素的供求精准匹配^⑥。政府公共数据平台涵盖生产数据、商业数据、政务数据以及经济社会统计数据等多维度内容，通过与人力资源部门和金融部门建立数据共享机制，可降低各类创新主体与这两部门之间的信息不对称，使创新主体能精准描绘城市人才供给情况并提高创新资金供给效率^⑦。最后，公共数据反映了政府配置要素资源的过程和成效，公共数据开放平台的设立有利于加强公众对政府行为的监督，倒逼政府提升治理能力、减少对要素市场的不当干预，进而提升城市

① Schumpeter A J, *The Theory of Economic Development*, Cambridge, MA: Harvard University Press, 1912.

② Glaeser E L, Kominers S D, Luca M, et al., “Big Data and Big Cities: The Promises and Limitations of Improved Measures of Urban Life”, in *Economic Inquiry*, 2018, Vol.56, No.1, pp.114–137.

③ Zhang Q, Yang L T, Chen Z, et al., “A Survey on Deep Learning for Big Data”, in *Information Fusion*, 2018, Vol.42, pp.146–157.

④ 杨秀云、韩奇：《公共数据开放能提升企业全要素生产率吗？》，《证券市场导报》2023年第12期，第18—30页。

⑤ 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——基于建立政府数据开放平台的准自然实验》，《经济管理》2024年第1期，第25—46页。

⑥ 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》，《管理世界》2023年第9期，第124—142页。

⑦ 鲁元平、张克中、欧阳洁：《土地财政阻碍了区域技术创新吗？——基于267个地级市面板数据的实证检验》，《金融研究》2018年第5期，第101—119页。

要素配置效率与创新水平^①。基于上述分析,本研究提出如下假设。

假设 2b: 公共数据开放可以通过改善要素配置扭曲促进城市创新。

3. 公共数据开放、数字发展环境与城市创新

数字经济作为具有渗透性、替代性以及协同性的新型经济形态,可在促进创新知识溢出、重塑创新模式、提供创新技术保障等方面发挥积极作用,已成为城市创新发展的关键动能^②。数字经济的渗透性可加速创新信息的传播和溢出,不仅能大幅降低城市各类创新主体的信息获取成本,还能使不同创新主体在信息溢出过程中形成互补性投资与合作,进而推动城市创新水平提升^③。数字经济的替代性表现为:摩尔定律、吉尔德定律以及金帆定律的叠加,导致数字技术产品价格持续下降、类型不断更迭,既降低了创新主体的数字创新成本和创新周期,又加剧了数据要素对传统要素的替代,重塑了组织创新、生产创新以及商业创新的逻辑与模式^④。数字经济的协同性表现为:数据要素不仅能推动其他创新要素的协同应用,还能使各类创新主体借助线上平台实现跨区域、跨行业、跨方向的协同合作,促进创新活动由地理空间集聚向数字空间集聚转变^⑤。

公共数据是驱动数字发展环境优化的关键要素^⑥。公共数据兼具数字技术和公共资源双重特质,通过开放能够优化城市数字发展环境,进而为城市创新发展提供支撑。从技术特质来看,公共数据开放使各类创新主体可获取更多数据生产要素,加速数据要素的价值释放,通过促进数实融合推动数字产业化和产业数字化,为城市数字经济和创新发展奠定基础^⑦。从公共资源特质来看,公共数据具备公共基础设施的基础性和普惠性,可降低创新数据要素的获取门槛、打破创新数据垄断,让数据要素开放惠及所有创新主体;同时,公共数据开放还具备公共基础设施的非竞争性和非排他性,相关数据可由不同创新主体同时重复使用。此外,公共数据开放还有利于推动政府和企业的数字化转型,吸引更多数字化人才涌入,进而为城市数字经济发展和创新水平提升创造条件^⑧。基于上述分析,本研究提出如下假设。

-
- ① 张吉昌、龙静:《政府数据开放何以赋能企业创新?》,《现代经济探讨》2024 年第 4 期,第 29—43+59 页。
- ② 李三希、李嘉琦、刘小鲁:《数据要素市场高质量发展的内涵特征与推进路径》,《改革》2023 年第 5 期,第 29—40 页。
- ③ Jones C I, Tonetti C, “Nonrivalry and the Economics of Data”, in *American Economic Review*, 2020, Vol.110, No.9, pp.2819—2858.
- ④ 张亚豪、李晓华、刘尚文:《数据要素赋能服务型制造发展:场景应用、作用机制与政策建议》,《改革》2024 年第 1 期,第 69—81 页。
- ⑤ Akcigit U, Liu Q, “The Role of Information in Innovation and Competition”, in *Journal of the European Economic Association*, 2016, Vol.14, No.4, pp.828—870.
- ⑥ 梁若冰:《财政分权下的晋升激励、部门利益与土地违法》,《经济学(季刊)》2010 年第 1 期,第 283—306 页。
- ⑦ Bloom N, Bond S, Van Reenen J, “Uncertainty and Investment Dynamics”, in *The Review of Economic Studies*, 2007, Vol.74, No.2, pp.391—415.
- ⑧ Gulen H, Ion M, “Policy Uncertainty and Corporate Investment”, in *The Review of Financial Studies*, 2016, Vol.29, No.3, pp.523—564.

假设 2c：公共数据开放可以通过优化数字发展环境促进城市创新。

二、研究方法 with 数据说明

（一）模型设定

基于上述分析，为考察公共数据开放对城市创新的影响效应，本研究设定渐进式双重差分模型如下：

$$Innov_{it} = \alpha + \beta Open_{it} + \lambda Control_{it} + \mu_i + \mu_t + \varepsilon_{it} \quad (1)$$

式中： i 、 t 分别表示城市和年份；被解释变量 $Innov_{it}$ 表示 i 城市第 t 年的创新水平；核心解释变量 $Open_{it}$ 表示 i 城市第 t 年是否设立公共数据开放平台； $Control_{it}$ 是一组控制变量； α 、 β 、 λ 表示估计系数； μ_i 代表城市固定效应； μ_t 代表时间固定效应； ε_{it} 为随机扰动项。

（二）变量说明

1. 被解释变量

城市创新水平（ $Innov$ ）。专利数量是衡量城市创新水平的重要维度，各城市的专利数量均在统一衡量口径下计算，使不同城市间的专利数据具有可比性。本研究借鉴姚战琪的研究^①，以各城市专利授权数的对数衡量城市创新水平，并在稳健性检验中以专利申请数的对数作为城市创新水平的替代指标。

2. 核心解释变量

公共数据开放程度（ $Open$ ）。借鉴方锦程等的研究^②，以地级市政府设立公共数据开放平台作为准自然实验。将设立数据开放平台的城市取值为 1，作为实验组；将未建立公共数据开放平台的城市取值为 0，作为对照组。将设立平台的城市平台设立后的年份虚拟变量赋值为 1，设立平台的城市在平台设立前及未设立平台的城市赋值为 0，并以时间虚拟变量和地区虚拟变量的交互项衡量公共数据开放的影响效应。

3. 控制变量

根据已有文献，加入可能影响城市创新的控制变量。具体包括：①政府科教支持力度（ Edu ），用政府科学技术支出与教育支出占 GDP 的比重衡量；②人力资本水平（ HR ），以普通高等学校在校生数所占总人口的比重表示；③产业结构（ $Indust$ ），用第二产业增加值占 GDP 的比重衡量；④人口密度（ Pop ），以每平方公里内的常住人口数（万人）表示；⑤经济发展水平（ $CGDP$ ），以城市 GDP 的对数表示。

（三）数据来源与描述性统计

本研究以 2010—2021 年中国 287 个地级市为研究对象，剔除核心变量缺失严重的城市样本。其中，城市层面的公共数据开放平台数据通过新闻报道和站长工具手工搜集获得，并与复旦大学发布的《中国地方政府数据开放报告》进行交叉认证；城市专利数据来源于

① 姚战琪：《中国服务业开放对区域创新的影响》，《改革》2020 年第 1 期，第 48—58 页。

② 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》，《管理世界》2023 年第 9 期，第 124—142 页。

中国研究数据服务平台（CNRDS）；其他数据来源于各城市统计年鉴和 EPS 数据库。为规避异常值的影响，对连续变量进行双边缩尾处理。主要变量的描述性统计结果如表 1 所示。

表 1 相关变量描述性统计结果

变量		样本量	平均值	标准差	最小值	最大值
被解释变量	<i>Innov</i>	3 444	0.325	0.472	0.001	3.366
解释变量	<i>Open</i>	3 444	0.088	0.283	0	1
控制变量	<i>Edu</i>	3 444	0.037	0.019	0.008	0.186
	<i>Indust</i>	3 444	1.202	0.632	0.187	9.197
	<i>HR</i>	3 444	0.019	0.026	0	0.195
	<i>Pop</i>	3 444	0.380	0.256	0.031	1.506
	<i>CGDP</i>	3 444	11.971	0.946	9.250	15.274

三、实证分析

（一）基准回归分析

渐进式双重差分模型下的基准回归结果如表 2 所示。表 2 列（1）为双向固定效应下仅加入核心解释变量的回归结果，*Open* 的回归系数在 1% 的水平下显著为正。表 2 列（2）~（6）是逐步加入控制变量后的回归结果，可见 *Open* 的回归系数始终显著为正，表明公共数据开放对城市创新具有显著促进作用。如前文所述，公共数据开放可为创新活动开展提供高质量的数据要素供给和方向指引，有利于形成开放合作的创新文化与氛围，进而助力城市创新水平提升。本研究假设 1 得以验证。

表 2 基准回归结果

变量	<i>Innov</i>					
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Open</i>	0.135*** (0.033)	0.135*** (0.033)	0.136*** (0.033)	0.128*** (0.032)	0.127*** (0.031)	0.117*** (0.031)
<i>Edu</i>	—	0.056 (0.795)	0.080 (0.801)	1.099 (0.805)	1.097 (0.803)	2.757*** (0.912)
<i>HR</i>	—	—	0.206 (0.967)	0.263 (0.976)	0.296 (0.975)	0.406 (0.947)
<i>Indust</i>	—	—	—	0.453*** (0.120)	0.451*** (0.119)	0.139 (0.137)
<i>Pop</i>	—	—	—	—	0.0175 (0.044)	0.0261 (0.043)
<i>CGDP</i>	—	—	—	—	—	0.208*** (0.045)
固定效应	是	是	是	是	是	是
R^2	0.443	0.443	0.443	0.450	0.450	0.463
样本量	3 444	3 444	3 444	3 444	3 444	3 444

注：*、** 和 *** 分别表示系数在 10%、5% 和 1% 的水平下显著；括号内为聚类到城市层面的标准误；下同。

(二) 平行趋势检验

渐进式双重差分模型要求政策实施前实验组和对照组保持同步变化趋势，即符合平行趋势检验假设。参考 Beck 等的方法^①，本研究设定如下模型进行平行趋势检验：

$$Innov_{it} = \rho_0 + \rho_k \sum_{k=-6}^4 Data_{it} + \lambda_j Control_{it} + \mu_i + \mu_t + \varepsilon_{it} \quad (2)$$

式中：变量 *Data* 表示政府数据平台设立的时间虚拟变量。

为规避时间虚拟变量间的多重共线性，以政府数据开放当期作为基期并在回归中剔除。由图 1 的检验结果可见，在公共数据平台设立前，实验组和对照组的城市创新水平无显著差异，即城市公共数据开放符合平行趋势假设；在公共数据开放平台设立后，相关回归系数显著为正且不断提升，说明公共数据开放平台的设立有利于提升城市创新水平，且随着平台的逐步完善，其创新促进效应呈持续增强趋势。

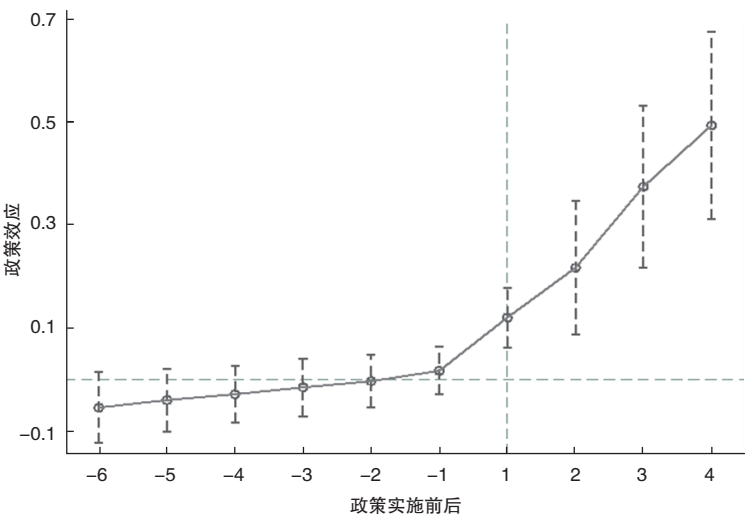


图 1 平行趋势检验结果

(三) 稳健性检验

1. 考虑政策预期效应

为尽可能规避各类创新主体预期对政策真实效应产生的混淆，本研究采用反事实方法，将政策实施年份提前并纳入模型。表 3 列（1）结果显示，核心解释变量 *Open* 的回归系数显著，而政策预期变量 *BOpen* 的回归系数不显著，表明公共数据开放对城市创新水平的提升作用未受到预期效应影响。

① Beck T, Levine R, Levkov A, “Big Bad Banks? The Winners and Losers from Bank Deregulation in the United States”, in *The Journal of Finance*, 2010, Vol.65, No.5, pp.1637—1667.

表 3 稳健性检验结果

变量	Innov						
	考虑预期效应	排除其他政策干扰	替换被解释变量		筛选样本		考虑数据开放质量
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<i>Open</i>	0.115*** (0.029)	0.247*** (0.043)	0.068*** (0.023)	0.231*** (0.039)	0.108*** (0.031)	0.094*** (0.027)	0.480*** (0.091)
<i>BOpen</i>	0.002 (0.018)	—	—	—	—	—	—
<i>Inpolicy</i>	—	0.132*** (0.046)	—	—	—	—	—
<i>Inform</i>	—	0.099*** (0.035)	—	—	—	—	—
控制变量	是	是	是	是	是	是	是
固定效应	是	是	是	是	是	是	是
<i>R</i> ²	0.463	0.693	0.471	0.843	0.460	0.491	0.184
样本量	3 444	3 444	3 444	3 444	3 396	3 264	300

2. 排除其他政策的干扰效应

在样本考察期内，国家陆续出台了“创新型城市”“信息惠民工程”等影响地区数字经济发展和城市创新转型的政策，可能对本研究估计结果产生干扰。为此，本研究在基准回归中同时加入“创新型城市”和“信息惠民工程”的虚拟变量 *Inpolicy* 和 *Inform*，以控制其他政策对结果的影响。表 3 列（2）结果显示，在纳入其他创新驱动政策之后，核心解释变量 *Open* 的回归系数仍然显著为正。

3. 替换被解释变量

为保证回归结果的稳健性，本研究对城市创新水平重新测度。一方面，采用专利申请数的对数作为城市创新指标进行重新估计；另一方面，采用复旦大学发布的《中国城市和产业创新力报告》中的创新指数作为城市创新的替代变量。表 3 列（3）和列（4）的回归结果显示，重新测度城市创新变量后，核心解释变量 *Open* 的估计系数仍然显著为正，主要结论未发生改变。

4. 筛选样本

考虑到直辖市和省会城市在经济发展与政策供给方面的优势，本研究先后剔除直辖市样本和省会城市样本进行稳健性检验，回归结果如表 3 列（5）和列（6）所示。可以发现，*Open* 的回归系数依然显著为正，再次验证了本研究实证结果的稳健性。

5. 考虑数据开放质量

前文仅考察城市公共数据是否开放，未涉及数据开放平台的数据质量。高质量的数据供给是城市创新的基础，但各城市政府开放平台的数据质量参差不齐，可能对城市创新产生差异化影响。基于此，本研究将利用复旦大学《中国地方政府数据开放报告》中的中国开放树林综合指数来衡量公共数据开放质量，进一步考察其对城市创新的影响。

表 3 列（7）展示了双向固定效应模型下的估计结果，可见核心解释变量 *Open* 的回归系数仍在 1% 水平上显著为正，即公共数据平台的开放质量越高，其对城市创新的促进作用越明显。

（四）作用机制检验

基准回归结果验证了公共数据开放对城市创新的影响效应，接下来本研究从政策不确定性、要素配置扭曲以及数字发展环境三个维度，借鉴江艇的两步检验法^①，对公共数据开放影响城市创新的作用机制进行实证验证。作用机制检验模型设定如下：

$$Innov_{it} = \varphi_0 + \varphi_1 W_{it} + \lambda_j Control_{it} + \mu_i + \mu_t + \varepsilon_{it} \quad (3)$$

$$W_{it} = \delta_0 + \delta_1 Open_{it} + \lambda_j Control_{it} + \mu_i + \mu_t + \varepsilon_{it} \quad (4)$$

式中： W_{it} 为机制变量，代表政策不确定性、要素配置扭曲和数字发展环境等变量；其他变量与模型（1）保持一致。实证结果如表 4 所示。

表 4 机制检验结果

变量	<i>Innov</i>	<i>CEPU</i>	<i>Innov</i>	<i>FDIST</i>	<i>Innov</i>	<i>DIGE</i>
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Open</i>	—	-0.066* (0.037)	—	-0.508** (0.205)	—	0.051** (0.025)
<i>CEPU</i>	-0.243*** (0.058)	—	—	—	—	—
<i>FDIST</i>	—	—	-0.008** (0.003)	—	—	—
<i>DIGE</i>	—	—	—	—	0.081*** (0.024)	—
固定效应	是	是	是	是	是	是
控制变量	是	是	是	是	是	是
R^2	0.469	0.288	0.686	0.098	0.456	0.687
样本量	3 444	3 444	3 408	3 408	3 444	3 444

1. 降低政策不确定性

参照赵春明的做法^②，构建城市层面的政策不确定性指数 *CEPU*。其基本思路是以 Baker 的全国层面经济政策不确定性指数为基准，以各城市财政支出总额占全国财政支出总额的份额为权重，对政策不确定性指数进行加权平均，计算出城市层面的政策不确定性指数。表 4 列（1）和列（2）报告的是基于政策不确定性的机制检验结果。其中，列（1）报告的是政策不确定性对城市创新的回归结果，可见 *CEPU* 的回归系数显著为负，说明政策不确定性的增加会制约城市创新发展。列（2）报告的是公共数据开放对政策不确定性的回归结果，可见 *Open* 的回归系数显著为负，表明公共数据开放可有效降低政策不

① 江艇：《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》2022 年第 5 期，第 100—120 页。
② 赵春明、范雅萌、熊珍琴：《贸易政策不确定性对中国地区产业结构升级的影响》，《亚太经济》2020 年第 5 期，第 116—125+152。

确定性,存在“公共数据开放→政策不确定性降低→城市创新水平提升”的间接传导机制,本研究假设 2a 得以验证。如前文所述,公共数据开放在提高政企间信息感知的同时,还能对政府政策实施进行有效监督和效果评估,赋能政府治理并降低政策不确定性,进而助力城市创新活动开展。

2. 缓解要素配置扭曲

参照赵新宇和郑国强的做法^①,构建城市层面的要素配置扭曲指数 $FDIST$ 。其基本思路是通过生产函数法测算资本配置扭曲指数 $DISK = Mp_K/r$,以及劳动力配置扭曲指数 $DISL = Mp_L/\omega$,其中 Mp_K 和 Mp_L 分别为资本和劳动力的边际产出, r 和 ω 分别为利率和工资,则要素配置扭曲指数 $FDIST = DISL^{\frac{\alpha}{\alpha+\beta}} \times DISK^{\frac{\beta}{\alpha+\beta}}$ 。表 4 列(3)和列(4)报告的是基于要素配置扭曲的机制检验结果。其中,列(3)报告的是要素配置扭曲对城市创新的回归结果,可见 $FDIST$ 的回归系数显著为负,说明要素配置扭曲不利于城市创新发展。列(4)报告的是公共数据开放对要素配置扭曲的回归结果,可见 $Open$ 的回归系数显著为负,表明公共数据开放可有效缓解要素配置扭曲,存在“公共数据开放→要素配置扭曲缓解→城市创新水平提升”的间接传导机制,本研究假设 2b 得以验证。如前文所述,公共数据开放有利于打破要素市场的信息壁垒,激发社会闲置要素活力,减少政府对要素市场的不当干预,进而提升城市要素配置效率和创新水平。

3. 优化数字发展环境

借鉴赵涛等的做法^②,从数字基础设施建设和数字交易两方面构建数字发展环境指数 $DIGE$ 。其中,数字基础设施建设采用城市互联网发展指数衡量,数字交易用蚂蚁金服公布的数字普惠金融指数衡量,再通过主成分分析法得到数字发展环境指数。表 4 列(5)和列(6)报告的是基于数字发展环境的机制检验结果。其中,列(5)报告的是数字发展环境对城市创新的回归结果,可见 $DIGE$ 的回归系数显著为正,说明城市数字发展环境的优化有助于赋能城市创新发展。列(6)报告的是公共数据开放对数字发展环境的回归结果,可见 $Open$ 的回归系数显著为正,表明公共数据开放有助于优化完善数字发展环境,存在“公共数据开放→数字发展环境优化→城市创新水平提升”的间接传导机制,本研究假设 2c 得以验证。如前文所述,公共数据具有数字技术和公共资源双重特性,既能通过促进数实融合推动数字产业化和产业数字化,为城市数字经济奠定基础,又能有效降低创新数据要素的获取门槛、打破创新数据垄断,为城市数字经济和创新发展提供支撑。

(五) 异质性分析

上述研究验证了公共数据开放对城市创新的影响效应及作用机制,但不同城市在政

① 赵新宇、郑国强:《劳动力市场扭曲如何影响城市全要素生产率?——基于配置扭曲和价格扭曲双重视角的实证研究》,《经济问题探索》2021 年第 6 期,第 112—124 页。

② 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》,《管理世界》2020 年第 10 期,第 65—76 页。

府支持、市场化发展以及一体化程度方面存在的差异，可能导致公共数据开放水平和应用程度不同，进而使公共数据的创新驱动效应存在区域层面的异质性。基于此，本研究将从政府数字关注度、市场化程度以及区域一体化水平三个方面，考察公共数据开放对城市创新的异质性影响，实证结果如表 5 所示。

表 5 异质性检验结果

变量	政府数字关注度		市场化程度		区域一体化水平	
	高数字关注度	低数字关注度	高市场化程度	低市场化程度	高一体化水平	低一体化水平
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Open</i>	0.046* (0.027)	0.028 (0.078)	0.070** (0.032)	0.036 (0.043)	0.121** (0.048)	0.094*** (0.031)
控制变量	是	是	是	是	是	是
固定效应	是	是	是	是	是	是
R^2	0.430	0.444	0.497	0.421	0.527	0.456
样本量	1 558	1 785	1 722	1 722	1 068	2 376

首先，公共数据开放的创新效应可能因地方政府的数字关注程度不同而存在异质性。政府对数据要素和数字技术发展的关注程度，不仅影响公共数据的开放与运营质量，还会影响政府的数字化创新偏好和创新扶持力度，进而导致公共数据开放的创新效应存在异质性。借鉴欧阳伊玲等对政府数字关注的衡量思路^①，本研究先构建有关数据要素和数字技术的词汇池，再借助 Python 软件对地级市的政府工作报告进行分词处理，通过统计相关词频数量量化政府数字关注度。根据政府关注度的中位数，将样本分为高政府数字关注度和低政府数字关注度两组，分组回归结果如表 5 列（1）和列（2）所示。可以发现，高政府数字关注度城市的公共数据开放具有更显著的创新驱动作用，而低政府数字关注度城市的公共数据开放未能有效促进城市创新。地方政府是公共数据开放的主体，其对数据要素和数字技术越重视，就越可能推动公共数据的汇聚融合、共享开放和开发利用，为数字化创新提供更多的财政支持和政策优惠，从而更好发挥公共数据开放的创新效应。

其次，公共数据开放的创新效应可能因市场化程度不同而存在异质性。一方面，数据要素在与劳动、资本和技术等要素相结合时，能更好地释放价值并驱动创新发展^②；另一方面，市场化发展可降低创新风险，为创新活动提供更多市场机遇。本研究采用王小鲁等编制的市场化指数表征区域市场化程度^③，2020—2021 年数据按年平均增长幅度预测得出。根据市场化程度的中位数，将样本分为高市场化程度和低市场化程度两组，分

① 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》，《世界经济》2024 年第 2 期，第 174—203 页。
② 武宵旭、任保平：《数字经济背景下要素资源配置机制重塑的路径与政策调整》，《经济体制改革》2022 年第 2 期，第 5—10 页。
③ 王小鲁、樊纲、胡李鹏：《中国分省份市场化指数报告（2018）》，北京：社会科学文献出版社，2019 年。

组回归结果如表 5 列（3）和列（4）所示。可以发现，在市场化程度较高的城市，*Open* 的回归系数显著为正，而在市场化程度较低的城市，*Open* 的回归系数不显著。这表明公共数据只有在市场化水平较高的城市才能有效驱动创新发展。较高的市场化水平加速了创新资源在城市的流动配置，促进了数据要素的融合渗透，有效降低了创新主体间的信息不对称，缓解了数据获取成本，为公共数据驱动创新发展提供了良好市场环境。

最后，公共数据开放的创新效应可能因区域一体化程度不同而存在异质性。互联互通是数据要素价值实现的重要方式，城市所在区域的一体化程度越高，越有利于各城市政务大数据体系实现互联互通，进而扩大公共数据供给规模，增强数据共享应用和开发利用能力，推动知识共享和创新合作，促进公共数据开放质量提升和创新效应发挥。本研究将京津冀、长三角、珠三角、成渝和长江中游城市群所在城市定义为高区域一体化组别，其他城市划分为低区域一体化组别并进行分组回归，结果如表 5 列（5）和列（6）所示。可以发现，*Open* 的回归系数在两组均显著为正，但一体化水平较高的城市估计系数更大，说明城市所在区域的一体化程度越高，公共数据开放的创新驱动效应越明显。区域一体化在资源整合、经济辐射、基础设施建设方面相对协调，打破了阻碍数据要素流动的藩篱，为数据要素自由流动提供了创新基础。

四、进一步分析

（一）公共数据开放与创新效率

在加快实现高水平科技自立自强的背景下，创新水平的提升不仅体现在数量和规模上，还应反映在质量和效率上。为此，本研究参照程惠芳和陆素俊的做法^①，采用 Malmquist 指数法测度各城市创新效率。其中，创新效率的产出指标为城市发明专利授权量，投入指标包括城市科学技术支出、教育支出和普通高等学校在校人数等。回归结果如表 6 列（1）所示，可见 *Open* 的回归系数在 5% 的水平上显著为正，表明公共数据开放不仅能增加城市创新数量，还能够提升城市创新效率，从而实现高层次的创新驱动效应。

表 6 进一步分析回归结果

变量	创新效率	空间溢出效应
	(1)	(2)
<i>Open</i>	0.056** (0.028)	0.587*** (0.071)
控制变量	是	是
固定效应	是	是
R^2	0.023	0.590
样本量	3 154	3 444

① 程惠芳、陆素俊：《知识资本对工业企业全要素生产率影响的实证分析》，《经济研究》2014 年第 5 期，第 174—187 页。

（二）公共数据开放的溢出效应

公共数据开放具有包容性和普惠性，基于电子政务和数字技术的普及运用，无论是城市辖区内还是其他地区的创新主体，均可以自由获取并使用开放数据。这使得公共数据开放平台不仅能服务本地创新主体，还能通过公共数据的跨区域配置和创新合作激发其他地区的创新活力，进而呈现显著的创新溢出效应。本研究参考余典范等的做法^①，以样本城市相邻各城市的专利授权量均值表征样本城市的创新溢出效应，并作为被解释变量代入模型（1）进行回归。表6列（2）回归结果显示，核心解释变量 *Open* 的回归系数在1%水平上显著为正，即公共数据开放具有较好的辐射性和带动性，能够对周边城市产生正向的创新溢出效应，有效推动周边城市的创新发展。

五、结论与建议

公共数据开放有利于激发数据要素活力，为城市创新发展和新质生产力培育筑牢基础。本研究将各地级市建立公共数据开放平台作为政府推动公共数据开放的准自然实验，基于2010—2021年中国地级市面板数据，采用渐进式双重差分法考察公共数据开放对城市创新的影响效应和作用机制，得到如下结论。第一，公共数据开放有助于提升城市创新水平，经过考虑政策预期、排除其他政策干扰、替换变量、筛选样本等一系列稳健性检验后，研究结论依然稳健。第二，降低城市创新主体的政策不确定性、改善城市要素配置扭曲状况以及优化数字发展环境，是公共数据开放促进城市创新的重要渠道。第三，在政府数字关注度较高、市场化程度较高以及区域一体化程度较高的城市，公共数据开放的创新驱动效应更为显著。进一步分析发现，公共数据开放除增加创新数量外，还能提升城市创新效率，并对周边城市的创新活动产生溢出效应。

本研究结论对推动政府公共数据开放和提升城市创新能力具有重要政策启示。第一，应深化公共数据开放的体制机制改革，加速公共数据授权运营管理，为城市创新提供强有力的数据支撑和制度保障。通过制定更加明确、细致的公共数据开放政策，明确授权运营的条件、程序、权限和责任，确保公共数据使用的合法性和规范性，鼓励和支持多元化、专业化的数据运营主体参与公共数据授权运营，确保数据开放政策与城市创新战略和创新需求相衔接。第二，应发挥公共数据在优化政府治理、改善要素配置以及促进数字经济发展中的积极作用。政府可建立公共数据驱动的决策支持系统和反馈机制，利用数字技术对公共数据进行深度挖掘和分析，为决策制定提供科学依据。第三，应推动公共数据在创新要素市场中的开放共享与应用，鼓励企业、科研机构和社会组织等利用公共数据开展产品研发、技术创新和 market 分析，进而优化创新资源配置。第四，还应打造公共数据驱动的数字经济创新生态，加大对数字经济创新项目的支持力度，鼓励创新主体利用公共数据进行技术创新和模式创新。第五，政府应提高对数字化转型和数字技术的认识与关注，在推动全国统一大市场建设中构建全国统一的区域公共数据共享平台，

① 余典范、龙睿、王超：《数字经济与边界地区污染治理》，《经济研究》2023年第11期，第172—189页。

促进公共数据要素在区域间、行业间自由流动，为各类创新主体提供便捷的数据获取和分析服务。

Public Data Openness and Urban Innovation: Impact Effects and Mechanisms

ZHENG Guoqiang, LYU Hang

Abstract: Data elements have become a crucial force in driving regional innovation and productivity transformation. Taking the public data open platforms established by local governments as quasi-natural experiments and based on panel data from 287 prefecture-level cities in China from 2010 to 2021, this study empirically examines the impact and mechanism of public data openness on urban innovation using a progressive difference-in-differences model. The research findings indicate that public data openness can effectively promote urban innovation, and this conclusion remains valid after a series of robustness tests. The mechanism test reveals that public data openness drives urban innovation primarily through reducing policy uncertainty, improving distortions in factor allocation, and optimizing digital development environment. Heterogeneity tests show that the innovation-driving effect of public data openness is more significant in cities with higher government digital attention, marketization, and regional integration. Further analysis reveals that in addition to increasing the quantity of innovation, the innovation-driving effect of public data openness is also reflected in the improvement of urban innovation efficiency and the spillover effect on the innovation development of neighboring cities. The research conclusions not only expand the research perspective on urban innovation drivers but also provide practical reference for the formulation of government policies on public data openness.

Keywords: public data openness; urban innovation; policy uncertainty; factor allocation; digital economy

(责任编辑: 李 玲)