

公共数据开放如何解决“创新者的窘境”： 基于数字生态系统视角

吉林大学经济学院

论文进展汇报

2024 年 12 月 15 日



- 1 引言
- 2 文献综述
- 3 政策背景与研究假设
- 4 研究设计
- 5 实证结果分析
- 6 稳健性检验
- 7 进一步讨论
- 8 研究结论与政策建议

1 引言

2 文献综述

3 政策背景与研究假设

4 研究设计

5 实证结果分析

6 稳健性检验

7 进一步讨论

8 研究结论与政策建议

研究背景

公共数据开放作为推进数字政府建设与数字经济战略的重要举措，在优化营商环境、完善公共数据要素供给上发挥了重要作用（Hughes and Coronado, 2019；蔡运坤等，2024）。数据显示，有 96% 的行业重点企业已实现数据场景化应用，然而，实现数据多主体复用增值的大企业仅占 8.3%，数据价值有待释放。2024 年 9 月 21 日，中共中央办公厅、国务院办公厅印发的《关于加快公共数据资源开发利用的意见》，指出要“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”。在此背景下，评估和解释公共数据开放如何破除企业创新的市场结构障碍，激励和引导企业进行创新成果转化，具有重要的现实意义与政策价值。

“颠覆式创新”的思考

颠覆式创新理论呈现出的一系列矛盾引发了我们的思考：一方面，在商业生态系统中，“创新者”面临着何种“窘境”？另一方面，数字技术的应用改造了传统商业生态，成功的“创新者”如何借助平台生态与创新生态实现了“颠覆”？

本文基于数字平台生态系统视角,以地方政府开通公共数据开放平台作为公共数据开放的准自然实验,选取 2008—2020 沪深 A 股上市公司作为研究样本构建双重差分模型,实证考察公共数据开放如何解决“创新者的窘境”。研究发现,公共数据开放有助于企业颠覆式创新。机制检验表明,公共数据开放主要通过实现主流产品创新突破、促进企业数字化转型升级、整合新兴市场价值网络等途径促进数字平台生态颠覆式创新。异质性检验发现,公共数据开放对企业颠覆式创新的促进作用受到区域数字基础设施建设与数字规制政策制定的正向影响,而且在非数字产业企业、国有企业、二职分离企业中更为明显。研究结论为如何破除企业创新的市场结构障碍,弥合企业数据要素获取与利用的“数字鸿沟”,激励和引导企业进行创新成果转化提供了政策启示。

边际贡献

相比于现有文献，本文可能的边际贡献主要体现在以下几个方面：第一，相比于已有的实证研究，本文从数字生态系统的角度验证了公共数据开放对颠覆式创新的外生催化作用，丰富了公共数据开放领域的研究成果。第二，本文刻画并识别数字生态系统中公共数据开放的内生创新机制，填补了国内基于数字生态系统的实证研究空白。第三，本文从数字生态系统的商业生态系统、创新生态系统、平台生态系统等多个角度解析了公共数据开放对于数字生态系统的赋能与重塑作用，加深了对于公共数据开放政策的理解。第四，本文从数字平台生态的硬环境与软环境产生的调节效应出发，多方面考察了公共数据开放政策对在不同企业性质、行业特征、股权结构等方面的异质性效应，为完善我国公共数据开放政策设计，营造更加良好的数字创新生态，实现数字生态价值共创提供了政策建议。

“创新者的困境”

Christensen and Raynor (2003) 在一系列技术创新研究基础上, 区别于渐进式创新, 提出了颠覆式创新理论, 认为“某种基于颠覆性技术的新产品服务于重视其非标准性能属性的细分市场, 进一步发展可以将颠覆性技术在主流属性上的表现提高到足以满足主流客户的水平, 最终在主流市场上取代了主流产品, 市场颠覆就发生了”。颠覆式创新通常是由新进入者开发和商业化的, 这是由于与现有在位企业相比, 新进入企业在颠覆式创新中有更大的成功机会, 因为它们对价值网络和当前技术模式的依赖更有限 (Christensen and Bower, 1996)。Christensen and Raynor (2003) 将其称为“创新者的窘境”, 以概括市场中的在位企业在颠覆式创新中的劣势。从市场角度来看, 颠覆式创新是一种过程, 通过破坏价值网络为新进入企业创造价值空间, 能够重构市场规则与竞争态势 (Christensen, 2006)。

“颠覆者的困境”

然而，由于平台市场中网络效应的存在（Katz and Shapiro, 1985），互补产品与创新成功商业化中起到关键作用，创新者需要依赖其互补产品从他们的创新中获得价值（Teece, 1986），学者们将由此形成的相互依赖的企业网络称为“商业生态系统”（Moore, 1996）。在商业生态系统中，每个企业的价值网络和技术模式，包括其各自的供应商、互补者、竞争对手和客户，重叠并交织在一起，产生多种互补的或竞争的主张，将创新引入现有生态系统的颠覆者需要将一系列交易、规则和角色整合在一起，以管理这些相互依赖关系（Jacobides et al., 2006; Adner, 2012）。Ansari 等人（2016）将这些问题总结为“颠覆者的窘境”，即在颠覆现有技术的同时，颠覆性技术也可能颠覆互补产品的供应商，而这些补充产品甚至在位企业的产品是新技术在市场上取得成功所必不可少的。如何处理在位企业和新进入企业之间的紧张关系，解决市场中的“创新者的窘境”，关键在于构建一个新的围绕颠覆式创新产品及其互补品的商业生态。

如何解决“创新者的困境”

与传统的“商业生态系统”不同，数字经济下的商业生态系统“由数字技术产生的相互依赖网络形成”（Kopalle et al., 2020），其特征在于平台市场中“由于客户需求、新兴技术和竞争行为的快速变化而导致环境中不可预测性的条件”，产生了围绕消费偏好的企业相互依赖和竞争以及相互关联的技术传播与采用（Hanelt et al., 2021）。数字生态系统为创新者带来了新的挑战 and 机遇（Teece, 2018）。一方面，平台竞争是由多个消费者群体所驱动的，而这些消费者群体本身的价值主张又受到间接网络效应的影响（Rochet and Tirole, 2003, 2006）。数字平台通过减小交易成本而增加了平台创新网络的连接性（刘洋等, 2020），使得在位企业可以通过对其主流产品和互补品的价值网络进行不断调整（Adner, 2017），实现在位企业对于创新产品的商业化，形成“创新生态系统”（Jacobides et al., 2018）。另一方面，平台作为一种分层和可分解的系统有助于减轻复杂性的影响，可以作为其他公司开发互补产品、技术或服务的基础（Gawer, 2009）。

如何解决“创新者的困境”

具体来说,随着企业在整个组织及其价值链中利用更标准化的工具来设计、生产和支持产品和服务,数字技术的应用可以跨组织边界共享更多的数据和流程(Yoo et al., 2012)。可通过共享或开源技术以及技术标准(如应用程序编程接口),将一系列外围公司连接到中心平台,形成技术意义上的“平台生态系统”(Jacobides et al., 2018)。数字技术的融合加剧了对知识资源异质性和动态整合的需求,利用数字技术进行几乎没有边界的重组已经成为创新的新来源(Nambisan et al., 2017),颠覆者能够在颠覆现有技术的同时,快速实现与其互补产品的兼容性,并破坏在位企业在上一代平台的基础价值,削弱其与“商业生态系统”的联系(Ozalp et al., 2018),实现颠覆式创新。

简言之，数字平台改变了市场中行业的竞争规则与产品的创新流程，实现了“数字平台生态系统”这一创新性的市场设计 (Kretschmer et al., 2020; 焦豪, 2023)，这为我们解决“创新者的窘境”的提供了可行方案：在数字化情景下，围绕平台中的多元参与者搭建开放互补的“平台生态系统”，实现与“商业生态系统”中用户和互补者的流量匹配以及“创新生态系统”中创新产品与互补产品的有效整合，既能激励在位企业通过与互补者和间接消费者的有效交互，开发基于颠覆性技术的新产品以实现网络效应，构建持续的竞争优势，破除“创新者的窘境”；又能帮助新进入企业迅速调整以应对与其在位者和互补者的紧张关系，使主流产品与必需互补品迅速适配颠覆式创新产品，将其关系定位从价值破坏者转变为价值创造者 (Ozalp et al., 2018)，将“逆风”转变为“顺风”，突破“颠覆者的窘境”。

文献评述

综上所述，公共数据平台作为一项数字生态变革改变了平台市场中企业的数据收集和交易策略，改变了企业数据要素利用的技术—经济路径（Farboodi and Veldkamp, 2020；杨俊等，2022），但现有研究多关注于公共数据开放对于企业生产经营决策的影响，鲜有文献考察公共数据开放对于企业创新活动的重塑，难以全面揭示其激励企业颠覆式创新的内在逻辑。本文将从数字生态系统视角识别公共数据开放对企业创新的影响，基于创新生态系统迭代突破、商业生态系统数字转型、平台生态系统市场整合等途径探讨其颠覆式创新背后的作用机制。

1 引言

2 文献综述

3 政策背景与研究假设

4 研究设计

5 实证结果分析

6 稳健性检验

7 进一步讨论

8 研究结论与政策建议

可见，公共数据开放是中国政府对公共数据资源开发利用的系统部署，旨在将公共数据视为国家基础性战略资源，并深化数据要素市场化配置改革。这一政策强调公共数据资源的重要性，旨在破除体制性和机制性障碍，公共数据开放平台的建设，对于促进政府、企业和公众之间的信息交流和互动，完善数据要素供给和激发企业创新活力具有重要意义。

理论分析与研究假设

整个数字平台生态系统不是静态的，而是平台架构持续延展和产业链群更新适应的动态演化过程。从平台内部自身架构来看，平台架构设计的延展特点强调平台中的各层级数字设计功能和技术工具具备弹性，可随着数字平台生态系统的发展实现功能扩展。数字平台生态系统可以通过开放平台接口促进不同平台产品和服务之间的共享以及为新的合作伙伴部署增值服务，可以促进数字平台生态系统治理机制的动态扩展（Jovanovic et al., 2022）。而环境选择、市场压力和资源要素供给状况的变化是推动产业链群生态体系演化的外部动力。产业链群生态体系是自适应循环的产业生态系统，能够随着内部突变和环境变化而发生演化，并通过不同的演化机制和演化路径实现生态系统内数据流的整合，最终实现动态平衡（余东华和李云汉，2021）。因此，数字平台生态系统可表现出自生演化的特征（如图1所示）。

- ① 开放端口接入
- ② 平台生态系统整合
- ③ 创新生态系统迭代
- ④ 商业生态系统转型

图 1: 数字生态系统视角下的公共数据开放与颠覆式创新示意图

理论分析与研究假设

可见，数字平台生态系统被来自宏观环境层面的外部力量所塑造，数字平台生态系统的内生属性，如技术架构和利益相关者治理需要与外部环境的动态变化相适应。同时，平台所有者的内生选择会影响利益相关者期望、促进协调并实现生态系统外部的兼容性。例如，数字平台会基于不同发展阶段打造自身的主流产品，在焦点企业初始阶段围绕自身独创的“明星产品”构建自身数字平台，随着其消费规模的扩大，企业将编排其平台资源并调整其平台参与者，进一步开发全新的场景化产品以满足来自不同场景的消费需求（焦豪，2023）。

理论分析与研究假设

基于自身创新生态的主流产品迭代升级需要协调企业自身原有的价值网络与可能的互补供应商，而公共数据开放降低了企业的信息搜寻成本并产生正向信息溢出效应。公共数据开放使得需求水平、产品质量以及成员信用等生态内原有的无法被观测的私人信息变为共有信息，有效改善了企业与互补品间信息共享水平，缓解与焦点企业的信息不对称，极大便利了企业双边交易。契约信息可验证性的提高使得企业对于供应链创新产品的验证成本降低（Goldfarb and Tucker, 2019）。由于企业可以分析合作者的历史信用，企业借助简单的筛选条件，更大程度地避免了与可能发生的投机行为（Teece, 2018）。同时，企业自身也可以通过公共数据释放声誉信号，营造相互信任的生态氛围，积累自身的生态网络资源，为实现颠覆式创新迭代积累资源基础（李玉花和简泽，2021）。

理论分析与研究假设

然而，数据生产要素与数字技术不可分割，公共数据要素涉及只有当具备一定规模时，才能产生持续积累的绝对生产力，放大数字技术的外溢性。这意味着后续投入越多的数据生产要素参与到再生产环节，前期投入的专用性资产越能得以充分摊销，单位使用成本才能随之降低。例如，对于数据模型精准度依赖高的先进数据应用领域，企业为了提高数据驱动的决策的精准度，会采集更多数据进行交叉验证，同时对数据进行更加细致的整理，从而提高数据处理成本。而且公共数据的时效性较强，更多的历史数据并不能为使数据要素具有更好的表现（王超贤等，2022）。

理论分析与研究假设

因此，数据资源越庞大，适配数据资源处理需求的数据存储、算力、算法等数字基础设施建设需求也将更高阶。这会占用企业大量研发经费投入，加之这些基础设施具有较强资产专用性，公共数据处理在企业出现颠覆性创新之前会占用大量的企业资源（Pedota, 2023），这又削弱了生态位中企业进行颠覆性创新的激励。此外，随着数据资本投入的持续增加以及累积，企业间的数据资本投入水平差距会拉大，而数据资本的规模经济、范围经济和网络效应会导致强者愈强的市场集中（唐要家等，2022；Bergemann and Bonatti, 2024），市场竞争程度的下降会降低市场中企业的创新。

机制分析

公共数据开放平台的设立有利于数字平台生态基础设施水平的提升和相关企业与生态位成员之间的网络资源的整合。随着公共数据的开放，地区数字基础设施的完善和数据安全保障技术的提升（Zuiderwijk and Janssen, 2014），能够有效激活消费者对数字化、信息化产品的市场需求（Hughes and Coronado, 2019），提高数字平台生态的创新绩效（Dawes et al., 2016）。同时，外部数字发展条件的完善有助于构建企业内外上下协同的数字生态圈和数据共享平台，发挥企业间的协同效应，降低了外部不确定性风险。首先，公共数据开放使得企业可以获得较多外部知识、信息资源和社会资本，增加企业创新的战略性资源拥有量，并使企业通过整合和有效利用外部数据资源准确预测市场发展趋势（Farboodi and Veldkamp, 2023），减少企业发展面临的不确定性。其次，公共数据开放有助于企业降低产品研发创新中的信息不对称（Veldkamp and Chung, 2024），提高企业决策质量和企业内生发展动力，将企业的颠覆式创新成果迅速转化为经济绩效。

机制分析

基于上述讨论，本文提出以下假说：

- 假设 2a：企业的数字化转型程度在公共数据开放对企业颠覆式创新的影响中起到了中介作用，这种影响呈现出对企业颠覆式创新的促进效应。
- 假设 2b：市场集中程度在公共数据开放对企业颠覆式创新的影响中起到了中介作用，这种影响呈现出对企业颠覆式创新的促进效应。

机制分析

数字创新生态中的创新过程依赖于数字环境的支持。数字环境是支持一个企业或者产业基本的数字技术与组织结构以及相关服务与设施，不仅包括计算机、移动设备和应用平台等数字基础设施，还包括数字社区相关组织和治理等数字规制政策（Tilson et al., 2010）。在数字平台生态系统中，数字基础设施不仅影响到企业创新的技术路径，改变企业间的竞合关系，还影响到了企业创新启动、开发、应用过程中各个主体之间的动态交互过程（刘洋等，2020）。

机制分析

基于上述讨论，本文提出以下假说：

- 假设 3a：公共数据开放与企业颠覆式创新的“U”型关系随着硬件基础设施的提高而变缓，表现为线性促进作用。
- 假设 3b：公共数据开放与企业颠覆式创新的“U”型关系随着有效治理强度的提高而变缓，表现为线性促进作用。

机制分析

企业在管理层激励、企业性质、数字化应用能力等方面的禀赋不同，导致不同企业对于数据要素的获取与利用能力存在着巨大的“数字鸿沟”，不利于实现数字平台生态的价值共创，阻碍了数字平台生态演化过程中实现颠覆式创新。据此，本文将从企业的股权结构、产权属性和数字化基础三个方面探索公共数据开放能否起到弥合数字鸿沟的作用。

机制分析

基于上述讨论，本文提出以下假说：

- 假设 4a：公共数据开放对数字平台生态颠覆式创新的促进作用呈现异质性，其对于二职分离企业相较于二职合一企业促进效应更强。
- 假设 4b：公共数据开放对数字平台生态颠覆式创新的促进作用呈现异质性，其对于国有企业相较于民营企业促进效应更强。
- 假设 4c：公共数据开放对数字平台生态颠覆式创新的促进作用呈现异质性，其对于非 ICT 行业企业相较于 ICT 行业企业促进效应更强。

- 1 引言
- 2 文献综述
- 3 政策背景与研究假设
- 4 研究设计
- 5 实证结果分析
- 6 稳健性检验
- 7 进一步讨论
- 8 研究结论与政策建议

模型设定

关于企业迭代式创新的突破机制的假说，主要考察企业过去积累的迭代式创新是否对当前的公共数据开放对企业颠覆式创新的冲击有影响。本文使用了企业的迭代式创新，将其作为解释变量加入基准回归中。如果发现，一个企业过去的迭代式创新与公共数据开放的交乘项对当前的颠覆式创新的影响 β_1 显著为正，而公共数据开放对于当前的颠覆式创新的直接影响 β_2 显著为负，则初步说明公共数据的开放对企业实现颠覆性创新具有显著的门槛效应。此外，为了验证这种倒 U 型关系是否真的存在，本文借鉴已有研究成果（Hansen, 1999）对倒 U 型关系进行了门槛回归：

$$\begin{aligned}
 \text{Innovation}_{D,it} = & \alpha + \eta_1 \text{Open}_{it} \times I(\text{L.innovation}_{C,it} \leq T) \\
 & + \eta_2 \text{Open}_{it} \times I(\text{L.innovation}_{C,it} > T) \\
 & + \eta_3 \text{L.innovation}_{C,it} + X_{it} + Z_{jt} + u_i + v_t + \varepsilon_{ijt}
 \end{aligned}$$

模型设定

具体做法是：首先，确定 U 型关系的最低点 T ，然后使用在该断点处对之前和之后的 $L.innovation_{C,it}$ 分别生成示性函数 $I(\cdot)$ 进行断点回归，如果在该点之前的非线性中介的系数显著为负，而且在该点之后的非线性中介系数显著为正，则可以证明 U 型关系。其次，考察 $L.innovation_{C,it}$ 的分布和 U 型中点位置之间的联系，考察非线性中介的取值范围是否涵盖了倒 U 型最高点位置且该位置是否较靠近分布的中心。最后，为进一步验证倒 U 型关系，本文在基准回归中加入示性变量进行了二次回归。

变量设定与数据说明

本文从公司基本治理属性和公司结构特征选取企业特征变量作为企业层面控制变量，选取企业规模（Size）、企业年龄（Age）、资产负债率（Lev）、总主营业务增长率（Grow）、资产收益率（ROA）、现金流占比（Cashflow）、资本密集度（Capital）、董事会规模（Board）、股权集中度（Tops）、机构投资者占比（Investor）、独立董事占比（Idr）等企业层面的控制变量。本文还考虑到地区经济因素对于企业数字技术创新的影响，选取区域人口规模（Pop）、经济发展水平（GDP）、教育资源（Edu）、人力资本（Human）、产业结构（Stru）、财政收入（Tax）、政府支出（Finance）等区域层面的控制变量。

变量设定与数据说明

关于企业数字化转型程度 (Trans)，本文借鉴已有研究成果 (吴非, 2021) 通过 Python 爬虫功能归集整理了样本内沪深 A 股上市企业的年度报告，并以此作为数据池供后续的特征词筛选。由于这类数据具有典型的“右偏性”特征，本文将其进行对数化处理，从而得到刻画企业数字化转型的整体指标。在稳健性检验中，本文根据技术的构成差异和运用状况，进行口径细分并重新进行回归检验。关于市场集中程度，本文选用赫芬达尔指数 (HHI) 来衡量企业面对的市场结构。为更好地反映行业间差异并且保证结果的稳健性，本文还选用行业集中度 (CR_n) 来衡量企业面对的市场结构，分析时采用 CR_4 和 CR_{10} 指标。

变量设定与数据说明

关于数字生态环境，本文借鉴已有研究成果（孙伟增等，2023）选取了人均电信业务总量（Tele）、互联网宽带接入用户数（Internet）和北京大学编制的数字普惠金融指数（Digital Fin ），表征企业数字发展所需要的基础设施水平（即硬环境）。此外，本文借助“北大法宝”整理汇总了各省份每年新颁布的与“数字法治”相关主题下的地方性法规、地方政府规章和地方规范性文件数（Digital Law ），反映企业数字发展所需要的有效规制强度（即软环境）。

① 引言

② 文献综述

③ 政策背景与研究假设

④ 研究设计

⑤ 实证结果分析

⑥ 稳健性检验

⑦ 进一步讨论

⑧ 研究结论与政策建议

基准回归结果

基准回归

首先，考察公共数据开放对企业颠覆式创新可能存在的非线性效应。图 2 列示了开放公共数据对企业创新影响的实证结果。第(4)列是在第(3)列控制公司和年份层面固定效应的基础上，增加了省份层面的固定效应。主要解释变量系数均在 1% 的基础上显著。

变量	(1)	(2)	(3)	(4)
<i>Open</i> × <i>L_innovation</i>	0.0729*** (9.7421)	0.03918*** (3.7704)	0.0426*** (4.0099)	0.0429*** (4.0434)
<i>Open</i>	-0.3039*** (-9.6647)	-0.1697*** (-4.7891)	-0.1480*** (-4.0655)	-0.1334*** (-3.6874)
<i>L_innovation_c</i>	0.3767*** (82.6484)	0.3133*** (23.3301)	-0.1480*** (-4.0655)	0.3158*** (23.8399)
控制变量	是	是	是	是
个体固定效应	否	是	是	是
年份固定效应	否	否	是	是
省份固定效应	否	否	否	是
<i>Constant</i>	-3.6889*** (-11.6426)	-5.0552*** (-3.6941)	-5.3247*** (-3.6503)	-2.6242 (-0.8956)
<i>N</i>	15561	15561	15561	15561
<i>R</i> ²	0.5263	0.7875	0.7896	0.7906

图 2: 基准回归结果

门槛效应检验

门槛效应

其次，通过对门槛模型来验证公共数据开放与企业颠覆式创新的 U 型关系。在门槛回归前，采用 Bootstrap 法进行 300 次反复抽样得出检验统计量验证是否存在着门槛效应，检验结果如图 3 所示。这里拒绝了三重门槛效应，单一门槛效应与双重门槛效应在 1% 的基础上显著。双重门槛与单一门槛第一门槛估计值均为 5.1818，双重门槛第二估计值为 6.6746。而且，图 4 显示，门槛值均位于渐进式创新取值范围内。

	Th1	Th2	Th3	F	P
Single	5.1818			186.64	0.0000
Double	5.1818	6.6746		73.46	0.0000
Triple	5.1818	6.6746	0.6931	61.17	0.4000

图 3: 门槛效应检验

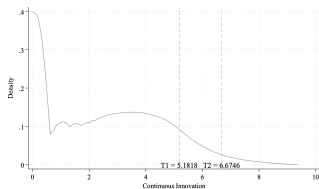


图 4: 渐进式创新核密度图

门槛回归结果

门槛回归

基于门槛值的系数估计结果如图 5 所示。这里的双重门槛效应显示，当企业积累的迭代式创新未突破第一门槛值 5.1818 时，公共数据开放对于企业颠覆式创新的影响在 1% 水平上显著为负；当积累的迭代式创新突破第一门槛值时，公共数据开放对于企业颠覆式创新的影响在 1% 水平上显著为正；突破第二门槛值时，公共数据开放对于企业颠覆式创新的影响在 1% 水平上显著为正，且系数估计值递增，基本验证了本文的假说 1

	(1) 单一门槛	(2) 双重门槛
Open ($L.innovation \leq T_1$)	-0.0865*** (-4.35)	-0.0822*** (-4.14)
Open ($T_1 < L.innovation \leq T_2$)	0.3068*** (9.65)	0.2098*** (6.21)
Open ($L.innovation > T_2$)	0.3068*** (9.65)	0.6969*** (12.29)
个体固定效应	是	是
年份固定效应	是	是
控制变量	是	是
Constant	-5.3090***	-5.1189***
N	15561	15561
R ²	0.5191	0.5255

图 5: 门槛回归结果

中介效应检验

中介效应检验

上述结果说明，公共数据开放能否提高企业颠覆式创新绩效的根本在于促进企业迭代式创新的产出积累，这种迭代式创新的积累会带来持续的颠覆式创新绩效改进。进一步，为验证公共数据开放与企业颠覆式创新的 U 型关系之间可能存在的作用机制，本文借鉴已有研究成果（温忠麟，2014），在模型基础上构建中介效应模型进行检验。中介机制分析结果如图 6 所示，Sobel 检验也证实存在这一中介机制。

变量	(1)	(2)	(3)	(4)
	<i>Trans</i>	<i>Innovation_p</i>	<i>HHI</i>	<i>Innovation_p</i>
<i>Open</i> × <i>L. innovation</i>	0.0378*** (5.45)	0.0411*** (6.42)	0.0028*** (3.88)	0.0425*** (6.62)
<i>Trans</i>		0.0429*** (5.56)		
<i>HHI</i>				0.2687*** (3.64)
控制变量	是	是	是	是
个体固定效应	是	是	是	是
年份固定效应	是	是	是	是
<i>Sobel</i>	0.0016*** (3.894)		0.0008*** (2.656)	
<i>Constant</i>	-5.4367*** (-6.03)	-5.0915*** (-6.11)	-5.3248*** (-6.39)	-5.4368*** (-6.03)
<i>N</i>	15561	15561	15561	15561
<i>R</i> ²	0.6990	0.7900	0.7896	0.6990

图 6: 中介效应检验结果

平行趋势检验

采用双重差分法的前提是实验组与对照组企业数字发明专利申请量在事件冲击前不存在明显差异。为了检验平行趋势是否成立，本文构建具体模型如下：

$$\text{Open}_{it} = \alpha + u_1 D_{it}^{-13} + \cdots + u_{22} D_{it}^8 + \gamma X_{it} + \delta Z_{it} + \varphi_i + \lambda_j + \theta_p + \mu_t + \varepsilon_{ijpt}$$

这里，为防止多重共线性干扰，本文将政策实施的前一期作为参照系赋值为零。

平行趋势检验

平行趋势检验

平行趋势检验如图 7 所示。结果显示，在设立公共数据开放平台前，实验组和对照组的企业颠覆式创新数量的变化趋势并未有显著差异。

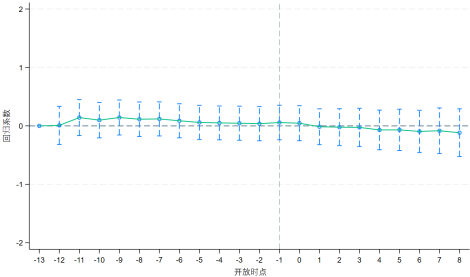


图 7: 平行趋势检验结果

- 1 引言
- 2 文献综述
- 3 政策背景与研究假设
- 4 研究设计
- 5 实证结果分析
- 6 稳健性检验
- 7 进一步讨论
- 8 研究结论与政策建议

安慰剂检验

安慰剂检验

鉴于被解释变量的统计显著性可能源于某些难以观察的随机因素，本文随机打乱实验组与对照组分组情况进行安慰剂测试。首先，利用计算机随机生成新的处理组，将上述过程重复进行 500 次回归；然后，画出所有回归结果中 t 值的核密度图。结果如图 8 所示。可见，模拟结果远远小于中与实际回归中的 t 值，结果仍旧稳健。

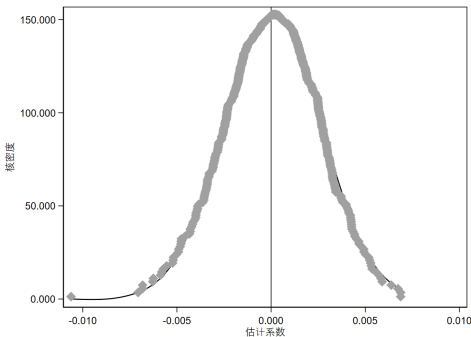


图 8: 安慰剂检验结果

调整样本结构

调整样本结构

为排除直辖市政策干扰，本文剔除了来自直辖市的样本，回归结果如图 9 所示。如第 (1) 列所示，仍旧与假设一致。2016 年，贵州省获批建立我国首个国家级大数据综合试验区，对促进贵州省大数据基建整合和数据资源汇聚应用具有促进作用，因此，本文剔除了贵州省的样本重新回归，结果如第 (2) 列所示，模型仍旧稳健。

变量	(1)	(2)
$Open \times L.innovation$	0.04898*** (3.7961)	0.04025*** (3.7211)
$Open$	-0.1621*** (-3.4736)	-0.1438*** (-3.9455)
$L.innovation_c$	0.3147*** (21.6343)	0.3171*** (23.6330)
控制变量	是	是
个体固定效应	是	是
年份固定效应	是	是
$Constant$	-2.9988* (-1.9230)	-5.2564*** (-3.5719)
N	12290	15379
R^2	0.7665	0.7903

图 9: 调整样本结构

匹配样本检验

匹配样本检验

公共数据资源开放相对企业数字技术创新满足外生条件，但不同企业的样本在组间可引起的选择性偏差。为了消除选择性偏差的干扰差的干扰，本文将利用倾向得分匹配控制样本的组间特征。具体而言，本文以基准回归中地区经济变量和企业控制变量为匹配变量进行和企业控制变量为匹配变量进行 1:4 的有放回近邻匹配的有放回近邻匹配。基于倾向得分匹配的样本检验回归结果如图 10 所示。公共数据开放的回归系数在 1% 水平上显著为正，这说明选择性偏差不会影响本文

变量	(1)
<i>Open</i> × <i>L. innovation</i>	0.05826*** (4.0508)
<i>Open</i>	-0.1973*** (-3.4029)
<i>L. innovation_c</i>	0.3101*** (21.1060)
控制变量	是
个体固定效应	是
年份固定效应	是
<i>Constant</i>	-0.6513 (-0.3714)
<i>N</i>	9833
<i>R</i> ²	0.7823

图 10: 匹配样本检验

外生政策冲击

调整样本结构

已有研究认为国家级大数据综合试验区会对企业创新行为产生影响，为排除国家级大数据综合实验区的政策干扰，本文将国家级大数据综合试验区作为控制变量 Policy 加入模型中，回归结果如图 11 所示。可见 Policy 的回归系数并不显著，而 Open 与 L.innovation 交乘项和 Open 的回归系数仍然在 1% 水平下显著为正。这充分证实了公共数据开放对企业颠覆式创新的激励作用不受到其他类似政策干扰。

变量	(1)
$Open \times L.innovation$	0.04307*** (4.0248)
Open	-0.1429*** (-3.8847)
L.innovation	0.3157*** (23.7454)
Policy	-0.04410 (-1.0596)
控制变量	是
个体固定效应	是
年份固定效应	否
Constant	-5.3104*** (-3.6971)
N	15470
R ²	0.7896

图 11: 外生政策冲击

1 引言

2 文献综述

3 政策背景与研究假设

4 研究设计

5 实证结果分析

6 稳健性检验

7 进一步讨论

8 研究结论与政策建议

异质性分析

激励异质性

第一，本文以两职合一作为股权激励水平的代理变量，将样本进行分组回归，结果如图 12 所示。可见，政策效应仅在两职分离的企业中显著，即公共数据开放对企业颠覆性创新存在股权激励水平的异质性作用。

变量	(1) 两职分离	(2) 两职合一
$Open_{it} \times L.innovation_{c,it}$	0.05105*** (4.5325)	0.01735 (0.5778)
$Open_{it}$	-0.1555*** (-3.9181)	-0.09426 (-0.9553)
$L.innovation_c$	0.3151*** (22.8183)	0.3059*** (6.3883)
控制变量	是	是
个体固定效应	是	是
年份固定效应	是	是
Constant	-4.5459*** (-2.8009)	-10.776*** (-3.1361)
N	12744	2190
R ²	0.8002	0.8084

图 12: 激励异质性

异质性分析

股权异质性

第二，本文以企业股权性质作为政企纽带的代理变量，将样本进行分组回归，结果如图 13 所示。可见，政策效应仅在国有控股的企业中显著，即公共数据开放对企业颠覆性创新存在股权性质的异质性作用。

变量	(1) 民营企业	(2) 国有企业
$Open_{it} \times L.innovation_{c,it}$	0.01886 (1.0182)	0.05790*** (4.5212)
$Open_{it}$	-0.04385 (-0.7387)	-0.1778*** (-3.8790)
$L.innovation_c$	0.2971*** (12.2964)	0.3169*** (19.6114)
控制变量	是	是
个体固定效应	是	是
年份固定效应	是	是
Constant	-2.4961 (-1.5260)	-6.0168** (-2.3390)
N	5464	10061
R ²	0.7694	0.8073

图 13: 股权异质性

异质性分析

行业异质性

第三，本文依照上市公司行业分类指引（2012 年版）将企业所属行业作为 ICT 行业的代理变量，将样本进行分组回归，结果如图 14 所示。可见，政策效应在两组企业中呈现方向的不同，在非 ICT 企业中呈现 U 型关系，在 ICT 企业中呈现倒 U 型关系。这在一定程度上表明，ICT 企业对于公共数据的开发呈现要素规模报酬递减的趋势，处于数据要素的“成型期”，而非 ICT 企业则呈现要素规模报酬递增的趋势，处于数据要素的“起步期”，公共数据开放对企业颠覆式创新存在行业属性

变量	(1) ICT 企业	(2) 非 ICT 企业
$Open_{it} \times L.innovation_{C, it}$	0.04716** (4.3966)	-0.1722** (-2.6655)
$Open_{it}$	-0.1558** (-4.2376)	0.5406** (2.2924)
$L.innovation_{C, it}$	0.3136** (23.5617)	0.3728** (4.9311)
控制变量	是	是
个体固定效应	是	是
年份固定效应	是	是
Constant	-4.8459** (-3.2774)	-4.9877 (-0.6008)
N	15245	314
R ²	0.7900	0.8905

图 14: 行业异质性

- 1 引言
- 2 文献综述
- 3 政策背景与研究假设
- 4 研究设计
- 5 实证结果分析
- 6 稳健性检验
- 7 进一步讨论
- 8 研究结论与政策建议

政策建议

为了推动经济的高质量发展，政府需要进一步优化公共数据开放政策，这包括破除体制性和机制性障碍，激发数据共享的开放动力，并优化公共数据资源的配置。同时，加强数字基础设施建设是提升企业利用公共数据资源能力的关键，这将为创新活动提供坚实的技术支撑。此外，制定和实施数字规制政策对于降低企业面临的网络效应壁垒至关重要，这有助于促进颠覆式创新成果的落地。政策制定者还需关注企业的异质性，根据不同企业的性质、行业特征和股权结构，制定差异化的政策，以实现更精准的政策支持。政府应通过政策激励和引导，帮助企业尤其是非数字产业企业、国有企业和二职分离企业进行创新成果转化。最终，通过公共数据开放政策，可以弥合企业间在数据要素获取与利用方面的“数字鸿沟”，这对于推动经济的数字化转型和持续增长具有重要意义。