

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

蔡运坤 周京奎 袁旺平*

摘要:公共数据开放是推进数字政府建设与数字经济战略的重要举措,其能否释放价值创造效应、促进社会创业与增强经济活力,是实践界与理论界共同关注的重大问题。本文利用地级市及以上城市的面板数据与企业工商注册信息数据,基于公共数据开放的准自然实验构建多期双重差分模型,系统检验了公共数据开放对城市创业活力的影响。研究发现,公共数据开放能显著促进城市创业活力提升。机制分析表明,公共数据开放通过制度性交易成本降低的政府治理逻辑和信息搜寻匹配效率优化的市场效益逻辑的途径来推进城市创业活力提升。异质性分析表明,区域要素禀赋与非正式制度环境强化了公共数据开放的创业促进效应,而且公共数据开放弥补了初始地区数字基础设施政策支持差距。此外,公共数据开放对生产性服务业的创业促进作用更明显,表明公共数据开放还促进了高质量创业。研究结论对于进一步推进公共数据开放、释放其价值创造效应、推动经济高质量发展具有重要的政策启示。

关键词:公共数据开放 城市创业活力 制度性交易成本 信息搜寻匹配效率

中图分类号:F293 **文献标识码:**A **文章编号:**1000-3894(2024)08-0005-21

一、引言

数据要素推动传统生产要素发生革命性聚变与裂变,通过数据挖掘、脱敏处理、整合分析,实现数据资源与供给、需求的精准对接,创新价值链流转方式,有助于培育新质生产力,推动经济高质量发展。相关数据显示,2017~2022年,中国数据产量从2.3ZB增长至8.1ZB,占全球数据总量的比例从8.8%增长至10.5%,位居全球第二^①。公共数据是其中比重最大,特别是涵盖政府部门在履行公共职责时产生的政府治理相关数据以及具有重大公共利益的企业数据等。中国政府把充分激发公共数据开放的价值创造效应置于重要战略位置。现阶段,激发国内经济活力是推进经济高质量发展的内在要求。企业是市场经济运行的主体,是推动国民经济发展的“压舱石”,充分激发并释放全

* 蔡运坤(通讯作者),博士研究生,南开大学经济学院,电子邮箱:caiyunkun_1997@163.com;周京奎,教授,南开大学经济学院、经济行为与政策模拟实验室,电子邮箱:zjk2004@nankai.edu.cn;袁旺平,博士研究生,南开大学经济学院,电子邮箱:1120231303@mail.nankai.edu.cn。本文获得国家社会科学基金重大项目(22&ZD129)、国家社会科学基金重点项目(22AJY004)、天津市哲学社会科学规划项目(TJYJ20-003)和南开大学文科发展基金项目(ZB22BZ0206)的资助。感谢匿名审稿专家的宝贵意见,文责自负。

① 数据来源:中国信息通信研究院与中国网络空间研究院发布的《国家数据资源调查报告(2022)》。

社会的创业动能更是深刻地扎根于经济社会发展的结构性变革之中(李宏彬等,2009;Kritikos,2015)。那么,作为国家重大战略举措的公共数据开放能否激发社会创业活力?其内在作用机制是什么?又有哪些因素限制或者强化公共数据开放的创业效应?针对上述一系列问题的剖析,将有助于深入理解公共数据开放战略的理论意义与实践价值,也从经验上为推进数据要素市场化建设,充分发挥数据要素作用这一命题提供政策启示。

创业是企业家重新组合生产要素,将资源与生产要素开拓为价值更高的新产品、新市场、新技术的“创造性破坏”过程(Schumpeter,1934)。但创业并非易事,政府逻辑与市场逻辑共存和互动的复杂环境关系是影响创业活动的重要因素(陈刚,2015;杜运周等,2020)。完善的市场机制为企业家开展创业提供有效信息与价值激励,以实现资源配置最优化。但是市场机制诱致的垄断、外部性与信息不对称等问题,将导致资源配置中的市场失灵。这一现象为公共利益理论者宣称政府管制能够避免市场失灵以实现资源配置效率更优化并提高社会福利提供了充分理由(庇古,2007)。但是公共选择理论认为,为维护在位企业的垄断地位与垄断租金,政府可能主动创造新企业进入租金与抽取租金的机制,政府管制将导致社会资源配置的低效率(De Soto,1989;Shleifer和Vishny,1993)。这一点在中国体现得更加明显,世界银行数据显示,2012年前,中国新企业进入市场至少要承担38天和149美元的制度性交易成本^①。此外,在信息不完全的现代市场体系下,企业家面临着较高的搜寻有效且准确信息而必须付出的财力、时间与承担风险的显性成本,尤其是大数据时代数据资源获取限制、数据垄断、数据产权模糊等引致的信息不对称,将造成数据要素市场配置资源高效流通机制失灵。

公共数据开放是推进数字政府建设与数字经济战略的重要举措。一方面,公共数据开放是变革政府治理模式、透明政府治理规范、融通公众与政府协同互动的重大战略。从政府治理逻辑来看,通过公开政府履行公共职责过程中产生、收集、整理的政务数据资源,在优化政务服务效率、推进公众与政府协调治理模式创新等方面具有重要意义,在深化简政放权改革、优化营商环境上发挥了重要作用(段盛华等,2020;Park和Gil-Garcia,2022)。一个典型案例是,截至2021年北京市通过公共数据资源平台向社会免费无条件开放共计579万条涉及公共服务事项与政府管理等领域数据,推动政府治理体系公开透明,提升公共服务有效性。另一方面,数据资源深刻渗透于经济社会发展的各个环节,变革市场主体生产模式、商业模式与思维模式。从市场效益逻辑来看,相较于传统公共基础资源,公共数据要素呈现出复制性强、公共品属性与非稀缺性资源等典型特征(陈晓红等,2022)。公共数据开放的海量数据中包括工商行政处罚、企业注册信息、客户信用状况、失信被执行人信息等具有重大公共利益的企业数据,由此降低市场主体可决策性数据资源获取的不对称性,缓解市场失灵。综上所述,从政府治理逻辑与市场效益逻辑出发,深入探讨公共数据开放对企业家创业动机的影响是一个重要的实证问题。然而,关注公共数据开放对经济社会的价值创造效应的实证分析还处于起步阶段,而且现有文献重点基于公共管理学领域视角,关注公共数据开放的内涵概念辨析、结构治理等方面的理论思辨(郑磊,2018;朱峥,2020;段盛华等,2020;胡峰等,2023),或者集中于关注公共数据开放的制度基础、开放程度与决定性因素等方面(Zuiderwijk和Janssen,2014;Donker和Loenen,2017;Jetzek等,2019;Wang和Lo,2020;Zhao等,2022)。公共数据开放对创业的影响缺乏系统性研究,尤其是在基于政府治理与市场效益的共生逻辑视角上。

基于此,本文以各城市公共数据平台上线作为一项准自然实验,利用2007~2021年296个地级市及以上城市的非平衡面板数据和工商企业注册数据,尝试基于政府治理与市场效益的共生逻辑视角,考察

^① 数据来源:《世界银行营商环境报告》。

公共数据开放对城市创业活力的影响及其作用渠道。实证发现,第一,公共数据开放显著促进城市创业活力提升。第二,机制检验结果表明,降低制度性交易成本和优化信息搜寻效率是公共数据开放促进城市创业活力提升的经济机制。具体而言,一方面,公共数据开放能够强化公众监督问责机制,提升政府公共服务效率,降低制度性交易成本,激发城市创业活力提升;另一方面,公共数据开放通过优化信息搜寻匹配效率,强化企业家客户锚定、市场需求定位与创业机会感知,为企业家创造了机会与条件,进而对提升城市创业活力产生积极影响。第三,异质性分析表明,区域要素集聚能力与非正式制度环境强化了公共数据开放对城市创业活力的促进作用,而且公共数据开放弥补了初始地区数字基础设施政策支持差距,且对生产性服务业的创业促进作用更显著,表明公共数据开放还进一步促进高质量创业。

本文的边际贡献体现在以下几点:第一,聚焦于公共数据开放推进数字政府建设与数字经济战略的功能定位,实证分析了公共数据开放对城市创业活力的影响,丰富了公共数据开放相关研究成果,也为深入理解公共数据开放战略的理论意义与实践价值提供经验证据。第二,进一步探讨了公共数据开放影响城市创业活力的内在机制,重点关注了制度性交易成本降低的政府治理逻辑与信息搜寻匹配效率优化的市场效益逻辑对创业的影响,解析公共数据开放战略通过政府治理与市场效益逻辑,促进城市创业提升,激发国内经济活力的传导路径,加深对激发社会创业问题的理解。同时,也为大数据时代进一步理解“有为政府”与“有效市场”的有机结合在数据要素市场化建设进程中的作用提供了新的理论逻辑。第三,进一步探讨了区域要素禀赋、非正式制度环境、地区初始数字基础设施政策支持差异与创业行业类型差异等协调因素的影响效应,对于理解公共数据开放价值创造效应、完善制度设计以提升城市创业活力具有重要的政策价值。

二、文献综述

相关研究主要体现在以下几方面:首先,本文的研究与政府角色的创业效应文献相关。经典文献强调,政府干预是影响创业活动的重要因素(Arrow, 1962; Romer, 1986)。因为企业家进行“创造性破坏”的重要激励在于资本追逐利润,但是市场机制的自发调节效应在有限理性与信息不完全条件下会产生市场失灵的结果,这将导致企业家资本逐利动机降低。在中国经济社会转型时期,较强的政府管制导致的较高制度性交易成本严重阻碍了创业活动的开展(陈刚, 2015),因为企业家会通过识别政府政策措施,判断制度与政策变化倾向对创业可行性以及创业回报的影响,进而决定是否开展创业活动(Baumol, 1990; Krueger和Brazeal, 1994)。因此,既有研究强调了政府应扮演协调者的角色,通过降低行政审批壁垒,优化政府行政职能,营造良好的企业进入的营商环境,降低企业家创新创业的制度性交易成本,由此激发企业家开展创业活动的内在动力(Lim等, 2010; 陈刚, 2015; 杜运周等, 2020)。

其次,本文的研究与数据要素的价值创造效应的文献紧密相关。既有文献主要围绕数字化转型、数字技术创新与数据要素利用等视角,探讨企业将数据要素纳入生产决策所获得的效益(蔡继明等, 2022、2023; 黄勃等, 2023)。随着以大数据为基础的数字经济在中国蓬勃发展,数字经济发展的创业激发效应得到了学者们的关注。一方面,数字经济持续扩散并广泛渗透于经济社会的各个领域,拓宽了企业家的创业资源获取渠道,激发企业家进行创业的热情(赵涛等, 2020),尤其是数字普惠金融兼具普惠性与包容性特征,对低社会资本的企业家的创业概率提升作用更显著(张勋等, 2019);另一方面,企业家可以从开放的数字平台中挖掘并识别商业机会与商业模式(Sigfusson和Chetty, 2013; Ross和Blumenstein, 2015)。另外一支文献探讨了数字政府建设的逻辑,认为大数据重构与强化了政府治理能力,驱动外部主体与政府协同合作格局的创新治理模式(赵云辉等, 2019; 段盛华等, 2020; 胡峰等, 2023);少量研究关注了数字治理的经济效应(何雨可等, 2023)。

最后,本文的研究与公共数据开放的文献最为紧密。2001年开始,中国政府推行了一系列旨在降低政府干预、优化政府职能的改革,如行政审批改革与商事制度改革,相关研究深入探讨了这两项改革对企业进入、企业创新、就业增长与经济增长等方面的影响(夏杰长和刘诚,2017;毕青苗等,2018;王永进和冯笑,2018;夏杰长和刘诚,2020;朱光顺等,2020;杜运周等,2020;毕青苗等,2024)。而公共数据开放是大数据时代将数据要素纳入治理体系创新模式的深化,通过强化市场主体与政府的协同治理,倒逼政府职能优化(郑磊,2018;朱峥,2020;段盛华等,2020;胡峰等,2023)。目前,关于公共数据开放的研究主要集中于沿用公共管理学的研究范式,定性地探讨公共数据开放的制度基础、开放程度、决定因素等。其中,制度基础方面的文献强调公共数据开放的表现形式在于以公众知情权为权利基础(朱峥,2020);开放程度方面的文献主要基于构建多情景、多因素的评估模型展开(Zuiderwijk和Janssen,2014;Donker和Loenen,2017);公共数据开放的决定因素则包括政府竞争压力与媒体压力等外部环境压力以及领导者参与和文化特征等内部决策者因素(Jetzek等,2019;Wang和Lo,2020;Zhao等,2022)。实证上,少量研究分析了公共数据开放对区域协调发展、企业全要素生产率的影响(方锦程等,2023;彭远怀,2023)。

综上所述,现有研究丰富地探讨了政府行为、数字经济等方面的经济效应,但尚未充分关注到兼具数字政府建设与数字经济战略意义的公共数据开放的创业效应。公共数据要素兼具政府治理逻辑与市场效益逻辑,但既有研究往往考虑数据要素的单方面效应,难以全面揭示其价值创造效应的内在逻辑机制。本文将从政府治理逻辑与市场效益逻辑的视角识别公共数据开放对创业的影响,基于制度性交易成本降低与信息搜寻匹配效率优化的途径探讨其影响创业背后的作用机制。

三、政策背景与研究假设

(一)政策背景

《中共中央 国务院关于构建数据基础制度更好发挥数据要素作用的意见》将公共数据界定为,各级党政机关、企事业单位依法履职或提供公共服务过程中产生的数据资源。充分激发公共数据开放的价值创造效应是中国政府的重要战略。2015年,《促进大数据发展行动纲要》首次强调了推动政府数据开放对提升政府治理能力与推动经济转型发展的重要意义;2021年,《国民经济和社会发展第十四个五年规划和2035年远景目标纲要》明确提出“加强公共数据开放共享”,“激活数据要素潜能,推进网络强国建设,加快建设数字经济、数字社会、数字政府,以数字化转型整体驱动生产方式、生活方式、治理方式变革”;2022年,《要素市场化配置综合改革试点总体方案》将完善公共数据开放共享机制列为探索数据要素市场化配置的关键环节。

在顶层设计指导下,各地方政府也积极探索公共数据开放的政策实践,推进公共数据开放平台上线建设,作为公共管理和服务机构之间因履行职责和提供公共服务产生的公共数据资源的载体。目前,公共数据开放平台建设主要有三种模式:一是省级统一建设,各地级市只需使用省级平台统一分配的账号开展公共数据开放填报与信息维护工作。例如,《山东省公共数据开放办法》指出,应依托省一体化大数据平台,建设统一的公共数据开放平台。二是省市分别建设,市级在省级平台开放目录框架中开放资源分散存储,在省级主管部门统一的标准约束下建设各地级市公共数据开放平台。例如,《广东省公共数据开放办法》指出,地级市以上公共数据主管部门应当根据国家和省公共数据分类分级相关规定,增补本地公共数据的分类分级规则。三是省市独立建设,省市平台间并无实质性关联,这也是目前各地方建设公共数据开放平台中普遍存在的模式。在开放容量上,省级公共数据开放平台数据资源总量可能更大,但是单个公共数据集平均总量可能小于市级。2017~

2021年,广东省公共数据开放总量远大于东莞市,但是东莞市单个公共数据集平均容量远超广东省^①。在上线时间上,部分市级公共数据开放平台早于省级。佛山市、湛江市、无锡市、武汉市于2014年上线,但广东省、江苏省、湖北省分别于2016年、2019年、2020年上线。截至2023年,陕西省、青海省内的所有城市以及海南省、河南省、吉林省、辽宁省和云南省内的绝大多数城市尚未上线公共数据开放平台。由此看出,省市独立建设是公共数据开放平台中普遍存在的模式。

从城市公共数据开放实践来看,上海市、北京市于2012年率先探索了公共数据开放平台上线实践,无锡市、佛山市、湛江市、武汉市等地于2014年上线,2015年《促进大数据发展行动纲要》发布后,城市公共数据开放平台数量显著增长。截至2023年底,城市公共数据开放平台数量已达到204个,占比达到60.53%,如附图1所示^②。在开放进展上,从2018~2021年城市开放数据的持续水平来看,前5名分别是上海市、贵阳市、青岛市、深圳市与济南市,而北京市排到第10名,广州市排到第13名^③。由此可见,相对落后城市也能持续、稳定地向社会提供公共数据资源。在开放内容上,公共数据平台主要开放包括社区治理、城乡建设、监管处罚、行政审批、企业注册信息、工商用户基本信息、信用体系失信、被执行人信息等与政府治理相关数据,以及具有重大公共利益的企业数据。为了充分激发公共数据的创新应用,武汉市、贵阳市、深圳市、哈尔滨市等城市开办了开放数据创新应用大赛,通过面向公众开放数据资源,挖掘数据价值,深化数据资源创新应用,以培育公共数据的衍生能力,激发创新潜力,集聚创业资源。一个典型案例是,2019年贵阳市开放数据应用创新大赛,通过运用公共数据开放平台提供的数据资源,参赛团队创新创造智慧化产品(原型)及解决方案,其中“房叮当”等项目已落地开展市场化应用,实现了数据资源的创新应用。由此可见,公共数据开放对激发城市创业活力具有重要意义。

(二)理论分析与研究假设

公共数据资源兼具政府逻辑与市场逻辑的共存和互动的复杂环境,此部分,将从政府治理逻辑与市场效益逻辑探讨公共数据开放对城市创业活力的影响机制。

第一,从政府治理逻辑来看,公共数据开放有助于优化政府行政职能与公共服务体系,降低市场主体创业面临的制度性交易成本。制度性交易成本是人为因素与交易环境交互影响下所产生的市场失灵现象。新制度经济学派强调,通过推进良好制度环境建设,降低制度性交易成本,对激发创业具有重要而积极的作用(Acemoglu等,2005)。公共数据中涵盖了政府部门在履行公共职责时产生的制作、收集与获取的数据资源储量,且能够更加充分地反映政府政策的真实信息(郑磊,2018)。一方面,市场主体通过收集整理与识别利用,能够更加有效地监督政府制度决策的产生机制、实施过程以及政策结果,形成并强化公众监督和参与的问责机制,倒逼地方政府提高管理运行的透明度,进而优化制度环境、畅通市场运行机制与提升资源配置效率(Zuiderwijk和Janssen,2014; Park和Gil-Garcia,2022;彭远怀,2023)。以厦门市公共数据开放平台为例,其向市场主体公布了行政审批统一服务审批过程信息、智慧审批共享文件实例信息、公共资源交易网招标项目信息,以此增强政府管理透明度,降低企业家创业的制度性交易成本。另一方面,市场主体对公共数据利用的反馈有助于形成双向互动治理模式,提升市场主体的内生治理效率,实现政府协调治理模式创新(段盛华等,2020;胡峰等,2023),继而进一步促进政务服务体制、简政放权、放管结合等公共政策的创新(Park和Gil-Garcia,2022),降低制度性交易成本。

第二,从市场效益逻辑来看,公共数据开放有助于优化市场主体的信息搜寻匹配效率,为企业家

① 参见复旦大学数字与移动治理实验室发布的《2021中国地方公共数据开放报告(城市)》。

② 本文附录详见《数量经济技术经济研究》杂志网站,下同。

③ 参见复旦大学数字与移动治理实验室发布的《2021中国地方公共数据开放报告(城市)》。

开展创业活动创造机会与条件。数据要素时代信息不完全现代市场体系下获取可决策性信息资源限制,以及数据要素市场不完善配置资源流通机制低效率的失灵现象,加剧企业家与消费者之间的信息不对称,不利于企业家识别客户、挖掘市场需求,进而降低企业家创业激情。公共数据开放为企业家实现信息资源的高效流通与实时共享,提升信息协同整合效率。根据动态能力理论,企业家能整合、构建和重新配置内外资源,增强动态复杂变化的适应环境能力,进而创造持续竞争优势(Teece等,1997;Augier和Teece,2009)。一方面,企业家通过分析种类多且实时性强的客户信用数据、工商行政处罚等公共数据,有效地降低了客户有限理性与减缓交易对象的机会主义行为,缓解行为主体摩擦问题引致的信息不对称(李唐等,2020;Farboodi和Veldkamp,2021);还进一步拓展了企业家搜寻消费者信息的潜在范围,有效地缓解供应链“长鞭效应”,优化客户信息传递机制。相关的案例包括,厦门市公共数据开放目录中涵盖了社会信用业务备案信息、法人基本信息、行政处罚信息等信用服务数据;河池市公共数据开放平台公布了失信被执行人信息等数据。企业家可以通过识别潜在的消费者信用与创业信息,有效地降低主体摩擦的信息不对称。另一方面,企业家能够从公共数据资源中挖掘企业进入选址信息,把握消费者偏好、企业创新销售等一系列信息进行商业性经济活动(Ghasemaghaei和Calic,2019、2020;Nagaraj,2022),提高企业家面临的市场需求评估效率。

此外,公共数据开放实质是数字技术创新的拓展应用,大数据数字技术创新的产品与流程能提升客户满意度与忠诚度(Balci,2021),进而促进企业家对市场机会的感知与把控。数字技术创新的产出成果正是企业家将数据要素纳入原有均衡体系的“创造性破坏”,这将进一步为其创造可开发、识别与利用的创业机会(Shane和Venkataraman,2003)。企业家能够将数字技术创新成果中大量隐含的创业机会与消费者需求信息整合匹配,并将其推向市场,实现公共数据开放的商业化价值(Shane和Venkataraman,2003;Nagaraj,2022)。

基于上述分析,提出如下待检验研究假设:公共数据开放可以通过降低制度性交易成本的政府治理逻辑与优化信息搜寻匹配效率的市场效益逻辑,显著促进城市创业活力提升。

四、研究设计

(一)模型设定

本文将各城市公共数据开放平台上线建设的个体与时间差异视为一项准自然实验,构建多期双重差分模型(Difference-in-Differences, DID),系统评估公共数据开放对城市创业活力的影响,具体模型设定为:

$$entre_activation_{it} = \beta_0 + \beta_1 open_{it} + \gamma controls_{it} + \eta_i + \mu_t + \varepsilon_{it} \quad (1)$$

其中,下标*i*表示城市,*t*表示年份;被解释变量 $entre_activation_{it}$ 表示城市*i*在第*t*年的创业活力;解释变量 $open_{it}$ 表示公共数据开放的虚拟变量,倘若城市*i*在第*t*年及以后上线公共数据开放平台则赋值为1,否则为0; $controls_{it}$ 表示影响城市创业活力的控制变量集合; η_i, μ_t 分别表示城市固定效应与年份固定效应,加以控制基准回归中不随城市与年份变化的混淆因素; ε_{it} 为随机误差项。此外,考虑到残差序列相关以及异方差造成的估计偏差,将所有回归标准误差聚类到城市层面。最关注的系数是 β_1 ,其衡量公共数据开放对城市创业活力影响的净效应,倘若 β_1 显著为正则表明公共数据开放能有效激发城市创业活力。

(二)变量定义

1.城市创业活力

借鉴杜运周等(2020)、袁礼和龚钰涵(2023)的研究思路,考虑到城市内企业规模异质性差异而产生的度量偏误,将城市人口作为标准化基数,采用城市每百人口中的新注册企业数量衡量城市创业活力。

2. 公共数据开放

公共数据开放的政策处理变量(*open*)表示城市*i*在第*t*年及以后是否上线公共数据开放平台。参考复旦大学数字与移动治理实验室发布的《2023中国地方公共数据开放报告(城市)》,整理出各城市是否上线公共数据开放平台的基础数据,并且将其与华中师范大学信息管理学院发布的《中国政府开放数据利用研究报告(2022)》进一步对照。同时,借鉴复旦大学数字与移动治理实验室发布的《2023中国地方公共数据开放报告(城市)》的方法,通过以下两步进一步进行检验:首先,将各地区公共数据开放平台域名限定于gov.cn,将其作为确认官方公共数据开放平台的依据,且必须由地级市以上行政级别的政府部门负责建设与运营,同时通过开设专门且统一平台或者由政府部门在官网设置专栏的形式开放;其次,通过在搜索引擎中对各地区名称加关键词(“数据开放”“数据资源”“公共资源”等)加以验证。据此,最后获得了截至2023年204个城市(含有直辖市、副省级城市与地级市)的公共数据平台上线时间,将上线公共数据平台的城市设定为*treat*,将首次上线时间以及之后年份设定为*post*,进而构建城市一年份层面城市是否上线公共数据开放平台的虚拟变量 $open(treat \times post)$ 。

3. 控制变量

城市层面控制变量包括:①经济发展水平。经济发展水平与生产要素集聚相关,影响企业家创业的吸引力与收益情况,采用城市人口密度的自然对数量化城市经济发展水平。②产业结构升级。资源禀赋基础与要素投入条件的差异是影响城市创业结构的重要因素,尤其是服务业与金融业的发展能够为创业活动开展提供服务型技术、人才与资金的支持,采用第三产业增加值与第二产业增加值比值衡量产业结构升级。③人力资本水平。人力资本是人才与技术集聚的重要因素,而人才与技术能够为激发城市创业活力提供高质量的要素,以人均普通本专科在校学生数衡量。④公共服务水平。公共服务关乎人才宜居、社会保障等多方面,也为企业进入提供坚实的基础条件,以每百人公共图书馆藏书量衡量城市公共服务水平。⑤传统基础设施。交通便利性与实体物流交易成本的节约为市场主体进入市场提供了较强的吸引力,采用铁路货运总量的自然对数量化城市传统基础设施。⑥互联网发展水平。互联网加速市场主体信息的传播速度,降低创业的信息成本,互联网基础设施建设还为创业运营模式提供了支持,采用电信业务总量的自然对数衡量互联网发展水平。⑦金融发展水平。金融资源可得性为企业家创业提供信贷资金来源,也有助于吸引生产要素集聚,采用年末金融机构各项贷款余额与地区生产总值比重衡量。⑧财政压力水平。财政压力水平有助于理解地方政府根据居民偏好与资源禀赋提供公共产品的行为策略,采用各地方财政一般预算内支出占地方财政一般预算内收入比重量化。⑨科学支出水平。科学支出反映地方政府对于科技创新环境与成果的重视,还为生产性服务业聚集提供良好的外部条件,以科学支出对数衡量城市科学支出水平。

此外,考虑到各城市公共数据开放平台上线行为可能不具随机性,各地方政府可能需要结合地区的经济发展水平、地理区位以及资源禀赋等条件作出抉择,这些因素随着时间变化可能对城市创业活力产生影响,从而导致采用多期双重差分模型识别假设不成立。因此,借鉴Edmonds等(2010)、赵涛等(2020)的研究思路,进一步纳入1995年人均粮食产量、1995年人均油料产量、1995年人均蔬菜产量、1995年人均普通高等学校数量以及时间趋势交互项,以控制影响城市是否建设公共数据开放上线平台的原有社会经济特征等,进而缓解实验组选择的估计偏差。

(三)数据来源

数据来源于以下几个方面:首先,城市数据来源于2008~2022年《中国城市统计年鉴》的面板数据,在剔除严重缺失值与无法识别的样本之后,进一步采用插值法对缺失值进行补充。其次,新注册企业数量来源于工商注册数据库,从市场监管总局网络获取了包括新注册企业的名称、类型、注

册地址、注册年份、注册资本和主营业务等基本信息在内的工商注册企业数据,利用注册地址以及注册年份,将工商注册企业数据合成得到城市一年份维度的面板数据。考虑到异常值的影响,将所有连续变量进行上下1%缩尾处理。各变量描述性统计如附表1所示。

五、实证结果分析

(一)基准回归结果

表1报告了公共数据开放对城市创业活力影响的回归结果。其中,第(1)列为控制城市固定效应与年份固定效应基础上的单变量回归,公共数据开放政策处理变量的回归系数在1%水平上显著为正,初步表明公共数据开放对所属城市的创业活力具有显著的正向影响。考虑到经济、教育、文化、基础设施与政府行为等特征对城市创业活力的影响,在前述单变量回归的基础上控制了经济发展水平、产业结构升级、人力资本水平、公共服务水平、传统基础设施、互联网发展水平、金融发展水平与财政压力水平等变量。第(2)列的回归结果显示,公共数据开放政策处理变量的回归系数仍然在1%水平上显著为正。此外,考虑到城市初始禀赋特征的基准变量可能影响城市是否选择开放公共数据,在前述基础上,进一步控制城市事前变量与时间趋势交互项。第(3)列的回归结果显示,核心解释变量的回归系数仍然通过1%水平检验显著为正,回归系数小于前述回归结果,说明较为有效地控制了城市事前变量与时间趋势交互项对政府是否开放公共数据对所属城市创业活力的冲击。根据第(3)列的回归结果,其经济含义在于,在其他条件不变的情况下,与尚未开放公共数据的城市相比,公共数据开放城市的每百人新注册企业数量将平均增加19.210家/万人,相当于均值的17.251(0.192/1.113)家/万人。上述回归结果意味着,公共数据开放显著促进城市创业活力提升,激发经济增长活力,这也意味着公共数据开放有助于降低制度性交易成本与优化信息搜寻匹配效率的理论逻辑可能成立。

表1 基准回归结果

| 变量 | 创业活力 | | |
|------------------|---------------------|---------------------|---------------------|
| | (1) | (2) | (3) |
| <i>open</i> | 0.268*** (0.071) | 0.205*** (0.068) | 0.192*** (0.066) |
| 控制变量 | 否 | 是 | 是 |
| 事前变量 | 否 | 否 | 是 |
| 固定效应 | 是 | 是 | 是 |
| 样本量 | 4343 | 4343 | 4343 |
| R ² 值 | 0.786 | 0.797 | 0.806 |

注: *、**、***分别表示在10%、5%、1%的水平上显著,括号内为聚类到城市层面的稳健性标准误。

(二)平行趋势检验

多期双重差分模型有效的一个重要前提假设是,公共数据开放前实验组与控制组间的城市创业活力的趋势不随时间变化而发生改变,整体态势应保持一致。本文构建如下回归方程进行平行趋势检验:

$$entre_activation_{it} = \beta_0 + \beta_k \sum_{k=-14}^9 open_{it}^k + \gamma controls_{it} + \eta_t + \mu_i + \varepsilon_{it} \quad (2)$$

其中, k 表示公共数据开放的第 k 年,由于公共数据开放最早于2012年,样本截至2021年, k 的最大取值为9,最小取值为-14。为避免多重共线性问题,以公共数据开放前一年为基期,系数 β_k 则表示相较于公共数据开放前一期,开放的第 k 年实验组与控制组城市创业活力的差异。如图1所示,在95%置信区间条件下,公共数据开放前的回归系数均在0附近波动且不显著,表明公共数据

开放前实验组与控制组的城市创业活力未呈现明显的趋势差异,满足平行趋势检验假设。公共数据开放之后政策效果存在滞后一期,随后政策效果逐渐加强并维持到较长的第3期,之后政策效果波动直至消失。因此,未来应更注重公共数据开放对城市创业活力影响的持续性。

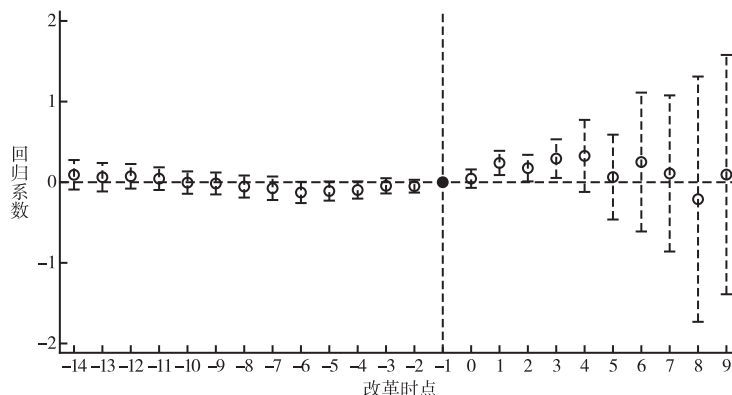


图1 平行趋势检验结果

(三) 异质性处理效应诊断与检验

多期双重差分模型双向固定效应估计量的异质性处理效应,可能导致同一处理对于不同个体产生效果差异性的偏差(刘冲等,2022;张子尧和黄炜,2023)。此部分,将在诊断双向固定效应估计量可能存在偏误的基础上,进行“异质性—稳健”估计量检验。

首先,采用Dacon分解进行偏误诊断。Goodman-Bacon(2021)认为,可以将 2×2 -DID估计量划分“新处理个体”以“未处理个体”为对照组、“新处理个体”以“尚未处理个体”为对照组、“新处理个体”以“已处理过的个体”为对照组三种情况。其中,第三种情况已含有处理效应,其匹配的对照组不具有科学性。根据Goodman-Bacon(2021)分解法进行诊断,在平衡面板数据下,从附表2发现,时变处理的实验组将“从未处理个体”视为对照组的权重约为68%;结合附图2结果可知,双向固定效应overall DD估计量为0.213,组内差异所占权重数值仅为2.93%,表明处理效应在组间存在一定程度的异质性,但不会对估计结果造成严重的偏差。

其次,采用De Chaisemartin和D’Haultfoeuille(2020)分解进行诊断。附表3汇报了分解结果,可以发现总权重为445,包括443个正权重和仅2个负权重,负权重占比仅为0.45%。因此,上述多期双重差分模型异质性处理效应诊断结果表明本文估计量并未存在严重偏误问题。

最后,基于Borusyak等(2022)、Cengiz等(2019)提出的方法进行检验。Borusyak等(2022)提出的“插补估计量法”以尚未接受政策影响的个体估计接受政策影响个体的反事实结果,然后将接受影响个体的真实结果与反事实结果相减,由此得到正确的处理效应估计结果。Cengiz等(2019)提出的“堆叠DID”将数据集重构为相对事件时间的平衡面板数据,在控制群组效应与时间固定效应的基础上,得到处理效应的加权平均值可以解决双向固定效应估计偏误。附图3与附图4显示了两种“异质性—稳健”估计量的事件研究图,可以发现在公共数据开放前,实验组与控制组的城市创业活力间并无显著趋势,而在公共数据开放后的部分年份,其显著促进城市创业活力提升。

(四) 稳健性检验

1. 预期效应排除

市场微观主体会对政府政策倾向产生预期效应,如根据公开政策文件预期政府事后的政策倾向、基于政府实施的政策形成对其他政策试点的预期或者基于政府部门工作会议对其政策进行判

断。为此,构建公共数据开放政策变量的实施前一年的变量 $open_1$,将其纳入基准模型回归以控制市场主体预期效应对其创业行为的影响。附录表4第(1)列结果显示,公共数据开放实施前一期变量 $open_1$ 不显著,而公共数据开放政策变量仍然显著为正;此外,进一步构建公共数据开放实施的前二期变量 $open_2$,将其与 $open_1$ 共同纳入基准模型回归。附录表4第(2)列结果显示,政策前变量 $open_1$ 、 $open_2$ 均不显著,而公共数据开放政策变量仍然显著为正,说明市场主体尚未对公共数据开放形成预期效应,也进一步地表明平行趋势检验结果的可靠性。

2. 安慰剂检验

考虑到不可观测的其他遗漏因素对城市创业活力的冲击,借鉴马述忠等(2023)的思路,采用间接安慰剂检验方法进一步检验。其基本逻辑在于从理论上寻找一个不会对城市创业活力产生冲击且具有随机性的错误政策变量 $open'$,从而产生一个错误估计系数 $\widehat{\beta}_1'$,重复上述回归过程1000次,并得到1000个系数 $\widehat{\beta}_1'$ 估计值的核密度分布。由附录图5所示,安慰剂检验中的回归系数 $\widehat{\beta}_1'$ 均值落在0附近并且近似服从正态分布,绝大部分回归系数并不显著,由此间接证明了不可观测的遗漏变量并不会影响结论的可靠性,符合安慰剂检验预期。

3. 对内生性问题的进一步处理

前述基准回归中控制了相关城市特征变量,还控制了前定变量与时间交互项以缓解实验组选择的估计偏差,并且通过了平行趋势检验,在多期DID异质性处理效应诊断与检验下仍然稳健成立,但公共数据开放城市与非公共数据开放城市在某些特征上呈现系统性差异,由此造成的内生性问题仍然对回归结果产生偏误。接下来,将采用多种方法对潜在的内生性问题做进一步处理。

(1)排除其他可能政策的影响。样本期间内涵盖了政府职能改革、政府数字治理以及数字经济发展等政策,可能冲击回归结果的可靠性。其一,政府职能改革。公共数据开放强化民众与政府协调治理,优化政府公共服务效率。相关政策包括深化行政审批制度改革,其目的在于最大限度地降低政府对微观事务的管理,放宽市场准入,激发市场主体活力;相关政策还包括自贸区制度,2019年《国务院关于在自由贸易试验区开展“证照分离”改革全覆盖试点的通知》要求各自贸区实现负面清单管理,按直接取消审批、审批改为备案、实行告知承诺、优化审批服务等方面推进改革,全面贯彻“证照分离”的改革。借鉴毕青苗等(2018)的做法与收集整理各城市设立自贸区的数据,设定试点城市 i 在第 t 年实施相关政策的政策虚拟变量纳入共同回归。附录表5第(1)列结果显示,纳入上述两个冲击政策因素后,公共数据开放的政策变量仍然显著为正。

其二,政务服务透明化与信息化影响。2014年的“信息惠民”政策、2016年的“互联网+政府服务”政策等同属于数字政府建设的环节。构建以上两个数字政府建设政策的试点城市 i 在第 t 年实施相关政策的政策虚拟变量纳入共同回归。附录表5第(2)列结果显示,控制相关数字政府建设冲击变量后,公共数据开放的政策变量仍然显著为正。

其三,数字经济政策冲击。公共数据开放是推进数字经济战略的重要抓手。同期内,政府部门部署了多项战略政策,以提高数字基础设施,建设新型数字公共服务。相关政策包括“国家创新型城市”“宽带中国”“大数据试验区”等,进一步构建以上三个数字经济政策的试点城市 i 在第 t 年实施相关政策的政策虚拟变量纳入共同回归。附录表5第(3)列结果显示,控制数字经济政策冲击后,公共数据开放的政策变量仍然显著为正。最后,将所有相关变量共同与公共数据开放变量纳入基准模型中回归。附录表5第(4)列结果显示,纳入上述可能影响因素后,公共数据开放的变量仍然显著为正,研究结论具有较强稳健性。

(2)考虑城市的时间变化趋势。基准回归控制了一系列可能影响公共数据开放的事前变量,但

不同城市的历史条件、资源禀赋和区位特征各不相同,这将导致具有时变特征的城市之间存在差异,进而影响前述基准回归结果的准确性和可靠性。为此,在基础模型中加入了城市一年度趋势项,以控制城市组别间固有特征年份趋势对城市创业活力的影响。附录表6第(1)列结果显示,在纳入城市一年度趋势项之后,公共数据开放对城市创业活力的影响仍然显著为正,基准回归结果仍然稳健。

(3)考虑省级公共数据开放平台的影响。在中国上下垂直的行政管理体制下,省级政府对辖区内城市发展具有重要的影响。市级公共数据开放平台上线时间与省级公共数据开放平台有一定的重合,而省级层面平台可能影响管辖区内市级平台上线行为,进而导致识别和回归结果有偏。为检验结果的稳健性,首先,在基准模型基础上控制省份与年份交互固定效应,以考虑省级公共数据开放平台及其他因素对回归结果的冲击。附录表6第(2)列结果显示,公共数据开放的回归系数与显著性有所下降,但市级公共数据开放的创业促进效应在控制省级平台的影响后仍然显著存在。其次,考虑到直辖市公共数据开放平台在地级市样本中的冲击,剔除直辖市样本进行检验。附录表6第(3)列结果显示,公共数据开放的回归系数仍然显著为正。最后,考虑到市级公共数据开放上线受到省级影响才设立的可能情况,重新构建一个回归样本。该样本中实验组为那些省级公共数据开放平台上线滞后于市级,或者完全尚未开放省级但开放市级层面平台的样本,控制组保持不变,将此样本重新进行检验。附录表6第(4)列结果显示,公共数据开放的回归系数仍然显著为正。上述结果意味着,尽管省级数据开放平台上线冲击市级公共数据开放对创业活力的作用,但是在尽可能的排除这一影响后,市级公共数据开放的创业促进效应仍然显著存在,回归结果仍然稳健。

(4)PSM-DID。多期双重差分法能够识别出政策实施的平均处理效应,但是公共数据开放可能存在选择性偏差导致的内生性问题,公共数据开放并非严格标准上的准自然实验。为检验基准回归结果的稳健性,进一步采用倾向得分匹配模型以缓解样本选择性偏差问题。首先,以经济发展水平、产业结构升级、人力资本水平、公共服务水平、铁路货物总量、电信业务总量以及一系列事前变量与年份交互项等控制变量作为所要匹配的变量;其次,采用卡尺为0.05的1:4邻近匹配法寻找所有满足公共数据开放共同支撑条件的最优控制组;最后,进一步重新采用多期双重差分法估计公共数据开放对城市创业活力的影响效应。附录表6第(5)列报告了采用满足共同支撑假设样本回归结果,公共数据开放回归系数仍然显著为正,基准回归结果具有稳健性。

(5)工具变量。前述研究从多个方面部分解决了存在的内生性问题,但仍然可能存在逆向因果或者遗漏变量等,导致估计结果偏误。例如,公共数据开放平台上线可能与地区经济发展水平显著正相关,而地区经济发展水平越高,其网络基础设施建设与信息透明度可能更高,这对于城市创业活力具有一定的促进作用。为处理逆向因果与遗漏变量问题,尝试选取工具变量进行两阶段最小二乘法检验。

地形特征是脱离经济系统的外生变量,地形特征通常被用于经济学研究中作为工具变量解决内生性问题。本文借鉴柏培文和张云(2021)的研究,采用地形起伏度与上一期全国互联网宽带接入端口对数交互项作为工具变量(IV)。其可行性在于:其一,地区地形的平坦程度与宽带网络、信息传输等网络基础设施建设直接相关,而公共数据开放平台建设的前提在于良好的网络基础设施,同时互联网宽带接入端口反映地区互联网服务供应规模,满足相关性假设;其二,地区地形起伏度是脱离经济系统的变量,对当期创业活跃度并无直接影响,上一期全国互联网宽带接入端口也不会直接影响市场主体创业行为,满足外生性假设。

如表2第(1)列与第(2)列所示,第一阶段系数为负,表明地区地形特征因素越陡峭,越不利于网络基础设施建设,对公共数据开放平台建设具有负向影响;第二阶段系数为正,符合预期,且Cragg-Donald Wald F值为79.253,大于经验法则10,排除弱工具变量问题。在此基础上,本文将检验工具变量

外生性条件,即工具变量仅通过影响公共数据开放作用于城市创业活力。借鉴方颖和赵扬(2011)的思路:其一,为证明工具变量仅在公共数据开放后才影响城市创业活力,本文剔除了2012年及后续样本,采用城市创业活力对工具变量进行回归,估计结果如表2第(3)列所示,验证在公共数据开放前,工具变量无法直接作用于城市创业活力。其二,为验证工具变量仅通过影响公共数据开放作用于城市创业活力,进一步将政策处理变量和工具变量同时与城市创业活力进行回归,如表2第(4)列所示,当控制公共数据开放政策时,工具变量对城市创业活力的估计系数不显著,而政策处理变量公共数据开放的系数仍然显著,说明工具变量不会通过其他渠道作用于城市创业活力,验证了工具变量外生性条件。

表2 工具变量回归结果

| 变量 | <i>open</i> (1) | 创业活力 (2) | 创业活力 (3) | 创业活力 (4) |
|------------------|----------------------|--------------------|------------------|---------------------|
| <i>IV</i> | -0.045*** (0.011) | | 0.010 (0.015) | -0.040 (0.025) |
| <i>open</i> | | 1.073** (0.541) | | 0.181*** (0.065) |
| Wald F 值 | 79.253 | | | |
| 控制变量 | 是 | 是 | 是 | 是 |
| 事前变量 | 是 | 是 | 是 | 是 |
| 固定效应 | 是 | 是 | 是 | 是 |
| 样本量 | 4288 | 4288 | 1415 | 4288 |
| R ² 值 | 0.508 | | 0.941 | 0.806 |

注:同表1。

(6)双重机器学习因果检验。选择性内生性问题仍然可能冲击回归结果的稳健性。此部分,借鉴Bodory等(2022)提出的双重机器学习法重新检验。与传统的因果推断方法相比,双重机器学习方法的优势体现在变量选择和模型设定上。一方面,企业家创业动机是经济社会互动与共生的复杂环境关系交互的结果,本文控制了多维度影响因素,但设定回归模型时可能因“维度诅咒”与多重共线性问题而导致估计结果偏误。通过机器学习及其正则化算法,双重机器学习方法将在预选项中自动筛选出精度较高的控制变量集合,由此避免高维控制变量冗余导致的“维度诅咒”,得到可靠的估计量。另一方面,双重机器学习的优势还体现在基于机器学习方法处理变量非线性问题上,由此可以缓解经济系统中变量非线性关系引致的模型设定偏误。

基于此,采用双重机器学习方法重新检验公共数据开放与城市创业活力间的因果关系。具体借鉴张涛和李均超(2023)思路,将双重机器学习方法样本分割比例设定为1:4,为了避免模型设定上的主观选择性,采用一般性交互式双重机器学习模型检验,使用随机森林算法预测方法,在控制城市与年份固定效应的基础上,逐步纳入控制变量集合的一次项、二次项与三次项。附录表7第(1)~(3)列结果显示,公共数据开放的回归系数均在1%水平上显著为正。此外,将双重机器学习方法样本分割比例设定为1:7,附录表7第(4)~(6)列结果显示,公共数据开放的回归系数均在1%水平上显著为正。上述结果表明采用双重机器学习方法缓解可能存在的内生性问题后,研究结论仍然稳健。

六、影响机制检验与异质性分析

(一)影响机制检验

1. 制度性交易成本

企业家创业面临的制度性交易成本主要体现在政府强管制、低效率治理与繁杂审批成本等方

面(陈刚,2015;毕青苗等,2018)。公共数据开放是推动数字政府治理的战略性部署,通过公开政府执行监督、管理、处罚等方面的数据记录,形成公众问责机制与优化政府运行效率,降低制度性交易成本。由于制度性交易成本概念涵盖范围较广,且定义不统一,对其直接衡量较为困难。现有文献多采用间接衡量方法反映制度性交易成本,比如采用上市公司层面的企业管理费用、销售费用与财务费用等占主营收入或者总资产的比重在城市层面的均值衡量制度性交易成本。然而,Cai等(2011)指出相关费用支出也包含了企业正常运行与维护企业正常关系的必要支出。

从以下三个方面构建测度指标,以检验制度性交易成本影响机制的存在性。首先,从政府服务效率视角刻画制度性交易成本。提供有效的政府服务是转变政府职能以激发市场活力的重要路径。借鉴刘大勇等(2023)的做法,采用《中国城市政商关系评价报告2021》的政商关系指数($zsgx$)、亲近指数($qjzs$)衡量。表3第(1)、(2)列结果表明,公共数据开放显著促进政府服务效率优化。其次,从政府治理效率视角刻画制度性交易成本。市场机制不健全与产权保护意识较弱将直接影响市场交易成本,优化政府治理效率能够提高信息透明度,降低市场主体面临的制度性交易成本(赵云辉等,2019)。借鉴赵云辉等(2019)的做法,以各城市年末总人数除以各城市公共管理、社会保障和社会组织就业人数作为政府治理效率的指标(gov)。表3第(3)列结果表明,公共数据开放显著促进政府治理效率优化。最后,从社会信用环境视角量化制度性交易成本。良好的社会信用是激发市场经济活力的前提条件,政府公信力与政务诚信更是直接影响到城市社会信用体系完善与运行效率。公共数据开放通过公开政府部门监督管理等收集的数据,强化公众问责机制,增强政府公信力与政务诚信,进而优化社会信用环境。本文采用中国城市商业信用环境指数量化政府公信力度($trust$)^①。表3第(4)列结果表明,公共数据开放能够增强政府部门公信力度,营造良好的社会信用环境。以上结果表明公共数据开放通过降低制度性交易成本,激发城市创业活力提升。

表3 制度性交易成本影响机制检验结果

| 变量 | $zsgx$ (1) | $qjzs$ (2) | gov (3) | $trust$ (4) |
|------------------|---------------------|---------------------|-------------------|---------------------|
| $open$ | 4.595*** (0.957) | 2.481*** (0.949) | 0.030* (0.015) | 0.967*** (0.222) |
| 控制变量 | 是 | 是 | 是 | 是 |
| 事前变量 | 是 | 是 | 是 | 是 |
| 固定效应 | 是 | 是 | 是 | 是 |
| 样本量 | 1111 | 1111 | 4339 | 3333 |
| R ² 值 | 0.822 | 0.825 | 0.930 | 0.832 |

注:第(1)、(2)列数据年限为2017~2020年;第(4)列数据年限为2010~2021年。其他同表1。

2. 信息搜寻匹配效率

大数据环境中信息搜寻匹配成本是企业进入市场的重要影响因素。从海量数据中,采集、挖掘、整合数据资源,并将数据资源与供给、需求进行精准对接,为企业家进入市场创造机会与条件。公共数据开放为企业家提供了海量免费的可决策性数据,优化企业家信息搜寻匹配效率。

为检验信息搜寻匹配效率机制的存在性,采用以下三种方法构建度量信息搜寻匹配效率的指标。首先,公共数据开放是推进数字经济战略的重要部署,有助于推动数字经济发展。数字经济发

^① 数据来源于《2020中国城市商业信用环境指数CEI蓝皮书》,其中2010年、2011年、2012年、2015年、2017年、2019年为原始数据,其余年份按照插值法补齐得到2010~2021年数据。

展水平较好的地区,企业家获取信息资源的难度越小,面临更低的信息搜寻匹配成本。借鉴赵涛等(2020)的研究思路,考虑到城市层面相关数据获取的限制,构建每百人互联网宽带接入用户数、计算机服务和软件业从业人员占城镇单位从业人员比重、人均电信业务总量和百人中移动电话用户数四个指标,通过主成分分析方法计算得到数字经济发展指数(*digital*)。表4第(1)列结果表明,公共数据开放显著促进城市数字经济发展水平提升,优化信息搜寻匹配效率。其次,信息获取可得性是降低市场主体信息壁垒,改善信息不对称程度的重要因素。结合中国地大物博以及数字经济时代的特征,以邮政业务总量除以行政区域土地面积量化每单位面积上的邮政部门为社会提供服务的可得性,以此量化城市内部的信息获取可得性(*yzym*)。表4第(2)列结果表明,公共数据开放显著增强信息获取可得性,降低城市内部的信息不对称。最后,由于地方政府权责被局限于一定的管辖区内,不同行政区划内的政策不一致性与竞争关系,阻碍信息资源在城市内的交流。因此,借鉴方锦程等(2023)的研究思路,以城市乡镇级行政区划数量除以城市总面积量化行政分割程度(*xzfg*)。表4第(3)列结果表明,公共数据开放显著降低行政分割程度,降低城市内信息壁垒。上述结果意味着,公共数据开放通过优化信息搜寻匹配效率,激发城市创业活力提升。

表4 信息搜寻效率影响机制检验结果

| 变量 | <i>digital</i> (1) | <i>yzym</i> (2) | <i>xzfg</i> (3) |
|------------------|-----------------------|---------------------|---------------------|
| <i>open</i> | 0.094** (0.040) | 0.229*** (0.060) | -0.001** (0.001) |
| 控制变量 | 是 | 是 | 是 |
| 事前变量 | 是 | 是 | 是 |
| 固定效应 | 是 | 是 | 是 |
| 样本量 | 4339 | 4343 | 4343 |
| R ² 值 | 0.725 | 0.918 | 0.970 |

注:同表1。

(二)异质性分析

1. 区域要素禀赋

地理区位特征与要素集聚程度相关,高质量生产要素(如人才)是社会创业的知识与创意产生、集聚与扩散的重要载体,尤其在制度环境不完善引致较高制度性交易成本的情况下,高质量人力资本开展创业活动的概率更大。公共数据开放推动政府服务透明化、强化民众问责机制,降低企业家面临的制度性交易成本,可能进一步激发高质量人力资本创业的内在动力。胡焕庸线是中国人口聚集程度与经济社会分布格局的分界线,其西北侧与东南侧人口分布与经济规模具有显著差异^①。将样本划分为胡焕庸线的西北侧与东南侧以表征地区聚集效应的高低,分样本探讨公共数据开放对区位特征异质的城市创业活力的影响。表5第(1)、(2)列结果表明,与胡焕庸线西北侧(低集聚)地区相比,公共数据开放对胡焕庸线东南侧(高集聚)地区的创业活力的促进作用更加明显。胡焕庸线西北侧地区的营商环境、数字互联网基础设施等发展程度较低,人才要素集聚吸引力也较低。而东南侧能够吸引要素资源在空间上集聚,在公共数据开放引致良好市场环境的加持下,高质量人才要素流入与资源配置效率优化,进一步强化公共数据开放的创业促进效应。

^① 胡焕庸线西北侧城市包括:新疆维吾尔自治区、青海省、甘肃省、宁夏回族自治区和内蒙古自治区(除通辽市、赤峰市外),其他则为胡焕庸线东南侧城市。

2. 非正式制度环境

既有研究指出,非正式制度通过深刻影响行为主体的信念偏好与行为决策,对创业选择产生重要影响(Zhang, 2020)。利用中国综合社会调查(Chinese General Social Survey, CGSS)数据计算得到2010~2021年的社会信任度^①,按中位数划分为非正式制度较强与较弱两组,通过分样本回归比较公共数据开放的创业效应在地区非正式制度环境上的差异。表5第(3)、(4)列结果表明,与非正式制度环境较弱地区相比,公共数据开放对非正式制度环境较强地区的创业活力提升作用更加明显。一方面,这表明公共数据开放所形成的数字政府治理优化的正式制度创新与非正式制度环境形成互补关系,助推城市创业活力提升;另一方面,非正式制度环境较强地区,企业家可以通过驱动社会网络效应,有效搜寻整合地区资源,赋能企业家创业选择,这进一步体现了公共数据开放影响城市创业活力的重要机制渠道。

表5 异质性分析回归结果

| 变量 | 区域要素禀赋 | | 非正式制度环境 | | 初始数字基础设施政策 | |
|------------------|-------------------|--------------------|------------------|---------------------|-------------------|------------------|
| | 较低 (1) | 较高 (2) | 较弱 (3) | 较强 (4) | 较低 (5) | 较高 (6) |
| <i>open</i> | -0.050 (0.060) | 0.165** (0.068) | 0.138 (0.090) | 0.266*** (0.072) | 0.142* (0.084) | 0.095 (0.094) |
| 控制变量 | 是 | 是 | 是 | 是 | 是 | 是 |
| 事前变量 | 是 | 是 | 是 | 是 | 是 | 是 |
| 固定效应 | 是 | 是 | 是 | 是 | 是 | 是 |
| 样本量 | 450 | 3893 | 1709 | 1725 | 2573 | 1770 |
| R ² 值 | 0.930 | 0.808 | 0.855 | 0.801 | 0.764 | 0.842 |

注:非正式制度环境部分采用2010~2021年的数据。其余同表1。

3. 初始数字基础设施政策

在数字强国战略下,地方政府积极出台了多项数字经济发展政策,旨在加强数字技术基础设施建设,强化数据资源要素供给,以此推动产业数字化与数字产业化转型。完善的数字基础设施为企业进入与公共数据平台上线提供支持和保障,这与公共数据开放促进数字经济发展环境的机制类似。基于省级政府工作报告,收集与数字经济相关词频,将其作为数字基础设施政策支持的代理变量^②,且根据2009~2011年三年的平均值,将样本划分为数字基础设施支持较高与数字基础设施支持较低的地区。表5第(5)、(6)列显示,相比于数字基础设施政策支持较高地区,公共数据开放对平台上线前的数字基础设施政策支持较低地区的创业促进作用更加明显。公共数据开放是建设数字强国的重要战略,地方政府免费向市场主体公开可决策性数据资源的公共品,在降低制度性交易成本与优化信息搜寻匹配效率的重要机制作用下,将能够弥补数字基础设施政策倾向力度的差异,对市场主体创业行为的促进作用更加明显。

4. 行业类型异质性

异质特征行业对市场环境的敏感度可能存在差异,进一步探讨公共数据开放对异质性行业类

^① 采用CGSS问卷中a33(总的来说,您同不同意在这个社会上,绝大多数人都是可以信任的?)的回复中“非常同意”和“比较同意”的人数占该省回复人数总数的比重衡量。考虑到数据统计年份间隔的情况,将2014年数据用2013年代替,2016年数据用2015年代替,2019年和2020年数据用2018年代替。考虑到数据有限,采用2010~2021年的数据进行回归。

^② 数字经济相关词频包括:数字经济、智能经济、信息经济、知识经济、智慧经济、数字化信息、现代信息网络、信息和通信技术、ICT、通信基础设施、互联网、云计算、区块链、物联网、数字化、数字乡村、数字产业、电子商务、5G、数字基础设施、人工智能、电商、大数据、数据化、产业数字化、数字产业化、数据资产化、智慧城市、云服务、云技术、云端、电子政务、移动支付、线上、信息产业、软件、信息基础设施、信息技术、数字生活。

型创业活力的影响,不仅有助于深入分析公共数据开放对高质量创业活力的影响,也能加强对本研究的机制逻辑的理解。按新注册企业类型划分为制造业与服务业,接着进一步将服务业划分为生产性服务业与非生产性服务业^①。表6第(1)、(2)列表明公共数据开放对服务业的创业活力的促进作用更加明显;第(3)、(4)列表明公共数据开放对生产性服务业的创业活力的促进作用更加明显。

表6 行业类型异质性分析回归结果

| 变量 | 制造业 (1) | 服务业 (2) | 生产性服务业 (3) | 非生产性服务业 (4) |
|------------------|------------------|---------------------|---------------------|---------------------|
| <i>open</i> | 0.004 (0.003) | 0.222*** (0.066) | 0.190*** (0.059) | 0.034*** (0.010) |
| 控制变量 | 是 | 是 | 是 | 是 |
| 事前变量 | 是 | 是 | 是 | 是 |
| 固定效应 | 是 | 是 | 是 | 是 |
| 样本量 | 4343 | 4343 | 4343 | 4343 |
| R ² 值 | 0.869 | 0.780 | 0.758 | 0.857 |

注:同表1。

上述结果表明,公共数据开放更多地促进服务业,尤其是生产性服务的创业,但是对实体经济制造业的影响并未体现。一方面,实体经济制造业对区位条件要求以及地区生产要素资源禀赋要求高,其新企业进入的门槛相对较高,而服务业企业进入市场的门槛相对较低。因此,在公共数据开放引致的制度性交易成本降低与信息搜寻匹配效率优化的作用机制下更易于服务业创业。另一方面,数据是一种新型生产要素,生产性服务业企业更加善于从数据中挖掘数据信息资源,将其作为生产要素融入生产体系,能够更加发挥出公共数据要素的创造价值;而实体经济制造业的生产要素往往以土地、资本等为主,数据要素尚未能够有效地融入制造业生产体系中,因此公共数据要素共享对生产性服务业创业的促进作用更加明显,公共数据开放还进一步显著促进高质量创业活力。同时,在构建良好市场环境的数据要素市场机制的进程中,政府部门应加强对实体经济制造企业创业的关注力度,充分发挥公共数据要素在制造业行业中的创业激发价值效应。

七、结论与政策启示

公共数据开放是推进数字政府建设,创新公众与政府互动协同治理模式的一项战略举措,也是推进市场主体数据要素价值转化,激发国内经济活力的必然选择。立足于公共数据开放兼具政府治理逻辑与市场效益逻辑的功能定位,推进经济高质量发展发挥赋能创造效应的基本事实。本文将各城市公共数据开放平台上线视为准自然实验,利用2007~2021年296个地级市及以上城市的非平衡面板数据与企业工商注册信息数据,构建多期双重差分模型,系统检验了公共数据开放对城市创业活力的影响及其作用机制。主要结论如下:第一,公共数据开放显著促进城市创业活力提升。第二,公共数据开放通过制度性交易成本降低的政府治理逻辑和信息搜寻匹配效率优化的市场效益逻辑的途径来促进城市创业活力提升。第三,异质性分析表明,区域要素集聚能力与非正式制度

^① 生产性服务业包括交通运输、仓储和邮政,信息传输、软件和信息技术服务业,金融业,租赁和商务服务业,科学研究和技术服务业,教育业,批发和零售等;非生产性服务业包括电力、热力、燃气及水生产和供应业,住宿和餐饮业,房地产业,水利、环境和公共设施管理业,居民服务、修理和其他服务业,卫生和社会工作文化、体育和娱乐业,公共管理、社会保障和社会企业等。

环境强化了公共数据开放对城市创业活力的促进作用,而且公共数据开放弥补了初始地区数字基础设施政策支持差距,呈现出包容性特征。第四,公共数据开放尚未有效地赋能制造业企业创业,而且对生产性服务业的促进作用更加明显,表明公共数据开放还进一步促进高质量创业。

基于本文的研究结论,提出如下政策建议。

第一,深化推进公共数据开放战略改革,加快培育数据要素市场。数据要素已经深刻地融入生产、分配、流通与消费各环节,是国家基础性战略资源,中国政府部门拥有着极其海量的数据资源。本文研究表明公共数据要素共享有助于提升市场微观主体创业活力,进而激发经济活力。有鉴于此,各级政府首先应积极探讨公共数据开放共享体制机制的建设,深化推进公共数据开放战略改革,依法主动开放社会发展迫切需要的高价值政府数据与重大公共利益的企业数据,强化高质量数据要素供给与畅通数据要素市场化流通机制,加快培育数据要素市场,以此激发经济增长新活力,推动经济高质量发展;其次应依托于国家层面的公共数据分类分级管理标准与开放原则,完善各级公共数据开放平台体系建设,加强数据分类分级安全管理,明确数据范围、细化业务分类、业务属性分类与确定分类规制,完善数据安全保护体系建设。

第二,充分有效地发挥“有为政府”与“有效市场”的有机结合,构建中国特色的数据要素市场化建设。本文研究表明,公共数据开放通过制度性交易成本降低的政府治理逻辑和信息搜寻匹配效率优化的市场效益逻辑的途径来促进城市创业活力提升。有鉴于此,政府部门首先应充分重视公共数据开放对于构建数字政府建设的作用,将“顶层设计”与“摸着石头过河”深化融合,以公共数据开放为底层逻辑支撑,构建各政府部门协同履行公共职责的数字政府治理体系机制,以此赋能国家治理体系与治理能力现代化;其次坚持以市场为导向,完善数据要素市场运营机制,切实维护微观市场主体的合法权益,疏通各市场微观主体数据采用收益权益不明确、数据垄断、数据泄露风险等问题,构建统一公平、竞争有序、成熟完备的数据要素市场体系,充分释放公共数据开放的价值创造;最后在数据要素市场分配机制中充分发挥政府作为以及市场监督作用,加强公共数据治理,建立健全国家公共数据的管理体系与监管规制。

第三,政府部门应根据区域经济发展以及行业的差异,拓展具有规范化、多样化的公共数据开放共享机制,以此发挥公共数据开放激发经济活力的重要作用。本文异质性分析发现,公共数据开放对初始数字基础设施政策支持较低地区的创业促进效应更加明显,而要素集聚能力与非正式制度环境则强化了公共数据开放的创业促进效应。结合前述降低制度性交易成本与优化信息搜寻匹配效率的作用机制,政府部门应依托于公共数据开放的制度创新,营造公平市场竞争环境,抑制政府对微观市场主体的无效干预;同时,加强对不同地区间的统筹协调,因地制宜地制定与落实公共数据开放共享机制,发挥公共数据开放的包容性特征,推动公共数据要素共享的创业效应的协调发展。更为重要的是,公共数据开放显著促进了高质量创业,但尚未在实体经济制造业企业创业中发挥价值创造效应。因此,政府部门在深化推进公共数据开放战略改革与建设数据要素市场化的过程中,应强化针对实体经济制造业行业创业的政策靶向设计,全面激发公共数据开放的价值创造效应,在促进高质量创业的同时更好地赋能实体经济高质量发展。

参 考 文 献

- [1]柏培文,张云.数字经济、人口红利下降与中低技能劳动者权益[J].经济研究,2021,(5):91~108.
- [2]毕青苗,陈希路,徐现祥,李书娟.行政审批改革与企业进入[J].经济研究,2018,(2):140~155.
- [3]毕青苗,徐现祥,杨海生.商事制度改革与就业增长[J].数量经济技术经济研究,2024,(1):130~150.

- [4]庇古.福利经济学[M].金镛译,华夏出版社,2007.
- [5]蔡继明,曹越洋,刘乐易.论数据要素按贡献参与分配的价值基础——基于广义价值论的视角[J].数量经济技术经济研究,2023,(8):5~24.
- [6]蔡继明,刘媛,高宏,陈臣.数据要素参与价值创造的途径——基于广义价值论的一般均衡分析[J].管理世界,2022,(7):108~121.
- [7]陈刚.管制与创业——来自中国的微观证据[J].管理世界,2015,(5):89~99+187~188.
- [8]陈晓红,李杨扬,宋丽洁,汪阳结.数字经济理论体系与研究展望[J].管理世界,2022,(2):208~224+13~16.
- [9]杜运周,刘秋辰,程建青.什么样的营商环境生态产生城市高创业活跃度?——基于制度组态的分析[J].管理世界,2020,(9):141~155.
- [10]段盛华,于凤霞,关乐宁.数据时代的政府治理创新——基于数据开放共享的视角[J].电子政务,2020,(9):74~83.
- [11]方锦程,刘颖,高昊宇,董纪昌,吕本富.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.
- [12]方颖,赵扬.寻找制度的工具变量:估计产权保护对中国经济增长的贡献[J].经济研究,2011,(5):138~148.
- [13]何雨可,牛耕,逯建,赵国昌.数字治理与城市创业活力——来自“信息惠民国家试点”政策的证据[J].数量经济技术经济研究,2024,(1):47~66.
- [14]胡峰,王秉,张思芊.从边界分野到跨界共轭:政府数据协同治理交互困境扫描与纾困路径探赜[J].电子政务,2023,(4):93~105.
- [15]黄勃,李海彤,刘俊岐,雷敬华.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.
- [16]李宏彬,李杏,姚先国,张海峰,张俊森.企业家的创业与创新精神对中国经济增长的影响[J].经济研究,2009,(10):99~108.
- [17]李唐,李青,陈楚霞.数据管理能力对企业生产率的影响效应——来自中国企业—劳动力匹配调查的新发现[J].中国工业经济,2020,(6):174~192.
- [18]刘冲,沙学康,张妍.交错双重差分:处理效应异质性与估计方法选择[J].数量经济技术经济研究,2022,(9):177~204.
- [19]刘大勇,徐晓轩,李妍,胡秋阳,段文斌.降低制度性交易成本如何影响城市间技术要素流转——基于政府服务与专利技术市场流动路径的分析[J].财经研究,2023,(11):110~124.
- [20]马述忠,吴鹏,房超.东道国数据保护是否会抑制中国电商跨境并购[J].中国工业经济,2023,(2):93~111.
- [21]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.
- [22]王永进,冯笑.行政审批制度改革与企业创新[J].中国工业经济,2018,(2):24~42.
- [23]夏杰长,刘诚.契约精神、商事改革与创新水平[J].管理世界,2020,(6):26~36+48+242.
- [24]夏杰长,刘诚.行政审批改革、交易费用与中国经济增长[J].管理世界,2017,(4):47~59.
- [25]袁礼,龚钰涵.专利质押融资对创业活跃度的影响[J].数量经济技术经济研究,2023,(11):202~224.
- [26]张涛,李均超.网络基础设施、包容性绿色增长与地区差距——基于双重机器学习的因果推断[J].数量经济技术经济研究,2023,(4):113~135.
- [27]张勋,万广华,张佳佳,何宗樾.数字经济、普惠金融与包容性增长[J].经济研究,2019,(8):71~86.
- [28]张子尧,黄炜.事件研究法的实现、问题和拓展[J].数量经济技术经济研究,2023,(9):71~92.
- [29]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):65~76.
- [30]赵云辉,张哲,冯泰文,陶克涛.大数据发展、制度环境与政府治理效率[J].管理世界,2019,(11):119~132.
- [31]郑磊.开放不等于公开、共享和交易:政府数据开放与相近概念的界定与辨析[J].南京社会科学,2018,(9):83~91.

- [32] 朱光顺, 张莉, 徐现祥. 行政审批改革与经济发展质量[J]. 经济学(季刊), 2020, (3): 1059~1080.
- [33] 朱峥. 政府数据开放的权利基础及其制度构建[J]. 电子政务, 2020, (10): 117~128.
- [34] Acemoglu D., Johnson S., Robinson J. A., 2005, *Institutions as a Fundamental Cause of Long-Run Growth* [J], *Handbook of Economic Growth*, 1, 385~472.
- [35] Arrow K.J., 1962, *The Economic Implications of Learning by Doing* [J], *Review of Economic Studies*, 29 (3), 155~173.
- [36] Augier M., Teece D. J., 2009, *Dynamic Capabilities and the Role of Managers in Business Strategy and Economic Performance* [J], *Organization Science*, 20 (2), 410~421.
- [37] Balci G., 2021, *Digitalization in Container Shipping: Do Perception and Satisfaction Regarding Digital Products in a Non-Technology Industry Affect Overall Customer Loyalty?* [J], *Technological Forecasting and Social Change*, 172, 121016.
- [38] Baumol W. J., 1990, *Entrepreneurship: Productive, Unproductive, and Destructive* [J], *Journal of Political Economy*, 98 (5), 893~921.
- [39] Bodory H., Martin H., Lukáš L., 2022, *Evaluating (Weighted) Dynamic Treatment Effects by Double Machine Learning* [J], *Econometrics Journal*, 25 (3), 628~648.
- [40] Borusyak K., Jaravel X., Spiess J., 2022, *Revisiting Event Study Designs: Robust and Efficient Estimation* [R], CEPR Discussion Paper, No.17247.
- [41] Cai H., Fang H., Xu L. C., 2011, *Eat, Drink, Firms, Government: An Investigation of Corruption from the Entertainment and Travel Costs of Chinese Firms* [J], *Journal of Law and Economics*, 54 (1), 55~78.
- [42] Cengiz D., Dube A., Lindner A., Zipperer B., 2019, *The Effect of Minimum Wages on Low-Wage Jobs* [J], *Quarterly Journal of Economics*, 134 (3), 1405~1454.
- [43] De Chaisemartin C., D'Haultfoeuille X., 2020, *Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects* [J], *American Economic Review*, 110 (9), 2964~2996.
- [44] De Soto H., 1989, *The Other Path: The Invisible Revolution in the Third World* [M], New York: Harper & Row.
- [45] Edmonds E. V., Pavcnik N., Topalova P., 2010, *Trade Adjustment and Human Capital Investments: Evidence from Indian Tariff Reform* [J], *American Economic Journal: Applied Economics*, 2 (4), 42~75.
- [46] Farboodi M., Veldkamp L., 2021, *A Model of the Data Economy* [R], National Bureau of Economic Research, No. 28427.
- [47] Donker F. W., Loenen B. V., 2017, *How to Assess the Success of the Open Data Ecosystem?* [J], *International Journal of Digital Earth*, 10 (3), 284~306.
- [48] Ghasemaghaei M., Calic G., 2019, *Does Dig Data Enhance Firm Innovation Competency? The Mediating Role of Data-Driven Insights* [J], *Journal of Business Research*, 104, 69~84.
- [49] Ghasemaghaei M., Calic G., 2020, *Assessing the Impact of Dig Data on Firm Innovation Performance: Big Data is not Always Better Data* [J], *Journal of Business Research*, 108, 147~162.
- [50] Goodman-Bacon A., 2021, *Difference-in-Differences with Variation in Treatment Timing* [J], *Journal of Econometrics*, 225 (2), 254~277.
- [51] Jetzek T., Avital M., Bjorn-Aandersen N., 2019, *The Sustainable Value of Open Government Data* [J], *Journal of the Association for Information Systems*, 20 (6), 702~734.
- [52] Kritikos A. S., 2015, *Entrepreneurship and Economic Growth* [A], in Wright J. D., ed., *International Encyclopedia of the Social and Behavioral Sciences (Second Edition)* [C], Oxford: Elsevier, 675~680.
- [53] Krueger J. N. F., Brazeal D. V., 1994, *Entrepreneurial Potential and Potential Entrepreneurs* [J], *Entrepreneurship Theory and Practice*, 18 (3), 91~104.
- [54] Lim D. S. K., Morse E. A., Mitchell R. K., Seawright K. K., 2010, *Institutional Environment and Entrepreneurial Cognitions: A Comparative Business Systems Perspective* [J], *Entrepreneurship Theory and Practice*, 34 (3), 491~516.
- [55] Nagaraj A., 2022, *The Private Impact of Public Data: Landsat Satellite Maps Increased Gold Discoveries and*

Encouraged Entry [J], *Management Science*, 68 (1), 564~582.

[56] Park S., Gil-Garcia J. R., 2022, *Open Data Innovation: Visualizations and Process Redesign as a Way to Bridge the Transparency-Accountability Gap* [J], *Government Information Quarterly*, 39 (1), 101456.

[57] Romer P. M., 1986, *Increasing Returns and Long-Run Growth* [J], *Journal of Political Economy*, 94 (5), 1002~1037.

[58] Ross P. K., Blumenstein M., 2015, *Cloud Computing as a Facilitator of SME Entrepreneurship* [J], *Technology Analysis & Strategic Management*, 27 (1), 87~101.

[59] Schumpeter J. A., 1934, *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest, and the Business Cycle* [M], Cambridge: Harvard University Press.

[60] Shane S., Venkataraman S., 2003, *Guest Editors' Introduction to the Special Issue on Technology Entrepreneurship* [J], *Research Policy*, 32 (2), 181~184.

[61] Shleifer A., Vishny R. W., 1993, *Corruption* [J], *Quarterly Journal of Economics*, 108 (3), 599~617.

[62] Sigfusson T., Chetty S., 2013, *Building International Entrepreneurial Virtual Networks in Cyberspace* [J], *Journal of World Business*, 48 (2), 260~270.

[63] Teece D. J., Pisano G., Shuen A., 1997, *Dynamic Capabilities and Strategic Management* [J], *Strategic Management Journal*, 18 (7), 509~533.

[64] Wang H. J., Lo J., 2019, *Factors Influencing the Adoption of Open Government Data at the Firm Level* [J], *IEEE Transactions on Engineering Management*, 67 (3), 670~682.

[65] Zhang C., 2020, *Clans, Entrepreneurship, and Development of the Private Sector in China* [J], *Journal of Comparative Economics*, 48 (1), 100~123.

[66] Zhao Y., Liang Y., Yao C., Han X., 2022, *Key Factors and Generation Mechanisms of Open Government Data Performance: A Mixed Methods Study in the Case of China* [J], *Government Information Quarterly*, 39 (4), 101717.

[67] Zuiderwijk A., Janssen M., 2014, *Open Data Policies, their Implementation and Impact: A Framework for Comparison* [J], *Government Information Quarterly*, 31(1), 17~29.

Data Factor Sharing and Urban Entrepreneurial Vitality : Empirical Evidence from Public Data Openness

CAI Yunkun¹ ZHOU Jingkui^{1,2} YUAN Wangping¹

(1.School of Economics, Nankai University;

2.The Laboratory for Economic Behaviors and Policy Simulation, Nankai University)

Summary: Through data mining, desensitisation, integration and analysis, data resources can be accurately matched with supply and demand, and the flow of the value chain can be innovated, thus helping to cultivate new productivity and promote high-quality economic development. China has extremely massive data resources, especially public data resources covering government governance-related and enterprises with significant public interests generated by government departments in the performance of their public duties, and it is an important initiative to promote the construction of digital government and digital economy strategy to public data openness. Currently, China's economic vitality is an inherent requirement for promoting high-quality economic development. The enterprises are the main body of the market economy, which is the 'ballast' to promote the development of the national economy, and to adequately stimulate and release the entrepreneurial energy of the whole society is deeply rooted in

the structural change of economic and social development. Then, whether the public data openness strategy can release the value creation effect, promote social entrepreneurship and enhance economic vitality is a major issue of common concern in the practical and theoretical circles.

Accordingly, based on the panel data and business enterprise registration data of 296 prefecture-level cities from 2007 to 2021, using the public data platform on-line in each city as a quasi-natural experiment, the paper builds a multi-period double-difference model to attempt to examine the impact of public data openness on urban entrepreneurial vitality and its channels of action based on the symbiotic logic perspective of government governance and market efficiency. The research finds that, first, public data openness significantly promotes urban entrepreneurial vitality. Second, the mechanism test indicates that the reduction of institutional transaction costs and the optimisation of information search efficiency are the economic mechanisms through which public data openness promotes the enhancement of urban entrepreneurial vitality. In particular, on the one hand, public data openness strengthens the public supervision and accountability mechanism, improves the efficiency of government public services, reduces institutional transaction costs, and then stimulates the enhancement of urban entrepreneurial vitality; on the other hand, public data openness strengthens the entrepreneurs' customer anchoring, market demand orientation, and entrepreneurial opportunity perception by optimising the efficiency of information search and matching, and then creates the opportunities and conditions for entrepreneurs, and then positively impacts on the enhancement of urban entrepreneurial vitality. Third, the heterogeneity analysis indicates that regional factor agglomeration capacity and informal institutional environment strengthen the role of public data openness in promoting urban entrepreneurial vitality, and that public data openness bridges the gap of policy support for digital infrastructure in the initial region, and promotes entrepreneurship in the productive service industry more significantly, which suggests that public data openness further promotes high-quality entrepreneurship.

The contribution of this research is that, firstly, focusing on the functional positioning of public data openness to promote the construction of digital government and digital economy strategy, we have empirically analysed the impact of public data openness on urban entrepreneurial vitality, which enriches the results of the research related to public data openness. Secondly, we further explore the internal mechanism of public data openness affecting urban entrepreneurial vitality, focusing on the impact on entrepreneurship of the government governance logic of systemic transaction cost reduction and the market efficiency logic of information search and matching efficiency optimisation. Meanwhile, a new theoretical logic is provided to further understand the role of the organic combination of 'active government' and 'effective market' in the construction of data factor marketisation process in the era of big data. Third, this study further explores the effects of regional factor endowments, informal institutional environments, differences in initial digital infrastructure policy support and heterogeneity of entrepreneurial industries, which are important policy values for understanding the value creation effects of public data openness and improving the institutional design to enhance the entrepreneurial vitality of cities.

Keywords: Public Data Openness; Urban Entrepreneurial Vitality; Institutional Transaction Costs; Information Search Matching Efficiency

JEL Classification: H11; O31; R58

(责任编辑:焦云霞;校对:石银凤)