引用格式：卞元超，苏佳豪. 公共数据开放能缓解新进入者劣势吗？——来自公共数据平台的准自然实验证据[J]. 技术经济，2026，45(7)：83-99.

Bian Yuanchao，Su Jiahao. Can the opening of public data alleviate the liability of newness? Empirical evidence from a quasi-natural experiment on a public data platform[J]. Journal of Technology Economics，2026，45(7)：83-99.

产业技术经济

公共数据开放能缓解新进入者劣势吗？

——来自公共数据平台的准自然实验证据

卞元超，苏佳豪

（南京师范大学商学院，南京 210023）

摘 要：数字经济时代的新进入者劣势主要表现为数据敏感型创业活动因缺乏数据而导致的无法进入的状态。立足公共数据的数据要素和公共服务双重属性，考察2010—2019年这一样本区间下公共数据开放对地区数据敏感型创业的影响，并利用地方政府上线公共数据平台的准自然实验，实证分析其中的作用效果。研究结果显示，公共数据开放能够显著促进地区的数据敏感型创业，弥补数字经济时代的新进入者劣势。机制分析表明，公共数据开放能够释放地区的市场需求潜能，为数据敏感型创业者提供更多优质的生产要素，优化地区的营商环境，从而激发数据敏感型创业者的市场进入，形成市场赋能效应、要素赋能效应和制度赋能效应。进一步分析发现，随着公共数据开放的质量和地区信息获取程度的提升，公共数据开放对地区数据敏感型创业的影响效应也随之提升，而且公共数据开放也有助于带动地区整体的创业活跃度；公共数据开放对市场分割程度较低的城市数据敏感型创业水平的促进作用更强。研究结论可能为破解数字经济时代的新进入者劣势、营造有助于创业活动的数字化环境提供有益启示。

关键词：公共数据开放；数据敏感型创业；新进入者劣势；市场赋能；要素赋能；制度赋能

中图分类号：F062.6 **文献标志码**：A **文章编号**：1002-980X(2026)07-0083-17

DOI：10.12404/j.issn.1002-980X.J26030415

一、引言

创业在本质上是创业者将各种创业资源整合并进行生产，将其转化为新产品、新市场与新技术的“创造性破坏”的过程[1]。其中，创业资源的获取便成为决定创业活动成功与否，避免新创企业落入新进入者劣势的关键。新进入者劣势指的是新创企业由于缺乏顾客关系、投资者关系及外部联系等外部创业资源而陷入短期无效率等问题[2]。因此，新进入者，即创业者所新创办的企业往往会由于创业资源的获取不足而无法生存，从而退出市场。进入数字经济时代，除了资本、劳动力、土地、技术外，数据要素成为一种基础性和战略性资源，从而也催生了对数据要素的规模和结构比较敏感的创业活动。相较于传统的生产要素，数据要素属于数字化的信息，不仅具有非稀缺性、强流动性、高融合性和非排他性[3]，还具有内生生成性，即只要有经济活动就会产生数据；经济活动越多，数据产生量就越多[4]，这决定了数据要素不是无处不在、无时不在、人人皆有，其往往会被少数大的平台垄断[5-7]。对于数据敏感型创业活动而言，新进入者劣势就不仅仅体现在其拥有的传统资源和能力的欠缺，还体现为数据要素及其获取能力的匮乏，这可能会引发市场失灵，导致

收稿日期：2026-03-04

基金项目：国家自然科学基金面上项目“数据要素治理驱动企业技术边界拓展的机理及效应研究：数实技术融合的视角”(72574104)；国家自然科学基金青年项目“国内市场整合驱动企业技术升级的机理与效应研究：地区技术多样化的视角”(72203107)；江苏省社会科学基金青年项目“创新策源地建设赋能江苏未来产业发展的路径研究”(25ZHC010)

作者简介：卞元超(1991—)，博士，南京师范大学商学院副教授，研究方向：区域创新与技术管理；苏佳豪(2000—)，南京师范大学商学院硕士研究生，研究方向：区域政策与区域经济。

创业活动低于社会最优水平。

近年来,为了进一步激发创业活力,营造有助于创业活动的数字化环境,弥补数据要素缺失所引发的新进入者劣势,各级政府在强化数据要素市场体系建设过程中,不断致力于公共数据开放共享,以此充分释放各类数据要素的经济价值。公共数据是政府公共管理和服务机构在依法履行公共管理职责或提供公共服务过程中产生、采集并处理的数据,包括但不限于本地区在经济建设、公共安全等领域的信息,以及各级党政机关在执行相关政策时所记录的并以电子化形式保存的信息。2015年出台的《促进大数据发展行动纲要》指出要稳步推动公共数据资源开放,2016年发布的《大数据产业发展规划(2016—2020年)》也指出在全国建设若干国家大数据综合试验区,在大数据制度创新、公共数据开放共享等方面开展系统性探索实验。国家"十四五"规划在"提高数字政府建设水平"一章中专门提出要加强公共数据开放共享。《中共中央关于制定国民经济和社会发展第十五个五年规划的建议》也提出要健全数据要素基础制度,建设开放共享安全的全国一体化数据市场,深化数据资源开发利用。目前,推进公共数据开放已经成为加快数据要素市场体系建设和数字中国建设的重要内容,对于推进国家治理体系和治理能力现代化亦具有重要意义。

公共数据具有数据要素和公共服务的双重属性。一方面,公共数据所记录的政府部门供给公共服务的信息在很大程度上能够反映本地区居民或企业对公共服务的使用状况与需求偏好,是一个巨大的"市场信号库"。推进公共数据的开放共享,能够降低数据敏感型创业者获取市场信息的壁垒与成本[8],缓解信息不对称[9],帮助它们挖掘更多潜在的商业需求和经济价值增长点,形成一种市场赋能效应。另一方面,公共数据能够促进资本要素结构升级,从数字化、智能化方面优化资本结构,进而为数据敏感型创业者提供更多优质的资本要素,帮助创业者实现更好的资源整合与"创造性破坏",形成一种要素赋能效应。不仅如此,推进公共数据开放既能够更加有效地向社会传达政府的施政信息,强化社会监督,提高政府透明度,也可以减少企业的机会主义行为,规范市场秩序,从而优化地区的营商环境,降低数据敏感型创业活动的制度性交易成本,形成一种制度赋能效应。其中,值得注意的是,公共数据作为政府向社会开放的公用数据,其本身就是一种公共资源。政府向社会输送海量且透明的数据能够带来大量的商业和社会价值[10],实现了公共价值的形成与释放,开辟了公共价值形成的新路径。同时,在数字经济时代下,区别于以往政府治理的单一职能,公共数据开放也是政府实现数字化治理的重要体现。政府通过公共数据开放,实现了企业、政府、消费者等多元主体的协同治理,形成了数字经济时代下政府在治理中的"赋能者"角色,通过公共数据开放赋能各类主体共创社会价值。

既有文献已经发现了数字经济时代创业活动面临的新特征和新问题[11-13],也关注到了政府所推进的公共数据开放对企业创新[14]、企业机会主义行为[15]、企业生产方式和效率[16-17],以及对宏观经济发展所产生的影响[18-19]。与本文相近的文献中,蔡运坤等[13]及郑志强和何佳俐[15]等学者都关注到了公共数据开放对创业活动的影响,验证了公共数据开放在促进新企业进入方面的积极效果。遗憾的是,这些研究并没有从公共数据的双重属性角度探究公共数据开放所形成的市场赋能效应、要素赋能效应和制度赋能效应,尤其是忽视了公共数据开放对数据敏感型创业的影响效应,这可能不利于更加全面地认识公共数据开放与数字经济时代下创业活动之间的关系,以及数字经济时代下政府作为"赋能者"通过公共数据开放实现公共价值的路径。本文立足数字经济时代创业活动的新进入者劣势,基于2010—2019年这一样本研究区间,将公共数据开放的市场赋能效应、要素赋能效应和制度赋能效应纳入一个统一的框架内,分析其对地区数据敏感型创业水平的影响效应,发现公共数据开放能够显著地激发地区的数据敏感型创业水平。

可能的边际贡献有三点:第一,相较于以往研究,本文紧密结合数字经济时代创业者因数据要素缺乏所产生的新进入者劣势这一问题,聚焦数据敏感型创业活动,探究政府主导的公共数据开放如何助力数据敏感型创业者破解新进入者劣势和市场失灵。本文研究不仅有助于更加全面地认识在数字经济时代影响创业活动的关键因素,而且揭示了公共数据作为公共资源释放公共价值的方式,从而拓展公共价值的生成路径,丰富了公共价值理论的研究视角。第二,从公共数据作为数据要素和公共服务的双重属性出发,从市场赋能效应、要素赋能效应和制度赋能效应三个方面深入分析了公共数据开放对地区数据敏感型创业的影响机理,有助于更加全面地理解政府主导的公共数据开放影响创业活动的内在机制。本文为理解数字经济时

代下政府的“赋能者”角色提供了新的视角，延伸了传统公共治理理论中政府的服务职能，丰富了公共治理理论的内涵。第三，相较于以往研究，本文不仅考察高质量的公共数据开放对数据敏感型创业活动的影响，还探究了公共数据开放是否影响了地区整体创业活动水平。这有助于更加全面地认识公共数据开放与地区创业活动之间的关系，为相关政策的科学制定提供针对性的启示。

二、文献综述和机制分析

（一）文献综述

早期关于影响创业活动因素的研究主要集中于企业家才能的视角，即创业者的能力与禀赋是决定创业活动的核心要素，如自有资源的数量和质量、识别市场机会的能力、利用创新的能力及容忍风险的能力等[20-21]，以及创业者个人的特质（如性别、年龄、受教育年限和工作经验等）[22-24]。然而，与在位企业相比，新创企业由于内部资源与能力相对较少，缺乏丰富的市场经验，往往会陷入新进入者劣势而资源枯竭死亡[2,25-26]。因此，对于新创企业而言，外部创业资源获取和创业环境便成为其创业决策的关键因素，此即制度基础观[27]。

制度基础观认为，企业家的供应在不同时期的社会中是相对稳定的，经济发展并不取决于一个社会企业家才能或企业家精神的多少，而是取决于政治、经济、社会、文化等制度与非制度环境对企业家精神的正确引导[27]。从制度环境角度出发，差异化的制度影响了创业者的行动选择，创业者通过偏离和遵守制度来有效地管理不确定性，从而表现为差异化制度中的选择多样性[28-29]。从营商环境角度出发，营商环境之间的差异使得创业者对政府效率、人力资源、金融服务及公共服务等创业资源获取存在差异性，从而影响创业行为[30-31]。其中，从金融服务环境的角度出发，金融包容能够为企业提供多元化的融资渠道，极大程度上缓解了企业的融资约束[32]，激发了企业家的创业精神。此外，基础设施作为公共服务的重要体现，其不仅能够提升当地的环境承载力从而催生更多创业机会[33]，而且能够吸引企业与高技能人才的流入，激发地区的创业活力[34-35]。

近年来，随着数字经济的快速发展，大数据分析、云计算及人工智能等数字技术不断普及，数据要素不断渗透于企业的生产之中[16]。数字经济时代的创业者（数据敏感型创业者）不仅关注传统的生产要素和资源[36]，还对数据要素的获取，以及数字经济发展环境更加关注。其中，公共数据作为数据要素的典型代表，从其公共资源的属性角度出发，公共数据的开放首先就是为社会提供了大量透明、公开、共享的公共信息。公共数据的开放能够降低投资者等经济行为主体获取信息的成本[8,16]、披露企业信息减少企业的机会主义行为[15]，以及缓和银企之间的信息不对称[9]。然而，数字经济时代下公共数据不仅是一种公共资源，更是一种数字技术创新应用[13,19]。

从数字技术创新应用的角度出发，公共数据的开放有助于实现数据要素与传统生产要素的结合[16]。数字经济时代下，数据要素作为高效率生产要素，通过促进生产结构转型升级来匹配技术进步方向，提高企业的生产效率，降低生产成本，从而推动企业的创业创新行为[14]，尤其是数据敏感型创业。同时，金融机构能够依托公共数据开放平台，将公共数据资源与其内部资源相结合，对企业相关数据进行精准把控，能够扩大金融服务的覆盖范围[37]，缓解小微企业的融资约束[38]，优化数据敏感型创业者的金融环境和金融基础设施环境，激发其创业意愿[39]。此外，随着数据要素在传统生产领域的渗透，公共数据资源的供给服务提升，公共服务也呈现出数字化的特征。例如，数字基础设施建设能够有效利用海量的优质公共数据资源，有利于吸引高技能创业人才，增强当地的经济活力[40]。

整体而言，现有文献对于影响创业活动的关键因素进行了比较充分的挖掘和探讨，尤其是以制度基础观为代表的研究强调了外部环境对创业活动的影响。伴随着数字经济的快速发展，越来越多的学者关注到数字经济发展和数字化环境对创业活动的影响，也有文献考察了数字经济发展过程中的公共数据开放所产生的经济效应，这对于本文的研究都具有非常直接的启示意义。遗憾的是，目前少有文献关注到政府主导的公共数据开放共享对地区数据敏感型创业活动的影响，尤其是没有考察公共数据的数据要素属性和公共服务属性在其中的影响，这势必不利于更加深入地识别数字经济时代下影响创业的关键因素，也不利于全

面考察公共数据开放的经济效应。因此,本文从政府主导的公共数据开放这一角度出发,立足公共数据的公共服务属性和数据要素属性,探讨公共数据开放对地区数据敏感型创业的影响效应。

(二)影响机制分析

传统的创业理论认为,影响创业者市场进入的核心要素包括创业资源、创业机会与创业成本[41]。相较于传统的工业社会阶段,数字经济时代下的创业者,不仅关注传统的劳动力、资本、技术等资源及外部的制度环境,还更加关注数据要素的获取成本,以及相应的数字化环境。公共数据具有数据要素和公共服务的双重属性[13,19],富含大量的关于市场需求的信息。政府主导的公共数据开放能够推动公共数据的社会共享,降低数据敏感型创业者对公共数据的获取成本,这既有助于创业者获取更多与市场需求相关的信息,也能够实现数字赋能创业企业的资本结构升级,更可以通过优化营商环境而为创业者提供更加完善的制度环境。也就是说,数字经济时代下,公共数据开放是政府治理职能转变,实现多元主体协同治理,共创社会价值的重要方式。政府作为数字经济时代下的"赋能者"角色,其通过公共数据开放对地区数据敏感型创业的赋能作用主要体现为市场赋能效应、要素赋能效应和制度赋能效应三个方面。

第一,公共数据开放能够帮助数据敏感型创业者获取更多关于市场需求的信息,增加创业机会,从而促进地区数据敏感型创业活动的开展,形成市场赋能效应。公共数据是政府公共管理和服务机构在依法履行公共管理职责或提供公共服务过程中产生、采集并处理的数据,包括但不限于本地区在经济建设、公共安全、教育科技、卫生健康、供水供电、供气供热、环境保护、公共交通、文化休闲等领域的信息。这些数据不仅记录了政府公共管理部门提供公共服务或公共产品的信息,更蕴含了本地区居民或企业使用公共服务的信息,反映了本地区居民或企业对公共服务的需求状况和偏好状况,是一个巨大的"市场信号库"。因此,政府推动公共数据的开放共享,不仅是提高政府透明度的重要举措[15],还可以帮助市场主体挖掘更多潜在的商业需求和经济价值增长点[19]。尤其是公共数据所具备的可复制性强和公共产品属性,推进公共数据开放共享有助于降低数据敏感型创业者获取市场信息的成本,破除创业者面临的市场信息壁垒,减少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[13],从而提升新创企业对于未来市场需求和市场份额的预期,以此带来更多的创业机会。因此,政府推动公共服务数据开放共享能够帮助数据敏感型创业者挖掘更多潜在的市场需求,激发地区的市场潜能,激励创业者的市场进入行为,此即市场赋能效应。

第二,公共数据开放所形成的海量数据能够促进资本要素结构升级,激励数据敏感型创业活动的开展,形成要素赋能效应。创业是企业家重新组合生产要素,开拓新产品、新市场和新技术的"创造性破坏"过程,拥有种类更多、规模更大、结构更好、质量更高的生产要素无疑对创业活动的产生具有重要的促进作用。对于公共数据而言,其不仅是一种信息资源,更是一种生产要素[16]。企业能够通过加工、筛选和处理等方式,将其转化为能够有效融入生产活动的数据要素。数据要素依托线上公共平台实现了大规模的共享并不断渗透于生产领域,这有助于促进数据要素的流动,提升数据资源的配置效率[42]。这一过程中,数据要素通过数字化、智能化设备等资本载体,将数据驱动决策、实时优化、路径规划等功能嵌入传统机械设备,从而实现资本形态的数智化升级。因此,数据要素与传统资本要素的融合有助于数据要素改善传统资本要素结构[16],促进资本结构的升级,推动数据要素赋能。在数字经济时代,数据正在成为一种新型的生产要素,也成为激励和支撑创业活动的重要因素。数据要素不断渗入传统生产领域,有助于促进资本要素结构的升级,为数据敏感型创业活动提供更多的优质资本要素。数据要素对资本结构升级的影响不是局部性的"技术取代资本",而是系统性地改变了企业资本配置的条件、环境与能力。更多数据要素与传统资本要素的结合有助于企业采用数字化的模式进行生产,也通过使用数字化技术与智能化设备能够更好地把控产品信息和生产风险,优化内部管理结构,降低企业管理中的冗余,降低不确定性感知度[43],进而减少创业过程中的风险,激发地区的数据敏感型创业。因此,公共数据开放能够降低创业者获取数据要素的成本,为数据敏感型创业者供给更多的数据要素,促进资本结构向数字化、智能化方向升级,形成要素赋能效应。

第三,公共数据开放能够优化地区营商环境,强化社会监督,促进数据敏感型创业活动的开展,产生制度赋能效应。就公共数据的范围来说,其包括了政府供给公共服务过程中所产生的数据和政策制定、政策实施、政策执行过程中的信息资源[44],推进公共数据开放在一定程度上来说也是优化营商环境、强化社会

监督、提高政府透明度的重要环节。一方面，公共数据开放共享有助于向社会公众更好地传达政府政策动向和实施效果，推动政府治理体系公开透明，减少政府对市场的不正当干预，提升政务服务的效率，也有助于推进社会公众共同参与和监督政府的行为，从而进一步优化营商环境；另一方面，公共数据开放披露了大量的与企业相关的数据，包括企业的注册信息、工商行政处罚、客户信用状况、失信被执行人信息等方面的数据，极大程度上减少了部分企业的虚假披露和机会主义行为[15]，能够有效规范市场秩序，优化营商环境。

数字经济时代下，营商环境的优化能够使新创的数据敏感型企业更加容易获取外部资源，在一定程度上克服了新进入者劣势[30]，对于提升地区数据敏感型创业具有重要作用。商业信用环境的优化能够减少企业的寻租行为，形成更加公平的竞争环境，避免企业无谓的社会机会成本的损失，降低创业活动的成本。不仅如此，由于营商环境的优化，新建企业所面临的创新创业方面的成本约束较之以往更小，企业能够更专注于创新行为，提升综合竞争力，而且能够实现内部资源的合理分配[30]，实现企业运营与企业创新行为之间的效益最大化，提升企业的生存能力，从而激发了数据敏感型创业。因此，政府主导的公共数据开放能够进一步优化地区的营商环境，为数据敏感型创业者的市场进入提供制度赋能。

基于以上分析，本文提出假设：

政府主导的公共数据开放能促进地区的数据敏感型创业（H1）；

政府主导的公共数据开放能够从市场赋能、要素赋能和制度赋能三个方面促进地区的数据敏感型创业（H2）。

三、研究设计

（一）模型构建

本文拟通过采用地方政府上线公共数据开放平台作为政策冲击，考察政府主导的公共数据开放对数据敏感型创业的影响。政府进行公共数据开放的初衷在于充分释放公共数据的经济和社会价值，提升行政效率和透明度，强化社会监督，外生于数据敏感型创业活动。因此，可以将其视为一项公共数据开放的准自然实验。本文借助双重差分模型，将上线公共开放平台的城市设置为处理组，将未上线公共开放平台的城市设置为控制组。考虑到各个城市上线公共开放平台的时间不一致，本文建立多时点双重差分模型进行实证检验，具体回归模型如式（1）所示。

$$entrepreneur_{it} = \beta_0 + \beta_1 public_data_{it} + \beta_2 controls + \sigma_i + \sigma_t + \varepsilon_{it} \quad (1)$$

其中：i 为城市；t 为年份；$entrepreneur_{it}$ 为城市 i 在 t 年的数据敏感型创业水平，采用每百人中信息传输、软件和信息技术服务业及科学研究和技术服务业的新创企业数量对其进行表征；$public_data_{it}$ 为公共数据开放平台上线的虚拟变量，若城市 i 在 t 年上线了公共数据开放平台则取 1，否则取 0；$controls_{it}$ 为一系列可能会影响地区创业活力的控制变量；σ_i 为城市固定效应；σ_t 为时间固定效应；ε_{it} 为随机扰动项；β 为待估系数。另外，将标准误聚类到城市层面来排除同一城市内随机扰动项的相关问题。

（二）变量定义

1. 数据敏感型创业

数据敏感型创业活动不仅是指那些以数据要素为核心要素的创业活动，还强调以数据生产作为核心业务。这些创业活动面临着较高的数据合规要求和安全风险。相较于国民经济类别中的其他行业，信息传输、软件和信息技术服务业及科学研究和技术服务业直接从事数据的采集、传输、存储、处理与分析，其产品和服务本身就是数据或信息技术，对数据要素的敏感性不言而喻。同时，该类行业的研发活动、实验测试、技术咨询与推广等业务，高度依赖各类科学数据、实验数据、技术文档和专利数据作为关键研发投入。在当代科研范式下，数据驱动的研发已成为该行业的核心特征。基于此，结合数据可获取性，本文将信息传输、软件和信息技术服务业及科学研究和技术服务业的创业作为数据敏感型创业。采用企查查数据库全国企业信用查询系统获取地级市层面每年在国家市场监督管理总局公示的信息传输、软件和信息技术服务业及科学研究和技术服务业的新建企业数量，形成了各城市相应年份的信息传输、软件和信息技术服务业及科学研究和技术服务业新

创企业数量的面板数据。为了消除区域规模差异,参考白俊红等[45]的方法,采用人口法将城市人口作为标准化基数,以每百人中信息传输、软件和信息技术服务业及科学研究和技术服务业的新创企业数量来衡量数据敏感型创业水平,最终得到了2010—2019年数据敏感型创业水平的面板数据①。

2. 公共数据开放

中国于2012年在上海市和北京市等地率先上线了公共数据开放平台,之后数量持续增长,截至2023年底已达到204个。其中,所有直辖市,以及福建省、广东省等10个省域下辖的所有城市都已上线了公共数据开放平台,形成中国最为密集的"开放数林"。本文将公共数据开放政策视为一项准自然实验,以$public_data_{it}$这一虚拟变量来衡量公共数据的开放。具体的,若城市在t年上线了公共数据开放平台,则设为1;反之,则设为0。若城市上线了公共数据开放平台,则视为处理组;若城市未上线公共数据开放平台,则视为控制组。

3. 控制变量

本文选取控制变量:教育发展水平(*education*)、金融支持水平(*finance*)、经济发展水平(*development*)、科学技术支出(*expenditure*)、地区产业结构(*industry*)、基础设施水平(*road*)、人口密度(*population*)。主要变量的描述性统计见表1。

表1 主要变量描述性统计

变量	观测数	平均值	标准差	最小值	中位数	最大值
entrepreneur	2780	0.0361	0.0691	0.0012	0.0163	1.3249
public_data	2780	0.0633	0.2436	0	0	1.0000
education	2780	1.8674	2.2655	0.0098	1.0283	13.1124
development	2780	10.6660	0.6011	8.5762	10.6412	15.6752
finance	2780	1.1297	0.9818	0.1322	0.8238	16.7426
expenditure	2780	0.0227	0.0291	0	0.0133	0.4237
industry	2780	2.2956	0.1502	1.8206	2.2848	2.8320
road	2780	9.4771	8.1848	0.2535	7.8800	108.3700
population	2780	5.7430	0.9178	1.6194	5.8849	7.8816

四、实证分析

(一)基准回归结果的分析

为了检验公共数据开放对数据敏感型创业的影响效应,对式(1)进行估计,回归结果见表2。基准回归结果表明,不管是在未控制时间固定效应、未加入控制变量、控制时间固定效应和加入控制变量的情况下,公共数据开放(*public_data*)的估计系数均在1%的水平下显著为正。根据(4)列的估计结果,公共数据开放的估计系数显著为0.0388,表明在其他情况不变的条件下,相较于未上线公共数据平台的城市,上线公共数据平台的城市的数据敏感型创业水平平均提高了约3.88%,即地方政府主导的公共数据开放能够显著促进地区的数据敏感型创业。

(二)平行趋势检验

使用双重差分模型的前提是处理组和对照组能够满足平行趋势假设,即在上线公共数据开放平台之前,处理组与控制组在数据敏感性创业水平的变化上不存在显著差异。本文通过构造如式(2)所示的模型进行平行趋势检验。

$$entrepreneur_{it} = \beta_0 + \sum_{t=-4}^{0} Bef_{it}\beta_t + \sum_{t=1}^{5} Aft_{it}\alpha_t + \theta controls_{it} + \sigma_i + \sigma_t + \varepsilon_{it} \quad (2)$$

① 选取2019年作为样本的截止年份而非往后延伸,原因在于本文所采用的地级市层面的变量数据在2020年及以后统计口径(如城市统计年鉴中相关数据)大多发生了改变。例如,人口数量的统计口径由年平均人口数量变为常住人口数量。一方面,本文的被解释变量经过了人口数量的标准化处理,人口数量统计口径的改变会对敏感型创业水平的衡量产生一定的影响,统计口径的改变会导致变量测算产生偏误;另一方面,教育发展水平、经济发展水平(人均GDP),以及人口密度等控制变量的测算也均需经过人口数量的标准化处理或直接以人口数量计算。此外,为进一步排除新冠肺炎疫情对样本的干扰,本文选取2019年作为样本的截止年份。

其中：Bef_{it} 与 Aft_{it} 分别为公共数据开放平台上线前、后第 t 年的虚拟变量；β_t 与 α_t 分别为公共数据开放平台上线前、后第 t 年时的处理组与控制组之间的数据敏感型创业水平的差异。本文将政策相对实施前一期作为基期进行平行趋势检验，检验结果见表 3。其中，*pre_t* 的回归系数与 β_t 相对应，*post_t* 的回归系数与 α_t 相对应，*current* 的回归系数为政策实施当期的回归系数。根据回归结果，在公共数据平台上线之前，处理组与控制组的数据敏感型创业水平不存在显著差异。系数 α_t 在相对上线公共数据开放平台 1~5 年内均显著为正。公共数据开放平台的上线对地区数据敏感型创业活力的作用随着相对开放年份的推移而呈现出上升的趋势。

为了更加直观地观察系数 β_t 与 α_t 的变化趋势，进一步绘制了平行趋势检验图（图 1）。图 1 中的实线描绘了由式（2）回归得到的试点前和试点后估计系数 β_t 和 α_t 的变化趋势，虚线范围内为相应的 95% 置信区间。*pre_t* 分别对应了图 1 中横轴上的 −4~−1 节点，*post_t* 分别对应了图中横轴上的 1~5 节点，*current* 对应了图 1 中横轴上的 0 节点。根据图 1 的输出结果，在公共数据平台上线之前意味着控制组与处理组在政策上线前不存在显著差异。系数 α_t 在相对上线公共数据开放平台 1~5 年内均显著为正，说明公共数据开放平台的上线对数据敏感型创业水平具有显著的促进作用。

但是，既有文献提出，事前趋势检验的低效力和以事前趋势检验结果为条件的分析会导致对事后处理效应估计的偏误，从而无法保证估计的无偏性[46]。基于此，参考 Rambachan 和 Roth[47] 提出的诚实双重差分方法进行平行趋势敏感性检验。该方法主要包括两个步骤：一是构造与平行趋势的最大偏离程度（*Mbar*）；二是构造与上述偏离程度相对应的事后处理效应估计量的置信区间。若在最大偏离程度下，处理后点估计量的

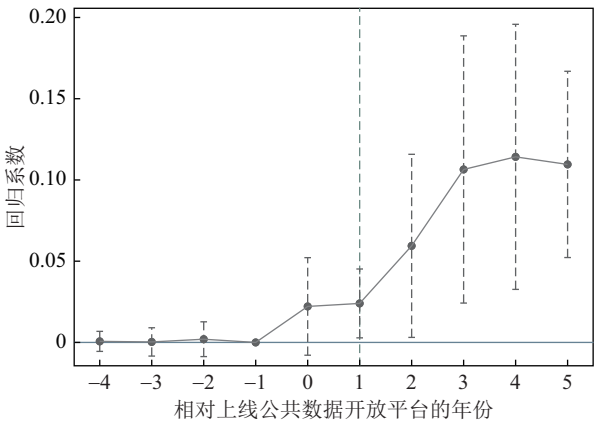


图 1　平行趋势检验结果

表 2　基准回归结果

变量	*entrepreneur*			
	（1）	（2）	（3）	（4）
public_data	0.0889***	0.0518***	0.0636***	0.0388***
	（6.07）	（3.65）	（5.09）	（3.44）
education			−0.0050***	−0.0039**
			（−2.71）	（−2.06）
development			0.0142***	−0.0171**
			（2.81）	（−2.50）
finance			0.0006	0.0033***
			（0.43）	（3.06）
expenditure			0.0466	0.1290**
			（1.24）	（2.25）
industry			0.1117***	−0.1008***
			（5.23）	（−4.22）
road			0.0006	0.0017***
			（1.39）	（2.89）
population			0.2804***	0.2543***
			（3.61）	（3.82）
城市固定效应	是	是	是	是
时间固定效应	否	是	否	是
观测值	2780	2780	2780	2780
调整 R^2	0.5576	0.6341	0.6246	0.6769

注：*、**、*** 分别表示在 10%、5%、1% 水平上显著；括号内为 *t* 值，标准误聚类到城市层面。

表 3　平行趋势检验中各期系数的回归结果

变量	*entrepreneur*
	（1）
pre_4	0.0007
	（0.21）
pre_3	0.0003
	（0.07）
pre_2	0.0020
	（0.37）
current	0.0222
	（1.45）
post_1	0.0240**
	（2.23）
post_2	0.0584**
	（2.08）
post_3	0.1065**
	（2.55）
post_4	0.1142***
	（2.76）
post_5	0.1096***
	（3.76）
控制变量	是
城市固定效应	是
时间固定效应	是
观测值	2780
调整 R^2	0.3970

注：*、**、*** 分别表示在 10%、5%、1% 水平上显著；括号内为 *t* 值，标准误聚类到城市层面。

置信区间不包含 0 值,则表明处理效应对平行趋势的偏离程度具有较好的稳健性。参考 Rambachan 和 Roth[47]的方法,设置 *Mbar*=1 × 标准误,对政策实施后的处理效应进行平行趋势敏感性检验。图 2~图 5 分别反映了政策实施后第 1~4 年的平行趋势敏感性检验结果。根据检验结果,政策实施后第 1~4 年的估计系数的 95%置信区间均不包含 0 值,说明在政策实施后第 1~4 年内给定不同的最大偏离程度下,公共数据开放所引致的数据敏感性创业效应依然稳健,这四年的估计系数具有稳健性。在一定冲击偏离下,公共数据开放依然能够显著提升数据敏感性和创业水平。根据平行趋势敏感性检验的结果,公共数据开放形成的数据敏感型创业效应具有稳健性。

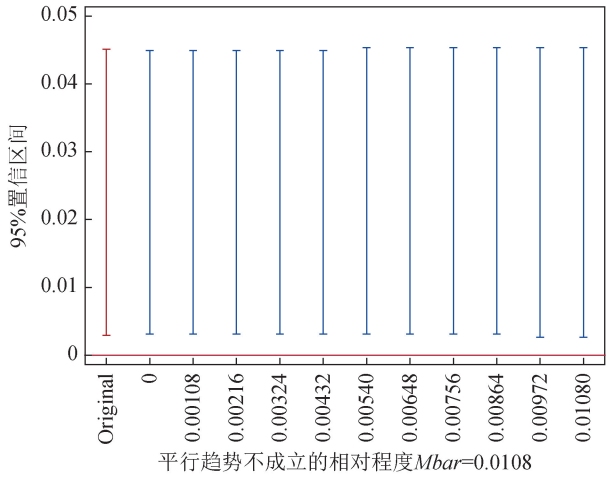


图 2　平行趋势敏感性检验(实施后一年)

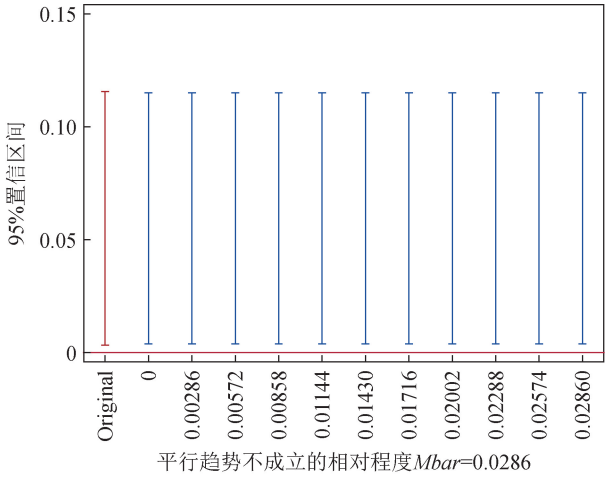


图 3　平行趋势敏感性检验(实施后两年)

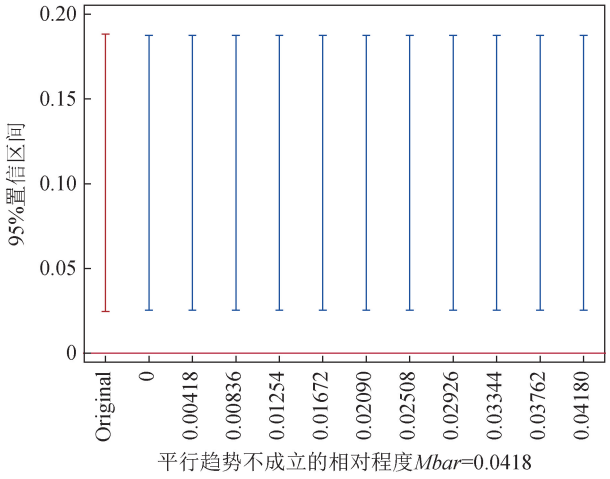


图 4　平行趋势敏感性检验(实施后三年)

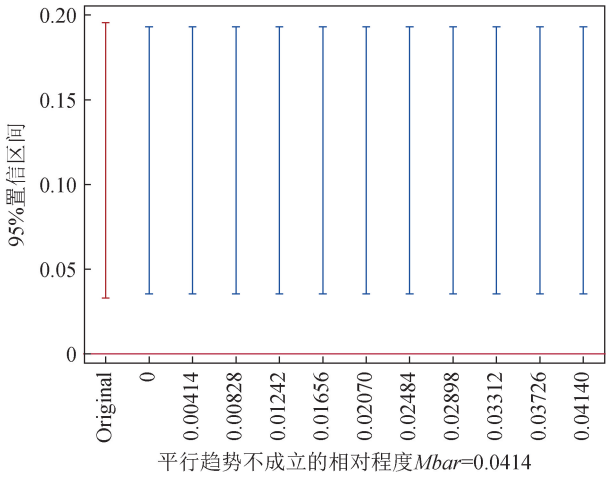


图 5　平行趋势敏感性检验(实施后四年)

(三)稳健性检验

1. 安慰剂检验

本文在 278 个地级市中随机抽取 110 个城市作为虚拟处理组,其余城市作为虚拟控制组进行回归,将此操作重复 500 次后得到 500 组回归系数与其 *P* 值,在图 7 中绘制出 500 组回归系数及其 *P* 值的核密度估计图。从图 7 可以看出,安慰剂检验所得到的回归系数核密度图近似标准正态分布,绝大部分回归系数所对应的 *P* 值大于 10%的显著性水平且安慰剂检验所得到的回归系数均不大于式(1)中的真实回归系数 0.0388。结果表明,随机抽取城市作为伪处理组进行回归并不会对地区数据敏感型创业水平产生显著影响,即随机扰动因素并不会干扰公共数据开放对地区数据敏感型创业水平的促进作用。

2. 替换地区经济活力衡量标准

将地区数据敏感型创业水平的衡量标准替换为未进行城市人口基准化的信息传输、软件和信息技术服务业及科学研究和技术服务业的新创企业数量，即选取数据敏感型新创企业数量（*number*）作为地区创业活力的衡量指标，从表 4 的（1）列可以看出，公共数据开放的回归系数在 1%的水平上显著为正，说明无论选取何种衡量指标，公共数据开放都对地区数据敏感型创业均有显著的促进作用，本文的基准估计结果具有稳健性。

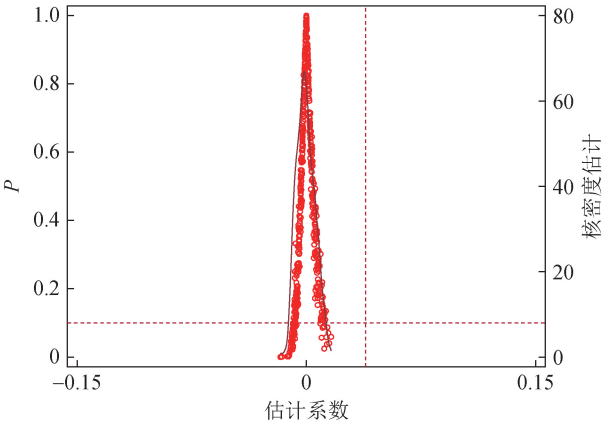


图 6　安慰剂检验结果

3. 排除其他政策冲击的影响

选取智慧城市建设（$policy_1$）与大数据综合试验区（$policy_2$）来进行实证检验。分别将智慧城市建设与大数据综合试验区放入式（1）中进行回归，结果见表 4 的（2）列和（3）列。$policy_1$ 与 $policy_2$ 的回归系数均不显著。本文进一步将这两项政策同时放入式（1）中进行回归。从表 4 的（4）列可以看出，$policy_1$ 与 $policy_2$ 的回归系数也均不显著。结果表明，公共数据开放对地区数据敏感型创业的促进作用不受智慧城市建设与大数据综合试验区的影响冲击，验证了结果的稳健性。

4. 剔除直辖市样本

剔除北京、上海、天津及重庆四个直辖市的数据样本，将式（1）重新回归的结果见表 4 的（5）列。在剔除了直辖市样本后，公共数据开放的估计系数在 1%的水平下显著为正，说明直辖市的特殊性并不会对公共数据开放的影响效应造成太大影响，结果具有稳健性。

5. 替换行业选择

本文基于反事实分析的思路，将行业选择由数据敏感型行业替换为非数据敏感型中的任意行业，将被解释变量替换为每百人中采矿业新创企业数量（*mining*），将式（1）重新回归的结果见表 4 的（6）列。回归结果表明，在替换行业选择后，公共数据开放的估计系数不显著，说明公共数据开放对非数据敏感型行业的新创企业数量不存在显著影响，从反向的角度验证了行业选择的稳健性。

表 4　稳健性检验结果

变量	替换变量	排除政策 1	排除政策 2	排除混合政策	剔除直辖市	替换行业选择
	number	*entrepreneur*	*entrepreneur*	*entrepreneur*	*entrepreneur*	*mining*
	（1）	（2）	（3）	（4）	（5）	（6）
public_data	2858.931***	0.0388***	0.0387***	0.0387***	0.0358***	0.0001
	（3.16）	（3.44）	（3.45）	（3.45）	（3.08）	（1.65）
$policy_1$		−0.0007		−0.0004		
		（−0.20）		（−0.12）		
$policy_2$			0.0064	0.0064		
			（0.80）	（0.79）		
控制变量	是	是	是	是	是	是
城市固定效应	是	是	是	是	是	是
时间固定效应	是	是	是	是	是	是
观测值	2780	2780	2780	2780	2740	2780
调整 R^2	0.7381	0.6768	0.6773	0.6771	0.6311	0.6367

注：*、**、*** 分别表示在 10%、5%、1%水平上显著；括号内为 t 值，标准误聚类到城市层面。

6. 异质性处理效应下的稳健性检验

De Chaisemartin 和 D'Haultfoeuille[48] 发现，使用多时点 DID 模型所得到的回归系数会受异质性处理效应的影响而存在估计偏差。他们将异质性处理效应引入双向固定效应模型，发现双向固定效应模型的真实估计系数 β_{fe} 可表示为所有受处理个体的处理效应的加权之和的期望值如式（3）所示。

$$\beta_{fe} = E\left(\sum_{(g,t):D_{g,t}=1} W_{g,t}\Delta_{g,t}\right) \quad (3)$$

其中：$W_{g,t}$ 为城市 g 在 t 年的处理效应的权重数；$\Delta_{g,t}$ 为城市 g 在 t 年的平均处理效应。负权重的处理效应过多可能会对基准回归的结果产生影响，检验结果见表5。如果异质性处理稳健性指标接近于1，说明原模型所得到的估计系数稳健。结果表明，大部分城市的各年度处理效应为正，个别城市的处理效应为负，说明公共数据开放政策虽然在整体上对地区的数据敏感型创业水平有显著的促进作用，但可能导致区域间的“虹吸效应”或“挤出效应”，表现为在样本期内公共数据开放对个别城市的平均处理效应为负。结果表明，异质性处理稳健性指标为2.0134，说明在整体上异质性处理效应对本文的估计系数不存在较大影响，本文的估计系数具有稳健性。参考 De Chaisemartin 和 D'Haultfoeuille[48]的做法，进一步剔除权重为负的控制组，估计出所有权重为正的控制组的平均处理效应后加权，得到修正后的平均处理效应。修正后的平均处理效应结果见表6。结果发现，修正后的平均处理效应为0.0447，在5%的水平下显著为正，验证了结果的稳健性。

表5　考虑异质性处理效应下的稳健性检验结果

处理变量：*public_data*	处理效应估计值个数	权重之和
正权重	173	1.0025
负权重	3	-0.0025
总计	176	1
双向固定效应系数（β_{fe}）= 0.0388		
与 β_{fe} 及 Δ_TR 相容的最小 $\sigma(\Delta)$ 值 = 0；0.1062		
与 β_{fe} 及符号相反的 Δ_TR 相容的最小 $\sigma(\Delta)$ 值：2.0134		

注：代表所有处理组样本实际的加权平均处理效应。

表6　修正后的平均处理效应结果

联合效应零假设检验：P = 0.04072976						
每个处理单元的平均累积（总）效应						
平均总效应	估计值	标准误	置信区间下限	置信区间上限	观测值	处理转换×时期数
	0.0447437	0.014181	0.0169495	0.0725379	2221	173
处理效应所累积的平均时期数 = 2.0693642						

7. 内生性问题处理

第一，使用 PSM-DID 模型。为了减小选择性偏差所带来的回归估计误差，基于多时点 PSM-DID 模型进行实证检验。具体地，使用卡尺最近邻匹配法将教育水平、地区发展水平等7个控制变量根据倾向得分匹配到截面，剔除匹配未成功的数据，接着根据倾向得分将处理组与其特征相似的控制组相匹配，最终得到匹配后的截面数据，匹配后的回归结果见表7的（1）列。结果表明，公共数据开放（*public_data*）的系数在5%的水平下显著为正，说明匹配后的公共数据开放依然能够促进地区数据敏感型创业水平的提高，验证了基准回归结果的稳健性。

第二，工具变量法。选取各地级市质心到“八纵八横”光缆骨干节点城市的最小距离的自然对数作为公共数据开放的工具变量。本文将其与各地级市政府工作报告中数字经济政策相关总词频占政府工作报告文本总长度的比重相交乘，赋予其时间趋势，最终得到了公共数据开放的工具变量（*IV*）。该工具变量的选择具有两方面的合理性。一方面，中国的宽带网络很大程度上是在1998年建设完成的“八纵八横”光缆通信干线网的基础上不断发展起来的。同时，城市距离这套光缆通信干线网的距离越近，互联网的扩散速度越快[49]。因此城市与“八纵八横”光缆通信干线网的距离会影响后续一系列以互联网为基础的措施，比如上线公共数据开放平台，工具变量满足相关性条件。另一方面，各城市与“八纵八横”光缆通信干线网之间的地理距离是在历史上客观存在的，并且随着现在各地信息基础设施的不断完善，互联网普及率也在不断提高。因此，这一地理距离对于信息获取壁垒的影响越来越小，并不会直接影响到地区的数据敏感型创业水平，故满足外生性条件。此外，进一步将2005年的历史数字基础设施水平与年份的交互项引入模型中，强化对该工具变量外生性假设的论证，以增强结果的稳健性。

基于此，使用两阶段最小二乘法进行工具变量检验，回归结果见表7。其中，（2）列为第一阶段回归结果，（3）列为第二阶段的回归结果。根据第一阶段的回归结果可以看出，所选取的工具变量（*IV*）与公共数据开放（*public_data*）存在显著的负向关系，且 Cragg-Donald Wald F 统计量为13.64，大于经验法则10，表明本文的弱工具变量问题较小。根据第二阶段的回归结果可以看出，*public_data_IV* 的回归系数在1%水平上显

表 7　内生性问题处理结果

变量	PSM-DID	工具变量法	
	entrepreneur	*public_data*	*entrepreneur*
	(1)	(2)	(3)
public_data	0.0254** (2.35)		
IV		−8.6244*** (−3.50)	
public_data_IV			0.4011*** (3.32)
控制变量	是	是	是
历史数字基础设施水平×年份	否	是	是
城市固定效应	是	是	是
时间固定效应	是	是	是
观测值	2679	2776	2776
Cragg-Donald Wald *F* 统计量		13.64	
调整 R^2	0.3350		

注：*、**、*** 分别表示在 10%、5%、1%水平上显著；括 7 号内为 *t* 值，标准误聚类到城市层面。

著为正，与式(1)的回归结果相同。根据两阶段最小二乘法的回归结果，本文所选取的工具变量能够在一定程度上缓解此模型的内生性问题。

五、机制检验

(一) 市场赋能效应

据理论分析中所述，公共数据开放能够帮助创业者获取更多关于市场需求的信息，提升新创数据敏感型企业对于未来市场需求和市场份额的预期，从而促进地区数据敏感型创业活动的开展，形成市场赋能效应。新创企业对于未来市场需求和市场份额的预期与当前市场所能达到的最大需求直接相关，即市场潜力。新创企业所能挖掘的市场潜力越大，其对于未来市场份额的预期就越高。考虑到企业对于市场份额的预期无法直接衡量，本文从市场需求潜力的角度对市场赋能效应进行了衡量。因此，参考韩峰和柯善咨[50]的测度方法来测算各地区的市场需求潜力(*potential*)，具体计算公式如式(5)所示。

$$potential_{it} = \sum_{i=1}^{n} \frac{Y_{it}}{d_{ij}^{\delta}} \tag{5}$$

其中：Y_{it} 为城市社会消费品零售总额；d_{ij} 为两个城市之间的距离；δ 为衰减参数，设为 1。为了避免 $d_{ii}=0$，令 $d_{ii}=2R_{ii}/3$，其中 R_{ii} 为城市半径。从表 8 可以看出，不管是否加入控制变量，(1)列和(3)列中公共数据开放(*public_data*)的回归系数均在 1%的水平上显著为正，说明公共数据的开放显著激发了地区的市场需求潜力。

表 8　市场赋能效应的检验结果

变量	*potential*	*entrepreneur*	*potential*	*entrepreneur*
	(1)	(2)	(3)	(4)
public_data	0.1861*** (3.55)	0.0352** (2.58)	0.1547*** (3.24)	0.0242** (2.18)
potential		0.0800*** (2.96)		0.0722*** (2.96)
控制变量	否	否	是	是
城市固定效应	是	是	是	是
时间固定效应	是	是	是	是
观测值	2560	2560	2560	2560
调整 R^2	0.9130	0.6856	0.9151	0.7212

注：*、**、*** 分别表示在 10%、5%、1%水平上显著；括号内为 *t* 值，标准误聚类到城市层面。

同时,(2)列和(4)列中市场需求潜力(*potential*)的回归系数均在 1%水平上显著为正,说明地区市场潜力的激发对地区数据敏感型创业水平的促进作用具有中介效应。据此,公共数据开放的市场赋能效应得以验证。

(二)要素赋能效应

公共数据作为数据要素本身,在不断渗透进生产领域的过程中,赋能资本要素向数字化、智能化方向转变及劳动力要素向高技能结构转变[51]。公共数据开放所引致的资本结构和劳动力结构转变本质上就是企业生产关系和生产方式的改变,主要表现为智能设备对劳动力的替代[52]。基于此,使用地区层面的工业机器人的渗透密度作为资本结构升级的衡量指标。根据许健等[53]的方法,本文首先将 IFR(international federation of robotics)数据中的行业分类与中国第二次经济普查数据的行业分类进行匹配,得到中国各行业在不同年份的工业机器人安装数量与库存数量。其次,根据各行业的机器人安装数量及库存数量与各行业的就业人数得到各行业的工业机器人密度。最后,将中国第二次经济普查年份 2008 年作为基准年份计算出各城市各行业的机器人密度的权重,将此权重与各行业的工业机器人密度相乘加权得到城市层面的工业机器人密度指标。具体计算公式见式(6)和式(7),式(6)和式(7)分别是用工业机器人安装数量与工业机器人存量计算的工业机器人密度指标。

$$install_{it} = \sum_{s=1}^{S} \frac{employ_{s,i,t=2008}}{employ_{i,t=2008}} \times \frac{INS_{st}}{employ_{s,t=2008}} \tag{6}$$

$$stock_{it} = \sum_{s=1}^{S} \frac{employ_{s,i,t=2008}}{employ_{i,t=2008}} \times \frac{STO_{st}}{employ_{s,t=2008}} \tag{7}$$

其中:s 为行业;INS_{st}、STO_{st} 分别为行业 s 在 t 年的工业机器人安装数量、工业机器人存量;$employ_{s,t=2008}$ 为行业 s 在 t 年的就业人数;$employ_{s,i,t=2008}$ 为城市 i 的行业 s 在 2008 年的就业人数;$employ_{i,t=2008}$ 为城市 i 在 2008 年的就业总人数。

从表 9 的(1)列和(3)列可以看出,公共数据开放(*public_data*)的回归系数均在 1%水平上显著为正,说明公共数据开放能够显著促进资本结构的升级。(2)列和(4)列中工业机器人渗透密度(*install*、*stock*)的回归系数均显著为正,说明资本结构升级能够显著促进地区的数据敏感型创业水平。据此,公共数据开放的要素赋能效应得以验证。

表 9　要素赋能效应的检验结果

变量	机器人安装量		机器人存量	
	install	*entrepreneur*	*stock*	*entrepreneur*
	(1)	(2)	(3)	(4)
public_data	1.6323***	0.0330***	10.2817***	0.0322**
	(5.17)	(2.73)	(5.29)	(2.46)
install		0.035***		
		(2.96)		
stock				0.0006**
				(2.14)
控制变量	是	是	是	是
城市固定效应	是	是	是	是
时间固定效应	是	是	是	是
观测值	2780	2780	2780	2780
调整 R^2	0.8517	0.6840	0.8268	0.6827

注:*、**、*** 分别表示在 10%、5%、1%水平上显著;括号内为 t 值,标准误聚类到城市层面。

(三)制度赋能效应

公共数据开放有助于进一步优化地区的营商环境,从而降低企业的制度性交易成本,以及企业的创业活动成本,促进了地区的数据敏感型创业。参考“中国城市营商环境评价研究”课题组等[54]的研究方法,结

合数据的可获得性并剔除与本文控制变量相关联的维度指标，从公共服务、人力资源、市场环境、创新环境、金融服务五个方面构建地区营商环境综合指数②。从表10可以看出，中国城市营商环境综合指数作为营商环境（*environment*）的衡量指标，公共数据开放（*public_data*）的回归系数均在1%水平上显著为正，说明公共数据开放显著优化了地区的营商环境。（2）列中营商环境（*environment*）的回归系数均在1%水平上显著为正，说明地区营商环境的优化能够显著提升地区的数据敏感型创业水平。据此，公共数据开放的制度赋能效应得以验证。

表10　制度赋能效应的检验结果

变量	*environment*	*entrepreneur*
	（1）	（2）
public_data	0.0112***	0.0331***
	（2.98）	（3.08）
environment		0.4833***
		（3.06）
控制变量	是	是
城市固定效应	是	是
时间固定效应	是	是
观测值	2740	2740
调整 R^2	0.8983	0.6885

注：*、**、*** 分别表示在10%、5%、1%水平上显著；括号内为 *t* 值，标准误聚类到城市层面。

六、进一步分析

（一）高质量的公共数据开放与地区数据敏感型创业

根据理论分析中所述，随着公共数据开放质量的提升，其作为数据要素的价值释放作用越强，从而能够进一步强化公共数据开放的数据敏感型创业效应。基于此，根据《中国地方政府数据开放报告（指标体系与标杆体系）》，本文通过中国城市开放数林综合指数来衡量各地区的公共数据开放质量（*quality*），该指数越高说明公共数据开放的质量越高。从表11可以看出，不管是在何种情况下，公共数据开放质量（*quality*）的回归系数均在1%的水平下显著为正，说明随着公共数据开放质量的提升，公共数据开放对地区数据敏感型创业的促进作用也随之增强。

表11　考虑公共数据开放质量的估计结果

变量	（1）	（2）	（3）	（4）
quality	0.0259***	0.0153***	0.0184***	0.0109***
	（5.82）	（3.42）	（4.79）	（3.10）
控制变量	否	否	是	是
城市固定效应	是	是	是	是
时间固定效应	否	是	否	是
观测值	2780	2780	2780	2780
调整 R^2	0.5557	0.6341	0.6233	0.6756

注：*、**、*** 分别表示在10%、5%、1%水平上显著；括号内为 *t* 值，标准误聚类到城市层面。

（二）公共数据开放与地区整体的创业活跃度

公共数据开放能够通过要素赋能效应促进企业的要素升级和管理优化，扩大了对高技能人才的需求，促进了创业活跃度的提升，尤其是数据敏感型创业[55]。基于此，进一步探讨公共数据开放能否提升地区整体的创业活跃度。选取每百人中新创企业数（*entre_activity*）作为地区创业活跃度的衡量指标，回归结果见表12。不论是在何种情形下，公共数据开放（*public_data*）的回归系数均在1%的水平下显著为正，说明公共数据开放能够推动地区整体创业活跃度的提升。

表12　考虑地区整体创业活力的估计结果

变量	（1）	（2）	（3）	（4）
public_data	0.9223***	0.3306***	0.5660***	0.2176***
	（9.36）	（3.11）	（5.92）	（2.71）
控制变量	否	否	是	是
城市固定效应	是	是	是	是
时间固定效应	否	是	否	是
观测值	2780	2780	2780	2780
调整 R^2	0.1062	0.7693	0.7288	0.7883

注：*、**、*** 分别表示在10%、5%、1%水平上显著；括号内为 *t* 值，标准误聚类到城市层面。

（三）信息获取能力与公共数据开放的数据敏感型创业效应

根据理论分析中所述，信息获取能力的提升不仅能够降低创业者获取数据要素的门槛，加速其数据价值的转化，而且能够减弱创业者对政策不确定性的感知程度，进而激发其创业意愿，尤其是数据敏感型创业[43]。基于此，选取各城市的互联网普及率（*internet*）及移动电话普及率（*phone*）来衡量各城市的信息获取程度。互联网普及率和移动电话普及率越高，意味着获取信息的程度就越强。本文分别将公共数据开放与互联网普及率的交互项和公共数据开放与移动电话普及率的交互项代入式（1）。从表13可以看出，除了（6）

② 公共服务包括供气能力、公共供水能力、工业供电能力、医疗卫生服务；人力资源包括年末单位从业人员、人口净流入、平均工资水平；市场环境包括固定资产投资总额、当年实际使用外资金额、当年新签项目个数；创新环境包括发明专利授权量；金融服务包括金融从业人员。

表13　考虑信息获取能力的估计结果

变量	(1)	(2)	(3)	(4)	(5)	(6)
public_data	−0.0245 (−0.97)	−0.0603** (−2.57)	−0.0477** (−2.30)	−0.0232 (−0.86)	−0.0481* (−1.70)	−0.0409 (−1.63)
internet	0.0007** (2.42)	−0.0002 (−0.80)	−0.0001 (−0.80)			
public_data×*internet*	0.0025*** (2.95)	0.0027*** (3.47)	0.0021*** (3.03)			
phone				0.0006*** (4.59)	0.0001 (0.53)	0.0002 (1.04)
public_data×*phone*				0.0007*** (3.35)	0.0006*** (2.77)	0.0005** (2.54)
控制变量	否	否	是	否	否	是
城市固定效应	是	是	是	是	是	是
时间固定效应	否	是	是	否	是	是
观测值	2780	2780	2780	2780	2780	2780
调整 R^2	0.6147	0.6719	0.6992	0.6365	0.6907	0.7119

注：*、**、*** 分别表示在10%、5%、1%水平上显著；括号内为 t 值，标准误聚类到城市层面。

列中的交互项系数在5%水平上显著为正以外，公共数据开放分别与互联网普及率和移动电话普及率的其余交互项系数均在1%水平下显著为正，说明随着地区信息获取程度的增强，公共数据开放对地区数据敏感型创业的促进作用也随之增强。

（四）公共数据开放与地区数据敏感型创业的城市市场分割程度异质性

探究城市市场分割程度的不同是否会对公共数据开放的数据敏感型创业效应产生异质性影响。具体而言，参考盛斌和毛其淋[56]的方法使用价格指数法来计算市场分割指数。若城市市场分割指数大于等于均值，则视为市场分割程度较高的城市，赋值为1；若城市市场分割指数小于均值，则视为市场分割程度低的城市，赋值为0。从表14可以看出，低市场分割城市的公共数据开放（*public_data*）系数在1%的水平上显著为正，而高市场分割城市的公共数据开放系数（*public_data*）不显著。为了进一步检验组间系数是否存在差异，采用费舍尔组合检验法进行检验。费舍尔组合检验的系数在10%水平上显著，说明组间系数存在显著差异。这说明公共数据开放对地区数据敏感创业的促进作用在低市场分割城市更明显。相比于高市场分割城市，低市场分割城市往往存在更加统一的要素市场和更低的市场壁垒，要素流通更加自由，具有更高的资源配置效率。因此，创业者能够以更高的效率将数据要素与其他生产要素相结合，实现数据要素成果的孵化，激发数据敏感型创业的意愿。

表14　考虑异质性的估计结果

变量	高市场分割城市	低市场分割城市
	(1)	(2)
public_data	0.0600 (1.49)	0.0360*** (3.08)
控制变量	是	是
城市固定效应	是	是
时间固定效应	是	是
费舍尔组合检验	−0.024*	
观测值	882	1892
调整 R^2	0.7635	0.6279

注：*、**、*** 分别表示在10%、5%、1%水平上显著；括号内为 t 值，标准误聚类到城市层面。

七、结论与建议

（一）研究结论

在2010—2019年的研究区间内，地方政府上线公共数据开放平台后，数据敏感型企业的新建数量会出现显著的增加，即公共数据开放能够显著提升地区的数据敏感型创业水平，并且该结果通过了一系列的稳健性检验。本文为理解数字经济时代下政府的“赋能者”角色提供了新的视角，延伸了传统公共治理理论中政府的服务职能，丰富了传统公共治理理论的内涵。同时，揭示了公共数据作为公共资源释放公共价值的方式，从而拓展公共价值的生成路径，丰富了公共价值理论的研究视角。随着公共数据开放质量和地区信息获取程度的提升，公共数据开放对地区数据敏感型创业活力的影响效应也随之提升。研究有助于更加全

面地认识公共数据开放与地区数据敏感型创业活动之间的关系，为相关政策的科学制定提供针对性的启示。

（二）政策建议

第一，聚焦公共数据质量评估与分类分级开放，推动与非公共数据的可信融合。各级政府应建立常态化的数据质量评估机制，从准确性、完整性、一致性、时效性等维度制定可量化的评估标准，并定期发布质量报告，确保开放数据的可用性与可靠性。在此基础上，依据数据安全与隐私风险，对公共数据进行精细化的分类分级，明确不同数据的开放等级，以及针对不同主体设定差异化的访问权限和使用规则。同时，积极探索公共数据与非公共数据的安全融合路径，通过可信数据空间、隐私计算、联合建模等技术手段，在不泄露原始数据的前提下，形成融合数据产品，服务于精准决策、市场研判、创新孵化等场景。

第二，不断完善数字基础设施建设，构建数据要素价值评估体系，推动高质量的公共数据开放，带动高质量创业。各级政府应致力于建设数字基础设施，提高互联网普及率，拓宽不同经济行为主体获取信息的渠道，在提升信息透明度的同时丰富经济主体对信息的获取途径。同时，各级政府应积极构建数据要素价值评估体系，对公开数据的真实性和时效性等方面的价值进行系统评估，同时加强公共数据的开放广度和开放深度，实现以高质量的数据开放带动高质量的创业，推动经济高质量发展。

（三）研究展望

本文采用信息传输、软件和信息技术服务业及科学研究和技术服务业新创企业数作为数据敏感型创业的代理变量，主要基于该类行业具有较高的数据密集性与技术依赖性，与数据要素的关联较为直接。然而，金融、零售、制造等其他行业中同样存在大量对数据高度敏感的创业活动，这些行业的创业行为同样受公共数据开放的显著影响。在识别过程中，考虑到数据的可获取性，以及数据敏感程度缺少统一的衡量指标，本文只选取了与数据要素的关联较为直接的行业而未将其他行业中对数据敏感的创业活动纳入本代理变量的覆盖范围。因此，对数据敏感型创业的度量可能存在部分行业选择偏差，从而在一定程度上低估或窄化了公共数据开放对数据敏感型创业的影响。希望未来研究可尝试构建统一识别指标，更加精准地识别数据敏感型创业。

参考文献

[1] SCHUMPETER J A. The theory of economic development: An inquiry into profits, capital, credit, interest, and the business cycle[M]. Cambridge: Harvard University Press, 1934.

[2] 杜运周，任兵，张玉利. 新进入缺陷、合法化战略与新企业成长[J]. 管理评论，2009，21(8)：57-65.

[3] 黄先海，党博远，宋安安，等. 新发展格局下数字化驱动中国战略性新兴产业高质量发展研究[J]. 经济学家，2023(1)：77-86.

[4] JONES C I, TONETTI C. Nonrivalry and the economics of data[J]. American Economic Review, 2020, 110(9): 2819-2858.

[5] AUTOR D, DORN D, KATZ L F, et al. The fall of the labor share and the rise of superstar firms[J]. Quarterly Journal of Economics, 2020, 135(2): 645-709.

[6] VELDKAMP L. Valuing data as an asset[J]. Review of Finance, 2023, 27(5): 1545-1562.

[7] 沈坤荣，林剑威. 数据垄断问题研究进展[J]. 经济学动态，2024(3)：129-144.

[8] 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，2024，47(2)：174-203.

[9] 李春涛，闫续文，宋敏，等. 金融科技与企业创新——新三板上市公司的证据[J]. 中国工业经济，2020(1)：81-98.

[10] 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.

[11] ROSS P K, BLUMENSTEIN M. Cloud computing as a facilitator of SME entrepreneurship[J]. Technology Analysis & Strategic Management, 2015, 27(1): 87-101.

[12] 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，2020，36(10)：65-76.

[13] 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，2024，41(8)：5-25.

[14] 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，2024，46(1)：25-46.

[15] 郑志强，何佳俐. 社会信用、交易成本与企业专业化分工[J]. 财经研究，2024，50(2)：139-153.

[16] GOLDFARB A, TUCKER C. Digital economics[J]. Journal of Economic Literature, 2019, 57(1): 3-43.

[17] 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，2023，40(9)：50-70.

[18] VERHULST S, ANDREW Y. The global impact of open data: Key findings from detailed case studies around the world[M]. Sebastopol, CA:

O'Reilly Media, 2016.
[19] 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, 39(9): 124-142.
[20] ALVAREZ S, BARNEY J B. Has the concept of opportunities been fruitful in the field of entrepreneurship?[J]. Academy of Management Perspectives, 2020, 34(3): 300-310.
[21] 胡厚全. 企业家精神与中国经济增长：基于历史传承的视角[J]. 系统工程理论与实践, 2022, 42(6): 1481-1496.
[22] DAVIDSSON P, HONIG B. The role of social and human capital among nascent entrepreneurs[J]. Journal of Business Venturing, 2003, 18(3): 301-331.
[23] SIQUEIRA A C O. Entrepreneurship and ethnicity: The role of human capital and family social capital[J]. Journal of Developmental Entrepreneurship, 2007, 12(1): 31-46.
[24] JIA R, LAN X, I MIQUEL G P. Doing business in China: Parental background and government intervention determine who owns business[J]. Journal of Development Economics, 2021, 151: 102670.
[25] 杜运周, 任兵, 张玉利. 新进入缺陷, 合法化战略与新企业成长[J]. 管理评论, 2009, 21(8): 57-65.
[26] BRADLEY S W, SHEPHERD D A, WIKLUND J. The importance of slack for new organizations facing "Tough" environments[J]. Journal of Management Studies, 2011, 48(5): 1071-1097.
[27] BAUMOL W J. Entrepreneurship: Productive, unproductive, and destructive[J]. Journal of Political Economy, 1990, 98(5): 893-921.
[28] 施丽芳, 廖飞, 丁德明. 制度对创业家行动的影响机理——基于不确定管理的视角[J]. 中国工业经济, 2014(12): 118-129.
[29] YE W, SU Z, AHLSTROM D. Bankruptcy laws, entrepreneurs' socio-cognitions, and the pursuit of innovative opportunities[J]. Small Business Economics, 2022, 59(3): 1005-1022.
[30] 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？[J]. 管理世界, 2020, 36(9): 141-154.
[31] LIM D S K, OH C H, DE CLERCQ D. Engagement in entrepreneurship in emerging economies: Interactive effects of individual-level factors and institutional conditions[J]. International Business Review, 2016, 25(4): 933-945.
[32] DABLA-NORRIS E, JI Y, TOWNSEND R M, et al. Distinguishing constraints on financial inclusion and their impact on GDP, TFP, and the distribution of income[J]. Journal of Monetary Economics, 2021, 117: 1-18.
[33] ALDRICH H E, RUEF M. Organizations evolving[M]. London: Sage, 2006.
[34] MCCOY D, LYONS S, MORGENROTH E, et al. The impact of broadband and other infrastructure on the location of new business establishments [J]. Journal of Regional Science, 2018, 58(3): 509-534.
[35] KERR S P, KERR W, ÖZDEN C, et al. High-skilled migration and agglomeration[J]. Annual Review of Economics, 2017, 9(1): 201-234.
[36] 朱秀梅, 李明芳. 创业网络特征对资源获取的动态影响——基于中国转型经济的证据[J]. 管理世界, 2011, 27(6): 105-115, 188.
[37] EREL I, LIEBERSOHN J. Can fintech reduce disparities in access to finance? Evidence from the paycheck protection program[J]. Journal of Financial Economics, 2022, 146(1): 90-118.
[38] GAMBACORTA L, HUANG Y, LI Z, et al. Data versus collateral[J]. Review of Finance, 2023, 27(2): 369-398.
[39] 谢绚丽, 沈艳, 张皓星, 等. 数字金融能促进创业吗？——来自中国的证据[J]. 经济学(季刊), 2018, 17(4): 1557-1580.
[40] 焦豪, 崔瑜, 张亚敏. 数字基础设施建设与城市高技能创业人才吸引[J]. 经济研究, 2023, 58(12): 150-166.
[41] TIMMONS A, SPINELLI S. New venture creation: Entrepreneurship for the 21st century[M]. New York: Mc Graw Hill, 2008.
[42] 刘诚, 夏杰长. 线上市场、数字平台与资源配置效率：价格机制与数据机制的作用[J]. 中国工业经济, 2023(7): 84-102.
[43] 何雨可, 牛耕, 逯建, 等. 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, 41(1): 47-66.
[44] 陈潭. 国家治理的数据赋能及其秩序生产[J]. 社会科学研究, 2023(6): 12-30.
[45] 白俊红, 张艺璇, 卞元超. 创新驱动政策是否提升城市创业活跃度——来自国家创新型城市试点政策的经验证据[J]. 中国工业经济, 2022(6): 61-78.
[46] ROTH J, SANT'ANNA P H C, BILINSKI A, et al. What's trending in difference-in-differences? A synthesis of the recent econometrics' literature[J]. Journal of Econometrics, 2023, 235(2): 2218-2244.
[47] RAMBACHAN A, ROTH J. A more credible approach to parallel trends[J]. The Review of Economic Studies, 2023, 90(5): 2555-2591.
[48] DE CHAISEMARTIN C, D' HAULTFOEUILLE X. Two-way fixed effects estimators with heterogeneous treatment effects[J]. American Economic Review, 2020, 110(9): 2964-2996.
[49] 吴思栩, 李杰伟. "数字经济"时代城市的未来——互联网对中国城市生产性服务业集聚的影响研究[J]. 经济学(季刊), 2024, 24(2): 431-447.
[50] 韩峰, 柯善咨. 追踪我国制造业集聚的空间来源：基于马歇尔外部性与新经济地理的综合视角[J]. 管理世界, 2012, 28(10): 55-70.
[51] 肖土盛, 孙瑞琦, 袁淳, 等. 企业数字化转型、人力资本结构调整与劳动收入份额[J]. 管理世界, 2022, 38(12): 220-237.
[52] ACEMOGLU D, RESTREPO P. The race between man and machine: Implications of technology for growth, factor shares, and employment[J].

American Economic Review, 2018, 108(6): 1488-1542.
[53] 许健, 季康先, 刘晓亭, 等. 工业机器人应用、性别工资差距与共同富裕[J]. 数量经济技术经济研究, 2022, 39(9): 134-156.
[54] "中国城市营商环境评价研究"课题组, 李志军, 张世国, 等. 中国城市营商环境评价的理论逻辑、比较分析及对策建议[J]. 管理世界, 2021, 37(5): 98-112, 8.
[55] 田鸽, 黄海, 张勋. 数字金融与创业高质量发展：来自中国的证据[J]. 金融研究, 2023(3): 74-92.
[56] 盛斌, 毛其淋. 贸易开放、国内市场一体化与中国省际经济增长：1985—2008 年[J]. 世界经济, 2011(11): 44-66.

Can the Opening of Public Data Alleviate the Liability of Newness? Empirical Evidence from a Quasi-natural Experiment on a Public Data Platform

Bian Yuanchao, Su Jiahao
(School of Business, Nanjing Normal University, Nanjing 210023, China)

Abstract: In the digital economy era, the liability of newness is mainly manifested as the inability of datasensitive entrepreneurial activities to enter the market due to data scarcity. Based on the dual attributes of public data as both a data factor and a public service, the effect of open public data on regional datasensitive entrepreneurship was examined over the sample period from 2010 to 2019. A quasinatural experiment based on the launch of open public data platforms by local governments was employed to identify the causal impact. The empirical results indicate that open public data significantly promotes regional datasensitive entrepreneurship and alleviates the liability of newness. Mechanism analysis reveals that open public data releases regional market demand potential, provides more highquality production factors, and improves the business environment. These effects stimulate the market entry of datasensitive entrepreneurs, forming marketenabling, factorenabling, and institutionalenabling effects. Further analysis shows that the positive effect is amplified with better data quality and higher regional information accessibility. Open public data also enhances overall regional entrepreneurial activity. The facilitating effect is stronger in cities with lower degrees of market segmentation. These findings may offer implications for overcoming the liability of newness in the digital economy and for creating a digital environment that supports entrepreneurial activities.

Keywords: public data opening; data-sensitive entrepreneurship; liability of newness; market empowerment; factor empowerment; institutional empowerment