

# 公共数据开放政策的就业效应研究

郝楠, 张昕怡

(安徽大学 经济学院, 安徽 合肥, 230601)

**摘要:** 就业是民生之本, 如何从政策层面有效推动企业就业是值得关注和研究的重要议题。基于2007—2021年A股上市企业的数据, 利用多期双重差分法探究公共数据开放政策对企业就业的影响。研究发现: (1) 公共数据开放政策的实施能够提升企业就业水平; (2) 该政策通过促进企业数字化转型、提升企业创新水平和营运效率来促进就业; (3) 金融发展水平越高的地区, 这种促进效应越明显; (4) 该政策对国有企业、服务业和东部地区的促进作用更明显; (5) 公共数据开放对高技能和低技能劳动力产生就业促进作用, 抑制中等技能劳动力的就业水平, 并扩大了企业内部的薪酬差距。本文的研究为实现稳就业和收入分配公平合理的目标提供了重要参考。

**关键词:** 公共数据开放; 数据要素; 就业水平; 技能结构; 薪酬差距

**中图分类号:** F241.4 **文献标识码:** A **文章编号:** 1008-2700 (2025) 04-0066-17

## 一、问题提出

随着信息技术、大数据、人工智能等领域的迅猛发展, 数据要素的重要性日益凸显。目前, 数据已成为国家基础战略竞争的核心资源, 亦是新质生产力发展的核心和关键要素<sup>[1]</sup>。公共数据开放政策是激活数据要素价值的关键入口, 其通过释放政府及公共部门的数据资源, 推动数据要素的高效流通与市场化配置, 为企业提供了优化决策、创新服务和拓展市场的基础性资源。数据要素的充分流动和深度应用能够赋能企业优化决策、创新商业模式、拓展市场空间, 进而通过产业升级和新兴业态的涌现创造更多就业岗位, 形成政策供给—要素赋能—就业升级的协同发展路径, 为数字时代下稳就业与经济增长提供新的政策着力点。

中国政府高度重视公共数据资源的制度化设计 (见表1)。2015年8月, 国务院发布《促进大数据发展行动纲要》, 强调要“加快政府数据开放共享”“稳步推动公共数据资源开放”。这也是政府数据开放政策首次被纳入国家级的政策文件中。为实现数据大国向数据强国的转变, 工业和信息化部在2016年12月发布的《大数据产业发展规划 (2016-2020年)》中强调了持续改进大数据标准体系的必要性, 并提出要在重点行业、领域、地区开展若干标准试验和试点示范, 充分利用现有数据资源, 合理规划大数据基础设施建设。2018年1月, 中央网信办、国家发展改革委与工业和信息化部联合发布了《公共信息资源开放试点工作方案》, 确定在北京市、上海市、浙江省、福建省、贵州省开展公共数据资源开放的试点工作, 并重点开放信用服务、社保就业等领域。为保证试点工作的顺利开展, 2020年发布的《中共中央国务院关于构建更加完善的要素市场化配置体制机制的意见》、2022年发布的《中共中央国务院关于构

收稿日期: 2024-12-11; 修回日期: 2025-05-28

基金项目: 安徽省社会科学创新发展研究课题“安徽新就业形态劳动过程与劳动关系协调机制研究”(2022CX052)

作者简介: 郝楠 (1983—), 女, 安徽大学经济学院副教授; 张昕怡 (2000—), 女, 安徽大学经济学院硕士研究生, 通信作者。

建数据基础制度更好发挥数据要素作用的意见》，均强调加快推进政府数据资源开放共享。2024 年 9 月，《中共中央 国务院关于加快公共数据资源开发利用的意见》的出台，进一步推动公共数据资源的深度开发与高效利用，充分释放其内在价值与要素潜能，助力经济社会的全面发展。

表 1 公共数据开放政策梳理

| 发文时间        | 政策文件名称                            | 主要内容                                                  |
|-------------|-----------------------------------|-------------------------------------------------------|
| 2015 年 8 月  | 《促进大数据发展行动纲要》                     | 明确设立国家政府数据统一开放平台的建设目标                                 |
| 2016 年 12 月 | 《大数据产业发展规划（2016–2020 年）》          | 推动跨行业、跨领域的數據应用，制定数据开放共享制度                             |
| 2017 年 2 月  | 《关于推进公共信息资源开放的若干意见》               | 协同推进公共数据开放共享，加强关键领域信息资源的整合利用，释放数据要素经济社会价值             |
| 2017 年 5 月  | 《政务信息系统整合共享实施方案》                  | 加速公共数据开放网站建设进程，确保民生、企业服务等重点领域数据资源的优先开放，并依法依规向社会公开     |
| 2017 年 6 月  | 《政务信息资源目录编制指南（试行）》                | 清晰界定公共数据的开放条件与共享范围，通过统一平台实现数据的标准化共享与社会化应用             |
| 2018 年 1 月  | 《公共信息资源开放试点工作方案》                  | 选定北京市、上海市、浙江省、福建省、贵州省为试点区域，开展公共信息资源开放试点工作，并明确重点开放领域   |
| 2020 年 4 月  | 《中共中央 国务院关于构建更加完善的要素市场化配置体制机制的意见》 | 推进政府数据开放共享，提升社会数据资源价值，同时加强数据资源整合与安全保护，制定出台新一批数据共享责任清单 |
| 2021 年 6 月  | 《中华人民共和国数据安全法》                    | 国家通过制定数据开放目录、构建统一规范且安全可控的开放平台，全面提升公共数据的开放质量与利用效率      |
| 2022 年 12 月 | 《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》 | 推动公共数据确权授权机制建设，加强数据汇聚、共享和开放开发，以激活数据要素潜能               |
| 2023 年 2 月  | 《数字中国建设整体布局规划》                    | 构建统一的国家公共数据开放平台和推动公共数据汇聚利用，实现公共数据资源的高效配置              |
| 2024 年 9 月  | 《中共中央 国务院关于加快公共数据资源开发利用的意见》       | 完善公共数据开放政策体系，提高开放数据的完整性、准确性、及时性和机器可读性                 |

在顶层设计的指导下，各级地方政府积极响应，加快推进公共数据开放平台的建设工作。2012 年上海市人民政府工作报告强调推动政府部门将可公开的经济和社会發展数据以及各类基础政务数据向社会开放的重要性，至此，上海市成为全国首个开放公共数据的城市。随后，无锡市、佛山市、湛江市、武汉市等城市于 2014 年上线公共数据平台，城市公共数据开放平台数量在 2017 年之后开始显著增长。截至 2023 年 8 月，已有 204 个城市建立了数据开放平台，占城市总数的 60.53%<sup>[2]</sup>。在这些城市中，上海市以最佳的综合表现位居榜首，紧随其后的是天津市、北京市和重庆市。未来平台开放数量将持续增加，数据资源将对经济社会发展发挥更重要的作用。虽然中国在政府数据开放方面已经取得了一些进步，但仍面临诸多问题和挑战。一方面，数据质量与价值未能契合公众期望，可利用性不高，常常出现缺少数据的情况，致使工作流于形式；另一方面，相关法律法规尚未完善，存在立法框架残缺、平台间协同困难、数据分布失衡等问题。此外，各地方数据开放平台在内容与标准上的不一致、平台与数据集数量的不协调以及隐私泄露风险等问题不容忽视。因此，中国应从政策制定、国家级平台建设、数据利用限制及与数据使用者互动等方面切入，以更好地推进地方公共数据开放工作。

面对当前中国经济所遭遇的需求减少、供给干扰和预期低迷等三重挑战，就业市场正遭受着空前的

压力和考验。一方面,随着人口红利逐渐消失,经济增长放缓,整体就业市场面临较大的就业总量压力。例如,每年有大量高校毕业生涌入就业市场,2025 年高校毕业生将达到 1 122 万人,未来很长一段时间都将保持每年千万级的规模,这给就业市场带来了巨大的挑战。另一方面,就业市场中不仅存在“供大于求”的现象,同时也存在“有活没人干”和“有人没活干”的结构性矛盾,具体表现为部分行业和地区存在劳动力短缺,但某些行业和岗位却面临劳动力过剩的问题。企业不仅是就业的创造主体,亦是重要的市场主体。开放政府公共数据平台能否为企业就业注入数据驱动的经济效益,进而成为促进就业的有效手段,对于维护中国劳动力就业市场的稳定和驱动经济实现高质量发展具有重要的理论意义和实践意义。

二、文献回顾与研究假设

(一) 文献回顾

中国公共数据开放平台的建设始于 2012 年,城市的行政级别、发展水平是影响开放绩效的关键因素<sup>[3]</sup>。随着公共数据开放政策的不断推进,近两年学术界相关主题的实证类文章开始涌现,其中关于数据开放平台效率的测算引发了热烈讨论。以上海市、北京市等 8 个具有代表性的政府开放数据网站为例,经过数据包络分析(DEA)测算后发现上海市在开放效率方面表现突出,全国的政府数据开放平台效率水平整体呈上升趋势<sup>[4]</sup>。目前,现有研究已扩展至政策评估方面。对于宏观主体而言,政府公共数据开放有利于提升城市创新水平<sup>[5]</sup>,并通过提高城市创业活跃度促进要素流动<sup>[6]</sup>,从而提高资源配置效率<sup>[7]</sup>,缩小城市内部不同区域的发展差距,推动区域间的协调发展<sup>[8]</sup>,对经济增长产生积极影响<sup>[9]</sup>。对于微观主体企业而言,公共数据开放平台是企业获取数据的重要途径,数据资产能力的提升有效增强了企业的核心价值能力,通过提升企业营运效率<sup>[10]</sup>和减轻资金筹集的难度<sup>[11]</sup>等方式推动企业创新,助力企业实现数字化转型<sup>[12]</sup>。此外,公共数据开放还能够有效减少企业和消费者的搜寻成本<sup>[13]</sup>,实现供需精准对接,在提高企业全要素生产率的同时推动商业领域新业态的产生和发展<sup>[14]</sup>。少数文献关注到公共数据开放对企业招聘和劳动者就业的影响,认为公共数据开放可以通过缓解融资约束、促进企业创新等渠道来扩大企业用工规模、优化劳动者就业结构<sup>[15-16]</sup>。

综上所述,关于公共数据开放研究的深度和广度正在不断拓展,已从初期的现状分析、影响因素分析、效果评估等定性研究层面转入实证分析,这些成果为本文以及未来相关研究提供了坚实的理论支撑。但从政策视角出发,关于公共数据开放政策的效果评估主要集中于经济效应,而对社会效应特别是就业效应的探讨相对较少;从研究内容来看,一些文献考察了公共数据开放政策的就业效应,但其关注的是城市级层面的宏观就业影响<sup>[17]</sup>,未能揭示其对企业就业水平的作用。部分研究虽然聚焦企业层面的就业效应,但其使用的是省级政府开放平台数据<sup>[16]</sup>,难以捕捉城市间政策实施的异质性特征,这会影响到政策效应评估的精细度。上述文献为本文的研究提供了可以拓展的空间,进一步探究公共数据开放如何赋能企业就业不仅有助于深化对数据要素价值的理解,而且可为后续探寻更具韧性的就业路径、制定精准就业促进政策提供思路。

本文的边际贡献可能在于以下三个方面。第一,提供就业影响因素研究新视角。就业作为头等民生大事,一直是学术研究领域关注的重点议题。已有研究围绕就业的影响因素展开了深入且广泛的学术研讨,主要涉及人工智能<sup>[18]</sup>、企业数字化转型<sup>[19]</sup>等方面。本文基于政府服务视角,将公共数据开放政策与企业就业纳入统一分析框架,厘清具体影响机制。第二,细化公共数据开放政策的社会效应研究。与现有研究使用的省级公共数据开放平台不同<sup>[15]</sup>,本文将各地级市公共数据开放平台的上线时间视为一项准自然实验,通过运用多期双重差分(DID)法分析其对企业就业水平的影响效应,为促进稳就业这一目标的实现提供参考。第三,阐明公共数据开放政策影响效应的异质性特征并赋能精准施策。本文从企业所有制性质、企业所属行业、企业所在城市等方面探讨公共数据开放对企业就业的异质性影响,为政府制定差异化的政策提供经验证据。



## (二) 公共数据开放对企业就业的直接影响效应

就业问题产生的根源之一是高信息搜寻成本。在就业市场中,劳动力的供给方与需求方都在努力寻找着彼此满意的合作伙伴,以完成最佳的匹配。但在劳动力市场信息不对称这一现实背景下,企业招聘面临双重成本约束:既要支付岗位空缺产生的机会成本,又要承担搜寻适配人才的信息成本。政府公共数据开放平台作为政府信息资源的传播媒介,不仅是数据存放和共享平台,更是一个连接企业和个人,推动数据资源共享、利用的重要桥梁,有效降低了企业的岗位匹配成本。公共数据开放为企业注入新动能,数据要素与企业核心业务的深度融合正重构着企业用工需求结构。企业通过挖掘行业数据资源,能够精准识别市场缺口、优化管理效能。在此过程中,具有数据解析能力的复合型人才成为企业竞逐的关键资源,促使企业增加对工业互联网工程技术人员、数字孪生应用工程师等职位的雇佣<sup>[18]</sup>。据此,本文提出研究假设1。

H1: 公共数据开放有助于提升企业就业水平。

## (三) 公共数据开放对企业就业的间接作用机制

### 1. 数字化转型效应

数据是企业数字化转型的核心驱动要素。公共数据开放能够加快数据要素流动,为企业数字化转型提供充足的数据资源,减轻生产过程中企业所面临的要素成本压力。同时,与传统生产要素不同,数据作为一种生产要素,具有可复制、非竞争和非排他的特性。政府主导的公共数据开放作为数据价值发挥的前提,具有极强的公共性、原始性和权威性,为需要使用的社会主体提供了广阔的整合、提炼与分析空间<sup>[20]</sup>,为企业打破数据垄断、联通“数据孤岛”、降低数据获取门槛注入强劲推力,正向推动企业数字化转型。此外,要将公开的数据资源转化为实际的生产要素,企业必须具备出色的数据识别、分析和应用技能,从而倒逼企业进行数字化转型。

一方面,企业数字化转型对企业就业产生替代效应<sup>[21]</sup>。当数字技术相较于人力展现出更大的竞争优势时,人力可能会被这些技术所取代。例如,企业使用自动化机械装置来减少对重复性高、程序复杂的岗位的雇佣,最终导致企业就业人数下降。另一方面,企业数字化转型对企业就业产生创造效应<sup>[19]</sup>。从成本角度来看,企业通过利用互联网、人工智能和云计算等现代技术进行数字化转型,提高了生产效率,产品的制造成本和售价随之降低,产品需求量上升。这种增长的市场需求有助于产业规模的扩大,从而增加劳动力雇佣数量<sup>[22]</sup>。此外,企业数字化转型代表数字经济与实体经济紧密结合的创新商业模式,提升了企业在研发、管理与运营等领域的劳动力需求,新增的软件开发师和网络安全专家等与高技能人才相匹配的岗位,进一步增强了企业吸纳劳动力的能力。鉴于中国目前正处于数字化转型的初期阶段,可以预测替代效应小于创造效应<sup>[23]</sup>。据此,本文提出研究假设2。

H2: 公共数据开放通过企业数字化转型促进就业。

### 2. 创新驱动效应

创新是一项依赖资源的消耗性活动,其实质是企业对生产要素进行新的配置和组合。政府推动的公共数据开放平台充当了公共数据发布者和使用者之间的桥梁,有利于消除生产要素在产业中形成的“租”<sup>[24]</sup>,从根本上促进公共数据向企业的集聚。企业获得更多开放数据在一定程度上说明其拥有更多的新型生产要素,通过与传统生产要素结合实现生产要素配置优化,进而充分激发数据价值和企业创新动力。

创新被视为推动经济增长和提高生产效率的关键因素,同时也是创造就业机会的重要途径。根据不同的类型,创新可以被细分为产品创新和流程创新<sup>[25]</sup>,两者都是企业不断创新与研发的结果。从就业创造机制来看,产品创新通过需求扩张效应成为就业增长的核心引擎。通过产品创新,企业不仅能够提升现有产品与服务的品质,而且能够扩展产品种类,进而打开新的市场空间,催生新的消费需求,实现销售额的增长和生产规模的扩大,为就业市场带来新的增长点。相较而言,流程创新主要通过效率提升效应影响就业。当企业利用开放数据进行生产流程智能化改造时,生产效率得以提升,即生产相同数量的

产品或服务所需的劳动力数量相应减少, 从而降低就业率。但现有研究已指出这一负面影响较小<sup>[26]</sup>。据此, 本文提出研究假设 3。

H3: 公共数据开放通过提升企业创新水平促进就业。

3. 营运效率效应

开放公共数据的核心目的在于减轻企业在获取数据这一生产要素时的成本负担, 企业无须经过复杂的授权过程, 即可免费获取数据, 有助于减少信息不对称现象, 从而增强管理效能, 如解决企业管理中决策者的有限理性问题<sup>[27]</sup>, 以及限制交易对手的机会主义行为。再者, 数据要素有别于资本和劳动力等传统生产要素, 它所具备的高流动性和规模报酬递增的特性, 使其成为一种高效的新型生产要素。生产要素的流动性越高, 越容易形成聚集、积累和乘数效应, 协同提升流动性弱的其他生产要素配置状态, 规模报酬递增特性则有助于实现范围经济<sup>[28]</sup>。与此同时, 政府数据开放涵盖产品市场数据, 可为企业提供有价值的市场洞察和趋势预测, 识别出市场中尚未被满足的需求空白, 有利于企业及时把握获利机会<sup>[29]</sup>, 为企业创造更大的利润空间。

面对融资约束这一现实难题, 企业内部产生的利润及其带来的资金积累成为企业应对融资约束、扩大经营规模的重要支撑<sup>[30]</sup>。一方面, 企业利润的增加通常意味着企业的经营状况良好, 市场需求旺盛, 这促使企业扩大生产规模以满足市场需求。根据规模经济理论, 当企业生产规模扩大时, 单位成本下降, 企业能够以更高的价格决定能力将成本转化为销售价格, 从而获得更高的利润。在这种情况下, 企业为了增加产量, 往往会增加劳动力投入, 从而直接促进就业。同时, 企业利润的增加也可能提高员工的工资和福利待遇, 吸引更多的劳动力加入, 进一步提升就业水平。另一方面, 企业利润的增加可能会增强企业对新技术和新市场的投资, 从而达到生产效率提升和企业就业人数提高的双赢。例如, 在智能制造领域, 新技术的研发和应用离不开工程技术人才在产品开发、应用调试和售后维护等环节的支持, 这种需求直接催生出大量的就业岗位, 从而提升就业水平。据此, 本文提出研究假设 4。

H4: 公共数据开放通过提升企业营运效率来拉动就业。

(四) 交互效应

根据最优金融结构理论, 企业的资金筹集能力不仅受企业内部条件的影响, 也受到地区金融发展水平的影响。地区金融发展水平的高低直接影响企业融资的难度<sup>[31]</sup>。在发展水平较高的地区, 金融机构和市场的成熟度更高, 能够更有效地利用公共数据来评估企业信用和风险, 从而降低企业的融资约束, 提升融资效率。企业获得资金后, 可加快投资与扩张的进程, 由此扩大就业规模, 进而增加就业机会<sup>[15]</sup>。

相比于金融欠发达地区, 发达地区良好的金融发展环境为政府开辟出多元化的融资渠道, 并在强化政府资金筹措效能的同时, 与公共数据开放政策产生积极的交互影响。具体来看, 政府依托多元化的融资渠道所获取的资金, 可以更有效地配置于公共数据平台的构建与优化, 这不仅有助于提高公共数据的质量和可用性, 同时给金融市场主体提供了更为丰富的数据资源, 为金融领域后续开展创新活动实现效率提升注入强劲动力。总体而言, 这种交互效应既增强了政策的实施绩效, 又建立起金融发展与政府政策支持之间的良性互动, 为就业创业提供了更为有利的条件。据此, 本文提出研究假设 5。

H5: 地区金融发展水平越高, 公共数据开放对就业的促进作用越强。

三、研究设计

(一) 基准模型设定

从理论层面分析, 公共数据开放的时间、范围等主要取决于各地区的政府部门, 企业在这一过程中并不具备影响的能力。在数据层面上, 尽管上海市 (2012 年) 与北京市 (2013 年) 率先启动政府数据开放平台建设, 但区域经济相对欠发达的荆门市 (2017 年)、阳江市 (2017 年) 等城市的数据开放进程却早于天津市 (2019 年)、杭州市 (2020 年) 等经济发达地区。这一现象表明, 公共数据开放政策的实施具有外生性与随机性特征。本文拟通过建立多期双重差分模型进行实证分析, 鉴于中国城市公共数

据平台开放的非同步性特征,以各城市平台上线时间作为政策冲击时点进行动态分组,参考韩奇和杨秀云(2024)<sup>[12]</sup>的研究,在研究区间内选择企业所在城市已上线公共数据开放平台作为实验组,未上线公共数据开放平台作为对照组。基准回归模型设定如下:

$$Employee_{it} = \alpha + \beta_1 did_{it} + X'_{it}\gamma + \mu_i + v_t + \varepsilon_{it} \quad (1)$$

其中,  $Employee_{it}$  为企业  $i$  在  $t$  年的就业水平;  $did_{it}$  代表政策虚拟变量,如果企业  $i$  所在城市在  $t$  年上线公共数据开放平台则取值为 1,否则为 0。  $X'_{it}$  为控制变量向量,包含财务特征、治理特征以及地区特征;  $\mu_i$  为企业层面的固定效应;  $v_t$  为年份层面的固定效应;  $\varepsilon_{it}$  表示随机误差项,并对企业层面进行聚类调整。  $\beta_1$  是核心解释变量的回归系数,如果这一系数显著为正,说明相较于控制组企业,处理组企业的就业水平得以提高,从而证明政府公共数据开放政策能够促进就业。

## (二) 变量说明

### 1. 被解释变量

被解释变量为企业就业水平 ( $Employee$ )。根据新古典生产函数理论,劳动力作为核心生产要素,其投入规模直接决定企业生产能力。相较于其他替代性指标,如工资总额或社保缴纳人数,员工人数能更直观地反映企业实际雇佣规模。参考刘达等(2024)<sup>[15]</sup>的研究,用企业员工人数的对数来衡量企业就业水平。

### 2. 核心解释变量

核心解释变量是政策虚拟变量 ( $did$ )。地方政府建立的数据开放平台可以为公共数据开放共享提供坚实基础,并为研究公共数据开放政策创造理想的准自然实验场景<sup>[12]</sup>。本文使用  $did$  ( $treat$  与  $post$  的交互项) 来表示政策虚拟变量。其中,  $treat$  指分组虚拟变量,将企业所属城市上线了政府公共数据开放平台设为实验组,则  $treat$  取值为 1;将企业所属城市未上线政府公共数据开放平台设为对照组,则  $treat$  取值为 0。  $post$  指时间虚拟变量,若政策已实施,则  $post$  取值为 1;若政策未实施,则  $post$  取值为 0。

本文参考方锦程等(2023)<sup>[8]</sup>的研究,按如下步骤处理数据。(1) 确定标准。限定各地区公共数据开放平台的域名为 gov.cn。(2) 利用百度等搜索引擎手动搜索城市名称以及数据开放、数据资源、公共数据等关键词,并检查市级政府网站数据相关板块中公共数据开放平台的跳转链接窗口,以补充和完善检索得到的上线时间信息。(3) 依据复旦大学数字与移动治理实验室发布的中国地方公共数据开放利用报告进行数据交叉验证,确认各城市公共数据平台的首次启动上线时间。

### 3. 控制变量

参考已有文献<sup>[15]</sup>,分别设置企业层面和地区层面的控制变量。企业层面包括财务特征和治理特征,其中财务特征变量主要包括资产报酬率 ( $ROA$ )、资产规模 ( $Size$ )、资产负债率 ( $Lev$ ) 和账面市值比 ( $BM$ ),治理特征变量主要包括企业成立年限 ( $FirmAge$ )、托宾  $Q$  值 ( $TobinQ$ )。为了排除宏观经济发展水平的影响,地区层面的控制变量选取经济发展水平 ( $RGDP$ ) 和政府财政收入 ( $G$ )。

### 4. 机制变量

#### (1) 企业数字化转型

参考韩奇和杨秀云(2024)<sup>[12]</sup>的研究,通过对样本企业在样本期内的年度报告进行文本分析来评估企业的数字化转型水平 ( $Dig$ )。具体步骤包括:(1) 根据数字化转型的核心特征,创建了一个包含 67 个词语的企业数字化转型词典,该词典涵盖人工智能、大数据、云计算、区块链技术以及数字技术应用等多个方面;(2) 下载相关年份的上市公司年报,运用文本分析工具对年报文本进行分词处理,并统计上述词汇在文本中的出现频率;(3) 鉴于文本数据呈现的右偏分布特性,为了更精确地反映企业的数字化转型程度,对词汇出现的总次数进行了对数转换处理,以此作为衡量企业数字化转型的指标。

#### (2) 创新水平

现有研究通常从创新投入或创新产出角度来衡量企业创新水平<sup>[32]</sup>。本文从创新产出视角出发,参考张吉昌和龙静(2024)<sup>[10]</sup>的研究,通过汇总实用新型专利、外观设计专利和发明专利的数量来评估企业



的创新水平 (*Inno*)。

(3) 营运效率

借鉴包耀东等 (2020)<sup>[33]</sup> 的研究, 选取营业利润的对数值作为企业营运效率的代替指标 (*OP*), 数值越大, 说明企业的盈利能力越强, 企业拥有良好的发展态势。

5. 交互效应变量

参考沈洪波等 (2010)<sup>[31]</sup> 的方法, 采用金融机构存贷款余额与地区生产总值 (GDP) 的比值来衡量地区金融发展水平 (*fd*), 数值越大, 表明金融市场成熟度越高。

具体变量及说明如表 2 所示。

表 2 变量定义

| 变量类别   | 变量名称     | 变量符号            | 衡量方式                            |
|--------|----------|-----------------|---------------------------------|
| 被解释变量  | 企业就业水平   | <i>Employee</i> | 企业员工数量的自然对数                     |
| 解释变量   | 公共数据开放政策 | <i>did</i>      | 若企业所在城市已开通政府数据开放平台则取值为 1, 否则为 0 |
| 控制变量   | 资产报酬率    | <i>ROA</i>      | 息税前利润与资产总额之比                    |
|        | 资产规模     | <i>Size</i>     | 资产总额的自然对数                       |
|        | 资产负债率    | <i>Lev</i>      | 负债总额与资产总额之比                     |
|        | 账面市值比    | <i>BM</i>       | 公司股东权益与市值之比                     |
|        | 企业成立年限   | <i>FirmAge</i>  | 企业成立年限的自然对数                     |
|        | 托宾 Q 值   | <i>TobinQ</i>   | 企业市值与资产重置成本之比                   |
|        | 经济发展水平   | <i>RGDP</i>     | 人均 GDP 的自然对数                    |
|        | 政府财政收入   | <i>G</i>        | 政府财政收入的自然对数                     |
| 机制变量   | 企业数字化转型  | <i>Dig</i>      | 文本分析法                           |
|        | 创新水平     | <i>Inno</i>     | 实用新型、外观设计和发明专利数量的对数             |
|        | 营运效率     | <i>OP</i>       | 营业利润的对数值                        |
| 交互效应变量 | 地区金融发展水平 | <i>fd</i>       | 金融机构存贷款余额与 GDP 的比值              |

(三) 样本选择与数据来源

本文所使用的企业层面数据来自万得 (Wind) 金融终端、深圳希施玛数据科技有限公司 CSMAR 中国经济金融研究数据库和中国研究数据服务平台 (CNRDS)。城市层面数据主要来源于《中国城市统计年鉴》和国家统计局官网。对于个别缺失的样本数据, 结合地方统计年鉴、政府发布的统计公报或采用线性插值法进行补充。城市公共数据开放平台的相关信息, 通过手动收集并利用中国地方公共数据开放利用报告进行数据的核对和验证。

本文以 2007—2021 年沪深 A 股上市企业作为研究对象, 通过将证券代码、公司名称等信息与企业的基本信息数据、财务数据、人员结构数据进行匹配, 利用企业所在城市将城市层面变量进行匹配, 初步形成非平衡面板数据集。为保证研究结果更准确, 对合并后的数据进行如下处理: 剔除 ST、PT、\*ST 等企业样本; 剔除金融行业的企业样本; 剔除总负债超过总资产的企业样本; 剔除关键数据缺失过多的样本; 为减少极端值的影响, 对所有连续变量进行 1% 和 99% 分位点的缩尾处理。经过以上步骤, 最终得到 31 474 个观测样本。

(四) 变量描述性统计

经过数据处理后, 主要变量的描述性统计结果如表 3 所示。其中, 企业员工人数对数 (*Employee*) 的

标准差为 1.280 0，表明样本企业在就业水平方面表现出较大差异。政府公共数据开放（*did*）的均值为 0.398 4，说明接近 40% 的样本企业所在城市开通了政府数据开放平台，这与彭远怀（2023）<sup>[34]</sup> 的估算结果较为一致。控制变量结果表明，不同企业、不同地区在发展水平上存在异质性。总体来看，本文的测算结果与已有文献<sup>[35]</sup> 的统计特征较为一致。

表 3 描述性统计结果

| 变量              | 观测值    | 均值       | 标准差     | 最小值      | 最大值      |
|-----------------|--------|----------|---------|----------|----------|
| <i>Employee</i> | 31 474 | 7.665 2  | 1.280 0 | 3.555 3  | 11.180 7 |
| <i>did</i>      | 31 474 | 0.398 4  | 0.489 6 | 0        | 1        |
| <i>ROA</i>      | 31 474 | 0.058 4  | 0.069 0 | -0.364 9 | 0.332 4  |
| <i>Size</i>     | 31 474 | 22.182 3 | 1.306 5 | 19.316 7 | 26.452 3 |
| <i>Lev</i>      | 31 474 | 0.436 4  | 0.205 1 | 0.027 4  | 0.907 9  |
| <i>BM</i>       | 31 474 | 0.615 1  | 0.249 1 | 0.064 1  | 1.246 2  |
| <i>FirmAge</i>  | 31 474 | 2.854 8  | 0.358 7 | 0.693 1  | 3.610 9  |
| <i>TobinQ</i>   | 31 474 | 2.072 4  | 1.386 9 | 0.802 4  | 15.606 7 |
| <i>RGDP</i>     | 31 474 | 11.339 6 | 0.577 7 | 9.645 9  | 12.223 4 |
| <i>G</i>        | 31 474 | 15.769 6 | 1.447 1 | 12.356 6 | 18.168 6 |

四、实证结果分析

（一）基准回归

基准回归结果如表 4 所示。可以发现，无论是否控制固定效应，核心解释变量公共数据开放政策的回归系数均显著为正，表明本文选取的政策实验外生性较强，回归结果受到其他潜在因素影响的可能性较小。以上结果初步说明公共数据开放对企业就业水平具有提升作用，即 H1 得到了初步验证。

表 4 基准回归结果

| 变量             | (1)                      | (2)                     |
|----------------|--------------------------|-------------------------|
| <i>did</i>     | 0.014 9*<br>(0.007 7)    | 0.062 7***<br>(0.015 4) |
| <i>ROA</i>     | -0.012 7<br>(0.046 2)    | -0.108 1<br>(0.081 2)   |
| <i>Size</i>    | 0.647 1***<br>(0.005 4)  | 0.650 0***<br>(0.020 4) |
| <i>Lev</i>     | 0.275 1***<br>(0.023 9)  | 0.254 9***<br>(0.066 6) |
| <i>BM</i>      | -0.106 1***<br>(0.021 3) | -0.105 8**<br>(0.044 7) |
| <i>FirmAge</i> | -0.243 9***<br>(0.018 6) | -0.000 2<br>(0.085 7)   |



表4( 续)

| 变量                    | (1)                      | (2)                      |
|-----------------------|--------------------------|--------------------------|
| <i>TobinQ</i>         | 0.004 7<br>(0.003 4)     | 0.001 7<br>(0.006 3)     |
| <i>RGDP</i>           | -0.039 0**<br>(0.015 9)  | 0.002 7<br>(0.042 4)     |
| <i>G</i>              | -0.059 0***<br>(0.008 6) | -0.041 1<br>(0.033 9)    |
| 常数项                   | -4.706 5***<br>(0.129 9) | -6.034 1***<br>(0.698 4) |
| 企业层面控制变量              | 控制                       | 控制                       |
| 城市层面控制变量              | 控制                       | 控制                       |
| 企业固定效应                | 未控制                      | 控制                       |
| 年份固定效应                | 未控制                      | 控制                       |
| <i>R</i> <sup>2</sup> | 0.434 2                  | 0.440 4                  |
| 样本数                   | 31 474                   | 31 474                   |

注：\*\*\* 表示  $P<0.01$ ，\*\* 表示  $P<0.05$ ，\* 表示  $P<0.1$ ，括号内为聚类到企业层面的稳健标准误。后表同。

（二）平行趋势检验

采用多时点双重差分模型的前提是满足平行趋势假设，即处理组和控制组在政策发生前具有相似的变化趋势。只有通过平行趋势假设，基准回归的估计结果才具有可信性。鉴于各试点城市受到政策影响的时间存在差异，本文为每个试点城市创建反映政策实施相对时间的虚拟变量。参考克拉克和塔皮亚-施特（Clarke & Tapia-Schyte，2021）<sup>[36]</sup>的研究方法，采用事件研究法检验政策的动态处理效应。构建模型如下：

$$Employee_{it} = \alpha + \beta_j \sum_{j=-9}^9 did_{it}^{j=t-t_0} + X'_{it} \gamma + \mu_i + v_t + \varepsilon_{it}$$

(2)

其中， $t_0$  为政策实施年份； $j$  为相对政策实施年份前后的时间间隔，由于公共数据开放最早的年份是 2012 年，样本期最后一年是 2021 年， $j$  的最大取值为 9，最小取值为-14。为避免多重共线性，将政府公共数据开放政策实施前一年作为基期，同时为保证政策实施前后研究样本的时间对称性，将政策实施前 9 年的数据汇总到第-9 期，因此  $\beta_j$  捕捉到了政策前 9 年到政策后 9 年的动态效应。其余变量定义与式（1）相同。

根据图 1 可以发现，在公共数据开放政策实施之前，各年份虚拟变量的回归系数均不显著，说明处理组企业和控制组企业在公共数据开放政策实施之前无明显差异，符合平行趋势假设。在政策的动态效应方面，随着时间推移，回归系数有增大的趋势，说明公共数据开放政策对就业的提升效应具有持续性，促就业效应随时间的推移逐渐增强。可能的解释是，随着公共数据开放政策的深入实施，数据资源转化为生产要素的过程更加高效，充分发挥了公共数据蕴含的经济价值，从而提升企业就业水平<sup>[37]</sup>。

（三）内生性处理

尽管在基准回归中控制了企业固定效应、年份固定效应以及企业和城市层面的相关特征变量，但仍然存在忽略相关变量的可能，从而导致内生性问题，本文通过工具变量法进行内生性处理。根据本文的研究主题，一方面，固定电话的使用和普及为后来的通信技术发展奠定了基础，促进了数据通信技术的进步，为政

府公共数据开放提供技术支持，符合相关性条件；另一方面，早期的固定电话作为社会基础设施，并不直接影响企业的就业规模，符合外生性条件。参考宋建和王怡静（2024）<sup>[38]</sup>的研究方法，构造企业所在城市 1984 年每百人固定电话数量与前一年全国互联网上网人数（时间序列变量）的交互项（IV）作为工具变量。通过两阶段最小二乘（2SLS）回归后，第一阶段工具变量的回归系数在 1% 的水平下显著为正，且 F 统计量为 143.60，大于 10% 水平下的临界值 16.38，排除弱工具变量的可能；P 值为 0，说明不存在识别不足的问题<sup>[39]</sup>。第二阶段回归结果与前文基准回归结果一致<sup>①</sup>，表明在考虑内生性问题后，先前的主要研究结论仍具有可信性。

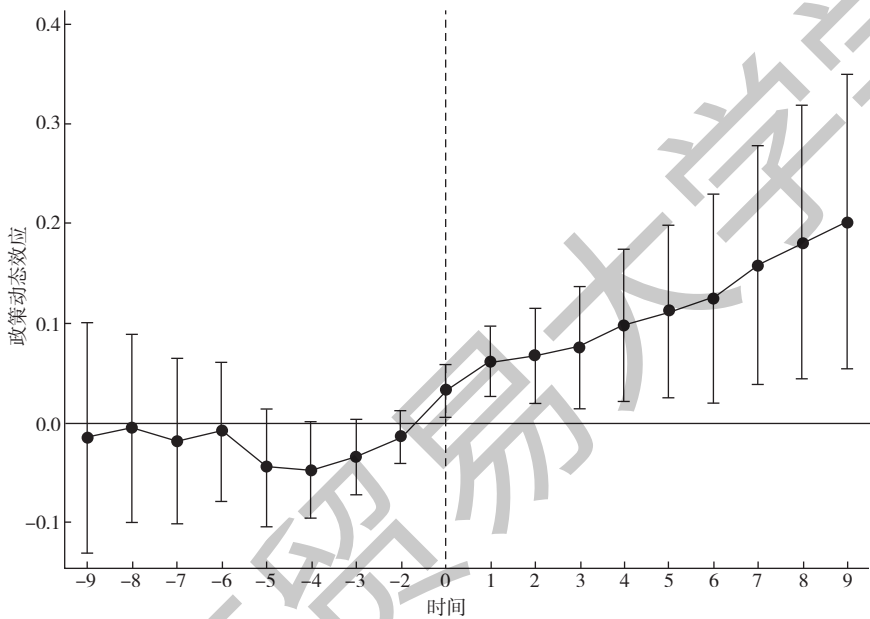


图 1 平行趋势检验结果

（四）稳健性检验

1. 安慰剂检验

尽管基准回归模型中已经加入大量控制变量，但仍存在忽略某些因素的可能性，从而导致公共数据开放促进就业这一结论有偏，因此借鉴韩奇和杨秀云（2024）<sup>[12]</sup>的研究方法，通过随机选取政府数据开放的城市和年份进行安慰剂检验，再利用式（1）重新回归，并将整个过程重复 1 000 次。由图 2 可知，随机处理后得到的估计系数大多围绕 0 值分布，近似正态分布的形状，而实际估计系数为 0.062 7，且在 1% 的水平下显著，与随机模拟结果有明显区别，表明安慰剂检验顺利通过，即核心结论具有稳健性。

2. 其他稳健性检验

本文以倾向得分匹配-双重差分（PSM-DID）、样本敏感性、变量滞后一期、排除其他政策干扰等方式进行回归分析。限于篇幅，下文简要报告稳健性检验结果<sup>②</sup>。

（1）PSM-DID 检验

为了缓解组间样本特征不可比引发的选择性偏差问题对研究结果的影响，本文采用多时点 PSM-DID 方法进一步进行稳健性检验。对样本进行 1：2 最近邻匹配，回归结果显示，公共数据开放政策的回归系数依然显著为正，与基准回归结果相一致，证明基本研究结论具有稳健性。

① 限于篇幅，省略具体回归结果，留存备索。  
② 限于篇幅，省略具体回归结果，留存备索。

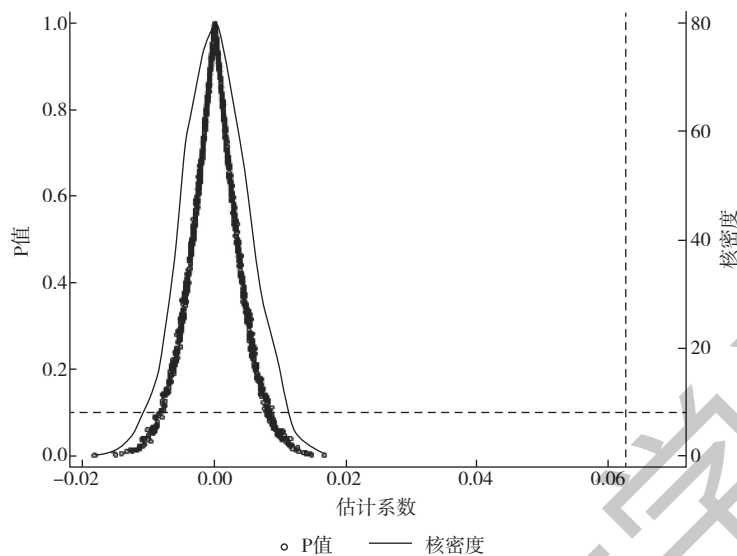


图 2 安慰剂检验结果

(2) 样本敏感性检验

考虑中国的直辖市在经济及就业方面具有特殊性,参考谭伟杰和庞钰标 (2023)<sup>[19]</sup>的做法,在回归分析中排除了属于北京市、上海市、天津市、重庆市 4 个直辖市的企业样本。结果显示,解释变量的回归系数依然显著为正,与基准回归的结果一致。

(3) 变量滞后一期

现实中政府公布的数据往往是过去的数 据,这可能导致公共数据开放对企业的经营活动及就业水平的影响存在一定的延迟。参考杨秀云和韩奇 (2023)<sup>[14]</sup>的做法,对被解释变量进行滞后一期处理后再次进行回归分析。结果显示,核心解释变量的回归系数在 1% 的水平下显著为正,再次证明基准回归结果的稳健性。

(4) 改变样本区间

为减少较长样本区间可能带来的政策干扰效应,采用 2010—2021 年的数据重新进行回归分析。实证结果表明,在调整样本区间后,政策变量的回归系数仍显著为正,且大小与基准回归结果相近,证实了基准回归结果的可靠性。

(5) 排除其他政策的干扰

在研究区间内,2009 年开始设立的国家自主创新示范区、2011 年起科技部开展的科技和金融结合试点政策、2012 年开始实施的国家智慧城市试点政策和知识产权试点政策以及 2014 年以来开展的电子商务示范政策与本文涉及的政策密切相关。为排除这些政策的干扰,借鉴陈艳利和蒋琪 (2024)<sup>[24]</sup>的研究方法,将这些政策的虚拟变量加入基准模型中。结果显示,公共数据开放政策的回归系数仍然显著为正,在一定程度上说明在排除了这些政策可能对结果产生的干扰后,公共数据开放政策仍然能够促进企业就业,证明基本研究结论足够稳健可信。

(五) 异质性分析

异质性分析的目的在于解答以下疑问:公共数据开放对企业就业水平的促进作用是否会因企业所有制性质以及企业所在行业的不同而有明显区别?是否会受到企业所在城市地理位置的影响?对此,本部分将从企业、行业和城市三个层面的视角分析公共数据开放对企业就业的影响差异。

1. 企业所有制性质异质性

依据 CSMAR 数据库中企业股权性质的不同,将样本企业划分为国有企业和非国有企业两类,分组进



行回归分析。表 5 结果显示，解释变量的回归系数均显著为正，表明公共数据开放对国有企业和非国有企业就业都有积极影响，但国有企业解释变量的回归系数更大，说明公共数据开放对国有企业的促进作用更强。两组系数差异的可能原因是：一方面，公共数据开放主要由各级政府主导实施，国有企业作为保障政策贯彻落实的主要阵地，与政府的联系更为紧密，更有动力去适应公共数据开放的需要<sup>[10]</sup>；另一方面，国有企业拥有更多资源优势，且在稳就业方面承担着更大的社会责任，因此促进作用更明显<sup>[40]</sup>。

2. 行业异质性

按照新注册企业类型将样本划分为制造业与服务业，进一步探讨公共数据开放对不同行业类型的就业效应。由表 5 可知，核心解释变量的回归系数都在 5% 的水平下显著为正，且对服务业的促进作用更加明显。一方面，制造业的生产要素主要以土地、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；另一方面，相比制造业，服务业具有更强的劳动密集特性和更高的灵活性，能够更快速地将公共数据开放资源转化为新的服务模式和业务形态，进而带动就业增长<sup>[15]</sup>。因此，服务业受到该政策的冲击更强。

3. 城市区位异质性

中国不同地区的资源分配存在差异，导致各区域在经济基础、政策制度环境以及产业结构等方面展现出独特的地域性特征。本文将企业所在城市划分为东部、中部和西部。由表 5 可知，东部地区的回归系数在 1% 的水平下显著为正，中部地区的回归系数为负且不显著，西部地区的回归系数不显著。可能的解释是东部地区在经济发展、人才技术、创新资源、数据基础设施等方面相对中部和西部地区具有明显优势<sup>[10]</sup>，能够更有效地将数据资源的价值转化为经济效益和就业机会。对中部和西部地区而言，在技术应用能力、数据整合利用水平以及相关配套政策完善度等方面的不足，导致公共数据开放政策难以有效转化为就业增长动力，甚至可能因企业短期数据管理成本上升而出现负向影响。

表 5 异质性分析结果

| 变量             | 企业所有制性质                  |                          | 行业                       |                          | 城市区位                     |                          |                           |
|----------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|---------------------------|
|                | 国有企业                     | 非国有企业                    | 制造业                      | 服务业                      | 东部地区                     | 中部地区                     | 西部地区                      |
| <i>did</i>     | 0.088 7***<br>(0.026 7)  | 0.038 7**<br>(0.018 2)   | 0.033 7**<br>(0.014 7)   | 0.067 4**<br>(0.027 4)   | 0.060 2***<br>(0.018 7)  | -0.061 6<br>(0.039 3)    | 0.022 0<br>(0.038 0)      |
| 常数项            | -5.209 8***<br>(1.439 2) | -6.393 2***<br>(0.867 7) | -6.982 4***<br>(0.582 8) | -3.926 2***<br>(1.314 3) | -3.832 8***<br>(0.810 2) | -5.491 6***<br>(1.513 8) | -11.777 0***<br>(1.242 2) |
| 企业层面控制变量       | 控制                       | 控制                       | 控制                       | 控制                       | 控制                       | 控制                       | 控制                        |
| 城市层面控制变量       | 控制                       | 控制                       | 控制                       | 控制                       | 控制                       | 控制                       | 控制                        |
| 企业固定效应         | 控制                       | 控制                       | 控制                       | 控制                       | 控制                       | 控制                       | 控制                        |
| 年份固定效应         | 控制                       | 控制                       | 控制                       | 控制                       | 控制                       | 控制                       | 控制                        |
| 样本数            | 12 990                   | 17 932                   | 19 622                   | 11 475                   | 21 784                   | 4 414                    | 5 276                     |
| R <sup>2</sup> | 0.368 5                  | 0.497 1                  | 0.503 5                  | 0.395 4                  | 0.422 7                  | 0.419 0                  | 0.569 9                   |
| 组间系数差异检验       | 0.050 0**<br>[0.026 0]   |                          | 0.034 0*<br>[0.067 0]    |                          |                          |                          |                           |

注：采用费舍尔组合检验方法抽样 1 000 次，方括号内数字为组间系数差异检验 P 值，P 值小于 0.1 说明结果显著。

## 五、机制检验

依据前文的实证检验,公共数据开放政策对企业就业水平的促进作用已被证实,但其中具体的作用路径仍未被检验。参考黄先海和虞柳明(2023)<sup>[5]</sup>的研究方法,构建如下计量模型:

$$M_{it} = \alpha + \beta_1 did_{it} + X'_{it}\gamma + \mu_i + v_t + \varepsilon_{it} \quad (3)$$

其中, $M_{it}$ 表示机制变量;根据前文的理论分析与变量说明, $M_{it}$ 分别代表企业数字化转型、创新水平和营运效率。

### 1. 促进企业数字化转型

通过前文理论部分分析政府公共数据开放的政策效应,本文预期该举措不仅可强化企业实施数字化转型的技术能力,而且有助于企业形成积极的转型态度,为企业数字化转型提供持续驱动力<sup>[35]</sup>。表6结果显示,核心解释变量的回归系数在1%的水平下显著为正,表明公共数据开放政策有利于促进企业数字化转型。企业数字化转型对就业的影响呈现“创造性破坏”现象,造成此现象的原因一方面是技术替代效应,数字技术凭借效率优势对程序化岗位产生替代,使这类岗位劳动力需求锐减,引发局部就业规模收缩,这是产业结构升级与数字化转型中生产要素重新配置带来的结果;另一方面是创造效应,数字化转型通过降本增效激发市场需求,促使产业规模扩大,企业增加劳动力雇佣,在这一过程中,就业创造效应占主导地位<sup>[23]</sup>。综上,H2得到了证实。

### 2. 推动企业创新

根据前文理论分析,公共数据作为关键性生产要素和战略资源,能够提升企业的创新水平<sup>[10]</sup>,通过促进产品创新和流程创新,最终带动企业用工规模的扩大<sup>[26]</sup>。表6结果显示,核心解释变量的回归系数同样显著为正,说明公共数据开放政策有利于企业创新水平的提升。在一定程度上,企业获得更多的公共开放数据象征其拥有更多的新型生产要素,通过与传统生产要素相结合,能够充分释放数据要素的潜在价值,推动提升企业的创新能力,并通过生产率渠道和市场需求规模变化渠道来影响劳动就业,其中创造效应大于破坏效应<sup>[26]</sup>。综上,H3得到了验证。

### 3. 提升企业营运效率

前文理论分析部分表明,公共数据通过降低企业数据获取成本、提升管理效能和增强市场洞察力<sup>[27]</sup>,助力企业经营规模的扩大,从而容纳更多劳动力<sup>[30]</sup>。表6结果显示,公共数据开放政策能够提升企业营业利润。企业利润增长通常代表经营态势良好、市场需求强劲,正向推动企业扩大生产以满足市场需求。此时,企业为增产会加大劳动力投入,直接推动就业。此外,利润增长刺激企业增加对新技术、新市场的投资,进而实现效率与就业的双赢。综上,H4得到了验证。

## 六、交互效应检验

为验证地区金融发展水平与政府公共数据开放政策影响企业就业的交互效应,参考张吉昌和龙静(2024)<sup>[10]</sup>的研究方法,建立如下分析模型:

$$Employee_{it} = \alpha + \beta_1 did_{it} + \beta_2 did_{it} \times fd_{it} + \beta_3 fd_{it} + X'_{it}\gamma + \mu_i + v_t + \varepsilon_{it} \quad (4)$$

其中, $fd_{it}$ 代表地区金融发展水平, $did_{it} \times fd_{it}$ 是交互效应变量;若 $\beta_2$ 大于0,说明地区金融发展水平增强了公共数据开放政策对企业就业的正向影响。

表6结果显示,核心解释变量与地区金融发展水平的交互项的回归系数为0.0284,且在1%的水平下显著,说明地区金融发展水平越高,公共数据开放对就业的促进作用越强。根据前文理论分析,企业融资能力同时取决于内部条件和区域金融环境两个维度<sup>[31]</sup>。通常情况下,金融发展水平高意味着金融市场的活跃度高、金融工具丰富多样、投资者信心充足。在这样的环境下,企业更易获取资金支持,能够缓解其面临的融资约束,进而促进资本投入和规模扩张,最终带动就业增长<sup>[15]</sup>;发达的金融发展水平对政府筹资能力具有积极的提升作用,为公共数据开放政策注入更强劲的资金支持,更有利于发挥公共数据

开放政策对就业的促进作用。综上，H5 得到了验证。

表 6 机制检验、交互效应检验结果

| 变量                     | 数字化转型                     | 创新水平                    | 营运效率                      | 企业就业水平                    |
|------------------------|---------------------------|-------------------------|---------------------------|---------------------------|
| <i>did</i>             | 0.095 7 ***<br>(0.022 6)  | 0.053 4 **<br>(0.024 1) | 0.036 2 **<br>(0.015 7)   |                           |
| <i>did</i> × <i>fd</i> |                           |                         |                           | 0.028 4 ***<br>(0.009 4)  |
| 常数项                    | -5.575 8 ***<br>(0.811 6) | -1.198 6<br>(1.247 8)   | -3.483 5 ***<br>(0.582 9) | -5.948 7 ***<br>(0.705 9) |
| 企业层面控制变量               | 控制                        | 控制                      | 控制                        | 控制                        |
| 城市层面控制变量               | 控制                        | 控制                      | 控制                        | 控制                        |
| 企业固定效应                 | 控制                        | 控制                      | 控制                        | 控制                        |
| 年份固定效应                 | 控制                        | 控制                      | 控制                        | 控制                        |
| R <sup>2</sup>         | 0.385 0                   | 0.039 5                 | 0.677 9                   | 0.441 3                   |
| 样本数                    | 31 332                    | 14 079                  | 27 180                    | 31 474                    |

七、拓展分析

为进一步探究公共数据开放的经济后果，本部分从就业技能和内部薪酬差距两个方面进行分析。参考李旭超等（2024）<sup>[30]</sup> 的变量衡量方法，按照岗位类型将财务人员与技术人员划分为高技能劳动力，将销售人员和生产人员划分为低技能劳动力，将其他行政人员归类为中等技能劳动力；通过比较管理层的平均薪酬和员工的平均薪酬来衡量企业内部的薪酬差距<sup>①</sup>。

将被解释变量依次更换为不同技能水平的员工人数和企业内部薪酬差距并进行回归分析。由表 7 可知，公共数据开放对于高技能和低技能劳动力的就业有推动作用，对于中等技能劳动力的就业则表现出抑制作用，与现有文献得出的结论一致<sup>[16-17]</sup>。造成这一结果的原因可能在于：一方面，高技能劳动力能够更好地与数据要素相结合，发挥其规模报酬递增的特性<sup>[41]</sup>；另一方面，如前文所述，公共数据开放会催生大量灵活性和随机性的岗位，这些岗位的人力资本专有性较弱，更易与低技能劳动力匹配，而数据要素的零成本特点促进了员工共享远程办公，企业为节约成本，可能会减少对行政类人员的雇佣<sup>[38]</sup>。表 7 最后一列核心解释变量的回归系数显著为正，说明公共数据开放加剧了企业内部薪酬差距。一般而言，企业设定的薪酬标准与员工的资质和能力相匹配，管理层在数据要素的利用、经营管理与决策方面比普通员工更有优势，因此其薪酬涨幅高于普通员工<sup>[42]</sup>，加剧了内部收入的不平等。

表 7 拓展分析结果

| 变量         | 高技能                       | 低技能                       | 中等技能                      | 内部薪酬差距                     |
|------------|---------------------------|---------------------------|---------------------------|----------------------------|
| <i>did</i> | 0.021 6 *<br>(0.012 2)    | 0.035 9 **<br>(0.014 3)   | -0.003 2<br>(0.016 0)     | 0.243 7 **<br>(0.113 3)    |
| 常数项        | -9.586 2 ***<br>(0.649 4) | -6.519 7 ***<br>(0.850 1) | -8.350 7 ***<br>(0.834 7) | -11.414 7 ***<br>(4.134 5) |

① 管理层的平均薪酬=(董事、监事和高级管理人员的年薪总和)/管理层总人数，员工平均薪酬=(应付职工薪酬总额的变动额+支付给职工以及为职工支付的现金-董事、监事和高级管理人员的年薪总和)/员工人数。

表7(续)

| 变量             | 高技能     | 低技能     | 中等技能    | 内部薪酬差距  |
|----------------|---------|---------|---------|---------|
| 企业层面控制变量       | 控制      | 控制      | 控制      | 控制      |
| 城市层面控制变量       | 控制      | 控制      | 控制      | 控制      |
| 企业固定效应         | 控制      | 控制      | 控制      | 控制      |
| 年份固定效应         | 控制      | 控制      | 控制      | 控制      |
| R <sup>2</sup> | 0.432 3 | 0.351 5 | 0.245 1 | 0.039 6 |
| 样本数            | 24 025  | 21 835  | 26 542  | 24 023  |

八、研究结论与政策建议

有条不紊地释放数据要素是发挥其赋能效应的先决条件，而公共数据开放政策是促进市场主体实现数据要素价值转换、激活国内经济活力的关键决策。本文将各城市上线公共数据开放平台视为一项准自然实验，以 2007—2021 年中国 A 股上市公司为研究对象，采用多期双重差分法考察了公共数据开放政策对企业就业水平的影响及其作用渠道。研究发现，该项政策能够提升企业就业水平，且该结论在经过一系列稳健性检验之后仍然成立。机制分析结果表明，公共数据开放政策通过促进企业数字化转型、提升企业创新水平和营运效率，进而促进就业，且地区的金融发展水平越高，促进作用越强。异质性分析发现，公共数据开放对国有企业、服务业和东部地区的就业促进作用更明显。进一步分析结果表明，公共数据开放促进高技能劳动力和低技能劳动力的就业，而对中等技能劳动力的就业产生负向作用。同时扩大了企业内部管理层与员工之间的薪酬差距。

基于以上研究结论，本文提出如下建议。

第一，要积极推进未开放城市上线数据开放平台，不断扩大政策的覆盖范围。在此基础上，着力推进就业服务领域的深层次开放，通过建立“纵贯省市县、横跨各部门”的智慧就业服务平台，形成基于数据的精准服务体系，最大限度地发挥数据要素赋能就业的作用。

第二，政府需统筹兼顾。一方面，激励企业顺应数字经济的发展趋势，加快推进数字化转型，增强其数据处理和应用的技能。同时，通过不断完善创新激励机制，如专利申请奖励、研发费用加计扣除等，促进企业创新水平的提升。另一方面，必须兼顾地区基础设施的建设，同时针对不同企业和区域的特点，实施差异化政策。

第三，考虑数据要素的开放是把“双刃剑”，在赋能企业发展、促进就业的同时，也伴随着内部薪酬差距的加剧以及岗位替代的风险，政府一方面可依托公共数据平台，建立员工薪酬满意度匿名反馈渠道，赋予员工发声权，促进企业内部协商机制的完善。另一方面，不断开放更多教育培训数据资源，提升劳动者的技能水平，以便更快地适应快速变化的就业环境，促进稳就业目标的顺利实现。

参考文献：

[1]段学慧,张娜.数据要素及其形成新质生产力的机理研究[J].经济纵横,2024(7):18-28.

[2]复旦大学数字与移动治理实验室.中国地方公共数据开放利用报告——城市(2023年度)[R/OL].(2023-11-01)[2025-04-03].  
<http://ifopendata.fudan.edu.cn/report>.

[3]王法硕.我国地方政府数据开放绩效的影响因素——基于定性比较分析的研究[J].情报理论与实践,2019,42(8):38-43.

[4]吴应强,李白杨,石乐怡.政府数据开放平台效率评估及影响因素实证研究——基于无导向SBM-Malmquist-Tobit模型[J].现代情报,2024,44(7):74-83.

[5]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[J].浙江学刊,2023(4):88-98.

[6]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.



- [7] 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[J]. 数量经济技术经济研究, 2021, 38(3): 64-83.
- [8] 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.
- [9] 张莉, 林安然. 中国地方政府数据开放与经济发展质量[J]. 公共行政评论, 2023, 16(2): 120-137.
- [10] 张吉昌, 龙静. 政府数据开放何以赋能企业创新? [J]. 现代经济探讨, 2024(4): 29-43.
- [11] JIA J, JIN G Z, WAGMAN L. The short-run effects of the general data protection regulation on technology venture investment[J]. *Marketing Science*, 2021, 40(4): 661-684.
- [12] 韩奇, 杨秀云. 公共数据开放能否促进企业数字化转型? [J]. 现代经济探讨, 2024(4): 44-59.
- [13] CARRIERE-SWALLOW Y, HAKSAR V. The economics and implications of data: an integrated perspective[Z]. *International Monetary Fund Departmental Paper No. 19/16*, 2019.
- [14] 杨秀云, 韩奇. 公共数据开放能提升企业全要素生产率吗? [J]. 证券市场导报, 2023(12): 18-30.
- [15] 刘达, 王晓丹, 石玉堂. 数据要素集聚对企业就业的影响——基于就业规模与就业结构的双重视角[J]. 山西财经大学学报, 2024, 46(4): 54-67.
- [16] 王淑瑶, 汤吉军, 刘达, 等. 政府数据开放对劳动力就业的影响研究——基于就业规模和就业结构的双重视角[J]. 华东经济管理, 2024, 38(9): 116-128.
- [17] 王春蕊, 李恩极, 彭兰凌, 等. 公共数据开放的就业效应[J]. 人口研究, 2024, 48(6): 3-17.
- [18] 李成明, 王霄, 李博. 城市智能化、居民劳动供给与包容性就业——来自准自然实验的证据[J]. 经济与管理研究, 2023, 44(3): 41-59.
- [19] 谭伟杰, 庞钰标. 企业数字化转型的稳就业效应[J]. 南方金融, 2023(11): 19-34.
- [20] 张晨, 万相昱, 姜智超, 等. 开放政府数据的经济增长效应研究[J]. 中国软科学, 2023(2): 1-11.
- [21] 王永钦, 董雯. 机器人的兴起如何影响中国劳动力市场? ——来自制造业上市公司的证据[J]. 经济研究, 2020, 55(10): 159-175.
- [22] ACEMOGLU D, RESTREPO P. Robots and jobs: evidence from US labor markets[J]. *Journal of Political Economy*, 2020, 128(6): 2188-2244.
- [23] 李磊, 王小霞, 包群. 机器人的就业效应: 机制与中国经验[J]. 管理世界, 2021, 37(9): 104-119.
- [24] 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.
- [25] GIULIODORI D, STUCCHI R. Innovation and job creation in a dual labor market: evidence from Spain[J]. *Economics of Innovation and New Technology*, 2012, 21(8): 801-813.
- [26] LICHT G, PETERS B. Do green innovations stimulate employment? —Firm-level evidence from Germany[Z]. *Working Paper No. 53*, 2014.
- [27] 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.
- [28] MORTON F S, BOUVIER P, EZRACHI A, et al. Committee for the study of digital platforms: market structure and antitrust subcommittee[R]. Chicago, IL: The University of Chicago Booth School of Business, 2019.
- [29] MAGALHAES G, ROSEIRA C. Open government data and the private sector: an empirical view on business models and value creation[J]. *Government Information Quarterly*, 2020, 37(3): 101248.
- [30] 李旭超, 张文怡, 赵婧. 企业数字化转型、人力资本升级与高质量就业[J]. 北京工商大学学报(社会科学版), 2024, 39(3): 40-52.
- [31] 沈红波, 寇宏, 张川. 金融发展、融资约束与企业投资的实证研究[J]. 中国工业经济, 2010(6): 55-64.
- [32] 徐悦, 刘运国, 蔡贵龙. 高管薪酬粘性与企业创新[J]. 会计研究, 2018(7): 43-49.
- [33] 包耀东, 高玉梅, 朱学义. 互联网对制造企业经营绩效的影响研究——以江苏上市制造企业为例[J]. 会计之友, 2020(11): 52-58.
- [34] 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.
- [35] 李佳林, 王禹, 赵西卜. 政府公共数据开放对企业数字化转型的影响研究[J]. 管理学报, 2024, 21(10): 1467-1475.
- [36] CLARKE D, TAPIA-SCHYTHE K. Implementing in the panel event study[J]. *Stata Journal*, 2021, 21(4): 853-884.
- [37] 胡金焱, 于露, 刘茵伟. 数字政府建设提升了企业绩效吗? ——基于数据要素开放视角[J]. 东岳论丛, 2024, 45(6): 109-122.
- [38] 宋建, 王怡静. 企业数字化转型是就业机遇还是替代危机: 来自中国上市公司文本分析的证据[J]. 中国软科学, 2024(4): 131-143.
- [39] 张顺, 吕风光. 企业数字化对劳动力市场就业风险的影响[J]. 当代经济科学, 2024, 46(4): 14-26.
- [40] 沈国兵, 袁征宇. 企业互联网化对中国企业创新及出口的影响[J]. 经济研究, 2020, 55(1): 33-48.
- [41] BUGHIN J, SEONG J, MANYIKA J, et al. Notes from the AI frontier: modeling the impact of AI on the world economy[R]. Chicago, IL: McKinsey Global Institute, 2018.
- [42] 高文书, 万诗婕. 企业数字化转型、人力资本与内部薪酬差距[J]. 北京工商大学学报(社会科学版), 2024, 39(2): 31-44.

# The Employment Effects of Public Data Openness Policy

HAO Nan, ZHANG Xinyi  
(Anhui University, Hefei 230601)

**Abstract:** Employment is one of the core issues in China, and how to effectively promote enterprise employment at the policy level is an important topic worthy of attention. Existing literature on public data openness mainly focuses on qualitative studies such as the analysis of the current situation, impact factors, and evaluation of the effect. However, few studies have utilized the public data openness policy to carry out empirical analyses, indicating that it is in the initial stage. Therefore, this paper regards the local government-led public data platform openness as an external policy shock and uses the multi-period difference-in-differences (DID) method to explore the impact of government data openness on enterprise employment based on the data of A-share listed enterprises from 2007 to 2021.

The findings reveal that the public data openness policy can increase enterprise employment. The mechanism test shows that this policy promotes employment by facilitating the digital transformation of enterprises and improving innovation and operational efficiency of enterprises. Moreover, the higher the level of financial development, the more significant the promotion effect is. The heterogeneity analysis indicates that the employment promotion effect is stronger for state-owned enterprises (SOEs), the service industry, and the eastern region compared to non-SOEs, the manufacturing, and the central and western regions. Further analysis demonstrates that the public data openness policy produces an employment promotion effect on high- and low-skilled labor, suppresses the employment level of middle-skilled labor, and widens the wage gap within enterprises.

Based on the findings, this paper proposes the following policy implications. First, it is necessary to encourage cities to implement open data platforms, thus maximizing the role of data elements in empowering employment. Second, the government needs to pay more attention to the digitalization of enterprises, and must take into account the construction of regional infrastructure, smoothing out the mechanism and channels for public data openness policy to promote enterprise employment. Meanwhile, it should formulate differentiated policies for different enterprises and regions. Finally, the government can strengthen the supervision of the enterprise salary system to ensure the fairness and reasonableness of salary distribution. Additionally, it can open up more education and training data resources to upgrade the skills of workers, so that they can adapt more quickly to the rapidly changing employment environment, thus facilitating the realization of the goals of “stable employment” and fair and reasonable income distribution.

The potential marginal contributions of this paper lie in the following aspects: providing a new perspective on the study of employment-related factors, refining research on the social effects of public data openness policies, elucidating the heterogeneous characteristics of the impact of public data openness policies, and empowering precise policy-making.

**Keywords:** public data openness; data element; employment level; skill structure; wage gap

(责任编辑: 蒋 琰)