

DOI:10.3969/j.issn.1674-8131.2023.02.002

大数据发展对企业数字技术创新的影响研究

——基于国家大数据综合试验区的准自然实验

戴艳娟,沈伟鹏,谭伟杰✉

(广东外语外贸大学 经济贸易学院,广东 广州 510006)

摘要:数字技术的发展和应用使数据成为数字经济的核心生产要素,而大数据的发展又会推动数字技术的进一步发展。将国家大数据综合试验区的设立作为一项准自然实验,用数字技术专利来衡量企业数字技术创新水平,采用多期双重差分模型对 2007—2020 年沪深两市 A 股上市公司的分析发现:设立国家大数据综合试验区显著促进了试验区内企业数字技术创新水平的提升,该结论在控制宏观经济因素及其他政策干扰、替换被解释变量、安慰剂检验、PSM-DID 检验、工具变量法等稳健性检验中均成立;设立国家大数据综合试验区对试验区内企业产生了显著的信息改善效应、治理改进效应和资源获取效应,从而有效促进了企业数字技术创新;设立国家大数据综合试验区对风险承担能力较强企业、非衰退期企业、与 ICT 关联度较高行业企业的数字技术创新促进作用较大。各地区应积极推动大数据发展,并充分发挥大数据综合试验区的示范引领作用与辐射带动作用,形成以数字技术进步推动大数据发展、以大数据发展促进数字技术创新的良性循环。

关键词:数据要素;大数据综合试验区;数字技术创新;数字经济;数字基础设施;大数据产业

中图分类号:F273.1;F127 **文献标志码:**A **文章编号:**1674-8131(2023)02-0016-13

引用格式:戴艳娟,沈伟鹏,谭伟杰. 大数据发展对企业数字技术创新的影响研究——基于国家大数据综合试验区的准自然实验[J]. 西部论坛,2023,33(2):16-28.

DAI Yan-juan, SHEN Wei-peng, TAN Wei-jie. Research on the impact of big-data development on enterprise digital technology innovation: quasi-natural experiment based on the national big data comprehensive pilot zone [J]. West Forum, 2023, 33(2):16-28.

* 收稿日期:2023-01-15;修回日期:2023-03-02

基金项目:国家自然科学基金(71873021)

作者简介:戴艳娟(1971),女,江苏金坛人;教授,博士,硕士生导师,主要从事投入产出分析和产业经济研究。沈伟鹏(1998),男,广东梅州人;硕士研究生,主要从事数字经济和投入产出研究。谭伟杰(1998),通信作者,男,广东江门人;硕士研究生,主要从事数字化发展研究,Email: 912566544@qq.com。

一、引言

在数字经济时代,数据要素成为新的核心生产要素,数字技术创新则成为驱动经济社会发展的新动能。企业是市场经济的微观主体,也是数字技术创新最重要的行为主体,因而促进企业数字技术创新能力和水平提升是推动数字经济发展的根本路径。然而,目前相关文献大多侧重于研究如何实现企业的数字化转型,而对如何促进企业数字技术创新的探究相对不足。数字化转型与数字技术创新具有不同的内涵和驱动因素,前者强调企业应用数字技术来创造新产品、改进生产过程、变革组织模式以及创新商业模式等(刘洋等,2020)^[1],而后者是指数字技术本身的新突破和新变化。可以认为,企业的数字技术创新是其数字化转型的技术支撑,决定着数字化转型的程度和效果。数字技术创新项目往往具有高投入、高风险、长周期等特点,导致企业开展数字技术创新的意愿和行为受到较多约束,同时,由于相关信息的庞杂笼统导致资本市场难以对企业不同的数字技术创新行为产生有针对性的激励,从而需要制定相关政策措施来改善企业数字技术创新的环境,以更有效地激励企业进行数字技术创新。因此,有必要深入研究影响企业数字技术创新的各种因素及其作用机制。

关于影响企业技术创新的因素已有大量研究。虽然数字技术创新是技术创新中的一种,影响技术创新的因素同样也会影响数字技术创新,但数字技术创新的特殊性使其可能受到特殊因素的影响,探究这些特殊因素的影响将更有利于科学把握数字技术创新的特殊规律,进而更有效地促进企业数字技术创新。然而,已有文献大多关注企业数字技术创新对企业高质量发展及全要素生产率等的影响(谢谦等,2022;罗佳等,2023;黄勃等,2023)^[2-4],鲜有研究对影响企业数字技术创新的因素进行专门探讨,更缺乏相关经验证据。对此,本文基于数据要素本身在数字技术创新中的特殊作用,试图分析地区大数据发展对企业数字技术创新的影响。2015年8月国务院印发的《促进大数据发展行动纲要》(国发[2015]50号,以下简称《纲要》)对“大数据”的定义是:以容量大、类型多、存取速度快、应用价值高为主要特征的数据集合,正快速发展为对数量巨大、来源分散、格式多样的数据进行采集、存储和关联分析,从中发现新知识、创造新价值、提升新能力的新一代信息技术和服务业态。目前,关于大数据发展包括哪些具体内容以及如何评价大数据发展水平还未形成统一认识,但从《纲要》提出的三大主要任务和10个重点工程来看^①,大数据发展至少包括数据要素市场和大数据(数字)基础设施的建设和完善以及大数据产业的集聚和发展等多方面的内容,而相关统计数据的缺乏也导致关于大数据发展的实证研究不足。

在《纲要》中提到,要“推进贵州等大数据综合试验区建设”。2016年2月,国家发展改革委、工业和信息化部、中央网信办发函批复,同意建设国家大数据(贵州)综合试验区,贵州省成为首个国家大数据综合试验区;同年10月,第二批国家大数据综合试验区获批建设(包括京津冀、珠江三角洲、上海、河南、重庆、沈阳和内蒙古)。国家大数据综合试验区将在大数据制度创新、公共数据开放共享、大数据创新应用、大数据产业聚集、大数据要素流通、数据中心整合利用、大数据国际交流合作等方面进行试验探索,并通过示范效应带动全国的大数据发展。因此,在相关统计数据还不完善的情况下,国家大数据综合试验区的设立使我们可以通过分析其政策效应来探究大数据发展的各种经济社会效应。现有文献也对此进行了一些探讨,其中,在企业发展层面主要分析了国家大数据综合试验区设立对企业“脱虚向实”、创

^① 主要任务有:加快政府数据开放共享,推动资源整合,提升治理能力;推动产业创新发展,培育新兴业态,助力经济转型;强化安全保障,提高管理水平,促进健康发展。重点工程包括:政府数据资源共享开放工程、国家大数据资源统筹发展工程、政府治理大数据工程、公共服务大数据工程、工业和新兴产业大数据工程、现代农业大数据工程、万众创新大数据工程、大数据关键技术及产品研发与产业化工程、大数据产业支撑能力提升工程、网络和大数据安全保障工程。

新及数字化转型等的影响(孙哲远,2022;陈文等,2022;王晓红等,2022;侯林岐等,2022)^[5-8],而对企业数字技术创新的影响还未论及。鉴于此,本文在已有研究的基础上,以国家大数据综合试验区的设立为准自然实验,基于其政策效应探究地区大数据发展对企业数字技术创新的影响及其机制。

与已有文献相比,本文的边际贡献主要在于:一是将支持大数据发展的宏观经济政策与微观企业数字技术创新活动置于同一分析框架内,探究地区大数据发展对企业数字技术创新的影响,丰富了大数据发展的政策效应研究和企业数字技术创新的影响因素研究;二是探讨了地区大数据发展赋能企业数字技术创新的信息改善、治理改进、资源获取效应以及企业和行业异质性,有利于深入认识大数据发展的积极作用进而充分利用其红利;三是基于数字技术专利测度企业数字技术创新水平,为国家大数据综合试验区的数字技术创新促进效应提供了经验证据,也为数字技术创新相关研究提供了方法借鉴。

二、理论分析与研究假说

从各个大数据试验区的建设方案和政策实践来看,主要是着力于以下三个方面:一是持续加快数据要素市场建设以有效破除数据要素流通壁垒、提高数据资源交易和配置效率,二是积极完善和升级大数据基础设施以促进各经济主体的数字化转型和地区数字经济发展,三是大力激励和支持大数据产业发展以为经济社会高质量发展提供新动能和新增长点。无论是数据要素市场的建设和完善,还是大数据基础设施的改善和升级,抑或是大数据产业的集聚和发展,都需要先进的数字技术的支持,因而在政府主导和推进的大数据试验区建设中,地方政府必然会通过一定的调控措施支持本地区的数字技术创新,其中对企业数字技术创新的政策激励和资源配置等是支持的重点。与此同时,地区大数据发展也为企业的数字技术创新提供了更好的市场环境和资源条件,从而可以更有效地促进企业的数字技术创新。比如:作为新型生产要素,大数据的增量提质能够打破信息壁垒和数据孤岛,为企业的数字技术创新带来更多路径和收益,同时,大数据的易复制、零边际成本和非损耗等特性颠覆了传统生产函数的定义,降低了数字技术创新成本,使企业数字技术创新的效益和全要素生产率得以提升(邱子迅等,2021)^[9],从而激发企业通过数字技术创新来增强竞争优势的意愿和活力;数据资源开放共享、数据中心整合等将推进大数据产业的发展和跨行业融合(陈文等,2022)^[6],形成数字产业集聚,为企业的数字技术创新合作提供了便利,同时,大数据简化了知识转移的流程,重塑数据要素整合、数字技术共创方式,弱化数字技术创新合作的时空约束,而不同企业间的合作能够大幅减少数字技术创新的成本和周期;大数据软硬件设施的建设和完善通过制度激励、服务优化、数据交易通畅等为数字技术创新提供了优越环境,有利于数字技术创新要素的高效整合,从而显著降低企业进行数字技术创新的成本和阻碍。

基于此,提出假说 H1:设立国家大数据综合试验区可以产生促进试验区内企业数字技术创新水平提升的政策效应。

进一步从阻碍企业数字技术创新的企业自身因素来看,地区大数据发展还可以通过影响这些阻碍因素来促进企业数字技术创新,主要体现在以下方面:一是信息改善效应。信息不对称是阻碍企业数字技术创新的重要因素之一,而地区大数据发展能够有效缓解企业内外部的信息不对称。大数据发展改善了企业的信息披露渠道与方式(Liu et al,2011)^[10],有利于外部投资者了解和掌握企业的经营状况与数字技术创新活动信息,从而改善企业的信息环境与市场预期(吴非等,2021)^[11],为企业带来更多创新资源,还可以促使企业将更多资源投入数字技术研发。同时,大数据、云计算等能够增强企业的信息处理能力(Dalenogare et al,2018)^[12],一方面可以更敏锐地识别和捕获市场信息,另一方面内部不同部门间的信息传递与资源整合也得以优化,进而通过提高资源利用效率、降低生产管理成本等提高经济效益

(申明浩等,2023)^[13],为数字技术创新提供更多资金保障。二是治理改进效应。内部治理不善也会阻碍企业的数字技术创新,而大数据发展可以提高企业财务管理与内部控制的透明度,缓解逆向选择和委托代理问题(Goldfarb et al,2019)^[14],提高内部控制质量与资源配置效率(Cohen et al,2013)^[15]。此外,大数据发展还有利于外部利益相关者参与公司治理和监督企业经营,并通过完善市场机制降低价格扭曲和外部交易成本,有利于企业治理能力的提升(Clemons et al,1993)^[16]。三是资源获取效应。创新资源短缺,尤其是融资约束会对企业数字技术创新形成较大抑制,大数据发展则可以提高企业的资源获取能力,降低融资成本,进而赋能企业数字技术创新。企业能否获得充足的资金来进行数字技术创新取决于其自身的融资能力和外部金融市场有效性(解维敏等,2011)^[17],而大数据发展降低了企业获得金融服务的门槛(Fuster et al,2019)^[18],并通过实时信息交换降低金融服务风险(黄静如等,2020)^[19],从而有效缓解企业的融资约束。此外,大数据试验区内的企业进行数字技术创新不仅响应了政策号召,还回应了市场预期,这也将有利于降低企业的融资成本。

基于此,提出假说 H2:设立国家大数据综合试验区会对试验区内企业产生信息改善效应、治理改进效应和资源获取效应,进而有利于企业数字技术创新水平的提升。

由于企业自身特征、行业特性的不同,地区大数据发展对不同类型企业数字技术创新的影响也可能不同。首先,从企业特征来看:一方面,由于企业数字技术创新具有高风险、高投入和长周期等特点,不仅需要企业具备统筹调配各方资源的能力和经历,也需要企业拥有较强的抗风险能力,因此在数字技术创新方面,风险承担能力较高的企业将更容易获取和利用地区大数据发展带来的红利;另一方面,相比成长和成熟期的企业,处于衰退期的企业往往财务状况欠佳、内部治理水平较低且制度僵化(李云鹤等,2011)^[20],导致其在技术创新方面较为保守,尽管有大数据试验区政策的外部激励,也较难处于数字技术创新的最前沿(童锦治等,2018)^[21]。其次,从行业特征来看:不同行业的技术进步与数据要素和数字技术的关联性不同。不同行业使用信息与通信技术(Information and Communications Technology, ICT)的强度是不同的,与 ICT 关联度较高行业的企业往往数字化程度较高,数字技术应用较多,数字技术创新的能力和意愿较强,因而受地区大数据发展的影响也较大。

基于此,提出假说 H3:设立国家大数据综合试验区对试验区内企业数字技术创新的影响存在异质性,表现为对风险承担能力较强企业、非衰退期企业、与 ICT 关联度较高行业企业的数字技术创新促进效应较大。

三、研究设计与数据说明

1. 模型构建与变量测度

为检验国家大数据综合试验区设立对试验区内企业数字技术创新水平的影响,构建如下多期 DID 模型:

$$Dinnov_{i,t} = \alpha + \beta Bigdata_{i,t} + \gamma X_{it} + Year + Firm + City + Industry \times year + u_{i,t}$$

其中, i 代表企业, t 代表年度, $Year$ 、 $Firm$ 、 $City$ 和 $Industry \times year$ 分别表示年度、企业、城市和行业×年度固定效应, $u_{i,t}$ 为随机扰动项。

(1)被解释变量($Dinnov$):“数字技术创新”。参考 Chen 等(2019)和 Yang(2022)的研究方法^[22-23],采用与数字技术相关的专利申请数量来衡量企业的数字技术创新水平。具体而言,首先利用 Python 软件基于《数字经济及其核心产业统计分类(2021)》、《浙江省数字经济核心产业统计分类目录》以及近年

政府工作报告等文件提取出数字技术关键词,然后与中国专利文本信息进行匹配得到与数字技术创新相关的 IPC(International Patent Classification,国际专利分类)分类号,剔除异常数据后将其与中国研究数据服务平台(CNRDS)数据库中上市公司发明专利和实用新型专利的 IPC 分类号进行匹配,最终得到样本企业的数字技术专利申请数目,并进行加 1 后取自然对数处理,以尽可能克服专利数据的右偏分布问题。

(2)核心解释变量(*Bigdata*):“大数据试验区”。双重差分项为企业是否位于国家大数据综合试验区的虚拟变量(政策虚拟变量与时间虚拟变量的交乘项)。其中,贵州省虽于 2016 年获批,但在 2015 年 9 月就已开始布局,故设定 2015 年为贵州省的政策时间节点,其余试验区的政策时间节点设定为 2016 年。

(3)控制变量(*X*)。参考向海凌等(2022)的研究^[24],选取以下控制变量:一是“企业规模”,采用总资产的自然对数来衡量;二是“现金流水平”,采用经营活动现金流量净额与负债总额的比值来衡量;三是“两职合一”,若董事长与总经理两职合一赋值为 1,否则赋值为 0;四是“员工规模”,采用员工人数的自然对数来衡量;五是“企业成长性”,采用固定资产增长率来衡量;六是“财务杠杆”,采用资产负债率来衡量;七是“盈利能力”,采用总资产回报率来衡量;八是“董事会规模”,采用董事会人数加 1 取自然对数来衡量;九是“委员会个数”;十是“企业价值”,采用托宾 Q 值来衡量。

2. 数据来源与描述性统计

本文的研究样本为 2007—2020 年沪深两市 A 股的上市公司,选择 2007 年作为研究起点是因为从当年开始数字经济规模开始加速扩张,数字技术的应用也快速增长(吴非等,2021)^[11]。剔除金融保险等行业以及主要数据缺失严重、上市时间不足三年、经营异常的企业样本,最终得到包含 2 485 个企业的 25 111 个观测值,其中来自试验区的有 10 678 个,来自非试验区的有 14 433 个。为了克服极端值对回归结果的影响,对所有连续型变量进行双侧 1%水平的缩尾处理。本文所用企业财务数据均来自国泰安(CSMAR)数据库和 Wind 数据库,城市层面数据来自相应年度的《中国城市统计年鉴》,专利数据来自中国研究数据服务平台(CNRDS)数据库。主要变量的描述性统计见表 1,可以发现不同企业之间的数字技术专利产出水平具有较大的差距。此外,试验区样本和非试验区样本的“数字技术创新”变量的均值分别为 1.028 4 和 0.718 4,表明试验区样本企业的数字技术创新水平总体上明显高于非试验区的样本企业。

表 1 主要变量的描述性统计

变 量	样本量	均值	标准差	最小值	最大值
被解释变量 数字技术创新	25 111	0.850 3	1.301 4	0	6.814 6
核心解释变量 大数据试验区	25 111	0.203 5	0.402 6	0	1
盈利能力	25 111	0.036 6	0.064 2	-0.291 4	0.212 4
财务杠杆	25 111	0.434 8	0.207 3	0.050 2	0.984 9
两职合一	25 111	0.239 3	0.426 7	0	1
现金流水平	25 111	0.185 2	0.315 7	-0.536 5	1.641 2
控制变量 企业规模	25 111	22.116 6	1.308 2	19.251 2	27.098 7
企业成长性	25 111	0.198 4	0.632 6	-0.591 5	4.876 7
员工规模	25 111	7.736 1	1.276 6	4.007 3	11.296 7
董事会规模	25 111	2.259 2	0.178 0	1.791 8	2.772 6
委员会个数	25 111	3.915 5	0.497 8	2	6
企业价值	25 111	2.098 3	1.364 7	0.871 8	8.758 0

四、实证结果与分析

1. 基准模型回归分析

应用双重差分法的前提条件是实验组和控制组在政策冲击前具有相同的发展趋势,就本文而言,需要满足试验区设立前实验组与控制组企业数字技术创新水平变化趋势基本相同的条件。本文采用事件研究法进行平行趋势检验,并考察政策实施后的动态效应。平行趋势检验结果如图 1 所示(置信区间为 95%)。在试验区设立之前,实验组与控制组的数字技术创新水平变化趋势并无显著差异,满足平行趋势假设;从动态效应看,试验区设立以后,实验组与控制组出现明显差异,政策效应较为稳定且持续性较强。表 2 报告了基准模型回归结果,核心解释变量“大数据试验区”的估计系数在 1%的水平上显著为正,表明国家大数据综合试验区的设立显著促进了试验区内企业数字技术创新水平的提高,假说 H1 得到验证。

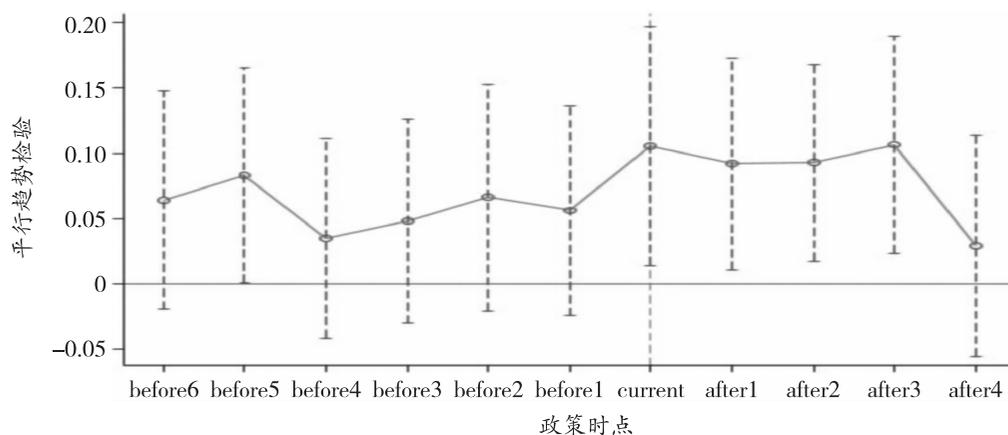


图 1 平行趋势检验结果

表 2 基准模型回归结果

变量	(1)	(2)	(3)	(4)
大数据试验区	0.782 0***(0.022 5)	0.633 6***(0.020 5)	0.081 8***(0.020 0)	0.108 3***(0.019 3)
盈利能力		-0.601 9***(0.120 6)	-0.051 8(0.094 9)	-0.000 6(0.094 9)
财务杠杆		-0.556 3***(0.051 0)	0.001 0(0.045 3)	0.007 5(0.044 7)
两职合一		0.105 8***(0.019 6)	0.038 2**(0.015 2)	0.045 0***(0.015 3)
现金流水平		-0.318 1***(0.027 3)	-0.035 9*(0.018 4)	-0.033 0*(0.018 6)
企业规模		0.252 4***(0.012 9)	0.296 5***(0.014 7)	0.290 9***(0.015 1)
企业成长性		0.071 6***(0.012 4)	0.002 9(0.008 4)	0.005 3(0.008 2)
员工规模		0.174 6***(0.011 3)	0.077 7***(0.012 1)	0.086 3***(0.011 8)
董事会规模		-0.395 3***(0.044 8)	0.020 3(0.047 3)	0.039 7(0.047 4)
委员会个数		-0.106 5***(0.014 5)	-0.031 3*(0.014 7)	-0.025 4*(0.014 6)
企业价值		0.094 5***(0.007 6)	0.042 8***(0.005 1)	0.040 1***(0.005 4)
常数项	0.406 3***(0.012 5)	-5.047 5***(0.245 7)	-6.368 8***(0.325 9)	-6.396 1***(0.336 0)

续表

变 量	(1)	(2)	(3)	(4)
企业固定效应	未控制	未控制	控制	控制
年度固定效应	未控制	未控制	控制	控制
城市固定效应	未控制	未控制	未控制	控制
行业×年度固定效应	未控制	未控制	未控制	控制
Adjusted R ²	0.088 6	0.200 0	0.748 5	0.753 5
观测值	25 111	25 111	25 111	25 111

注:括号内数值为企业层面聚类标准误,*、**和***分别表示在10%、5%和1%的水平上显著,下表同。

2. 稳健性检验

为了确保前文分析结果稳健,进行以下稳健性检验:

(1)控制宏观经济因素和其他政策的影响。在宏观经济因素方面,选取6个变量:企业所在城市的“人均GDP(取自然对数)”“第二产业占比”“实际利用外资额(取自然对数)”“1985年邮政所数与互联网普及率的交乘项”、企业所在省份的“市场化指数”和企业所在行业的“赫芬达尔指数”。在其他政策的影响方面,选取同时期对企业数字技术创新影响较大的3个国家政策,即国家创新型城市试点政策、“宽带中国”示范城市政策、国家知识产权试点和示范城市政策,并分别构建双重差分项。在基准模型基础上加入上述变量的估计结果见表3,“大数据试验区”的估计系数依然在1%的水平上显著为正。

表3 稳健性检验结果1

变量	(1)	(2)	(3)
大数据试验区	0.094 0***(0.021 1)	0.144 8***(0.036 5)	0.222 6***(0.043 6)
人均GDP	0.009 9(0.085 0)		-0.009 8(0.085 2)
第二产业占比	0.004 3(0.002 6)		0.005 2*(0.002 7)
实际利用外资	0.017 2**(0.008 4)		0.015 9*(0.008 5)
1985年邮政所数×当年互联网普及率	0.274 1**(0.116 4)		0.214 2*(0.117 8)
市场化指数	0.004 7(0.013 0)		0.005 2(0.013 0)
赫芬达尔指数	0.108 1*(0.061 0)		0.113 2*(0.060 9)
国家创新型城市试点		0.019 3(0.020 9)	0.016 4(0.024 0)
“宽带中国”示范城市		0.360 4***(0.109 1)	0.215 1*(0.117 7)
国家知识产权试点和示范城市		-0.041 8(0.035 2)	-0.140 0***(0.042 2)
Adjusted R ²	0.763 7	0.753 8	0.763 9
观测值	20 214	25 103	20 214

注:本表模型均控制了前文的控制变量以及企业、年度、城市、行业×年度固定效应,限于篇幅,控制变量和常数项的估计结果略,备索,下表同。

(2)安慰剂检验。首先,通过改变政策实施时间进行反事实检验,即将设立试验区的时间提前2年重新构建双重差分项来进行模型检验。回归结果见表4的(1)列,反事实的政策变量估计系数并不显著,表明真实的政策效应是显著存在的。然后,通过随机生成实验组的方法进行1 000次回归,1 000次

回归结果的 t 统计量分布见图 2。 t 统计量均在 0 值附近,与基准模型回归结果相差较大,且近似于正态分布,说明基准模型的回归结果几乎没有受到随机因素的干扰,本文的结论是稳健的。

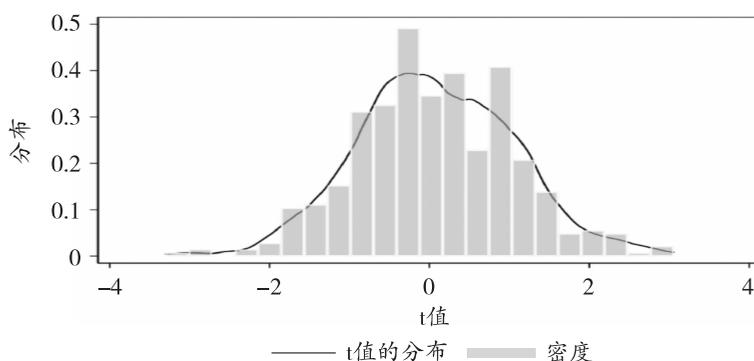


图 2 安慰剂检验结果

(3) 删除特殊样本和替换被解释变量。一是考虑到直辖市拥有较高的经济管理权限和较多的政治经济资源,因而剔除注册地位于直辖市的样本企业重新进行模型检验,估计结果见表 4 的(2)列;二是考虑到创业板上市公司多为高新技术企业,因而剔除创业板的样本企业重新进行模型检验,估计结果见表 4 的(3)列;三是考虑到专利从申请到授权需要经过一定的审查程序,因而采企业数字技术专利的获批数量(加 1 取自然对数)来替换原被解释变量,重新估计的结果见表 4 的(4)列。上述检验的核心解释变量估计系数仍然显著为正,表明本文分析结果较为稳健。

表 4 稳健性检验结果 2

变量	(1)	(2)	(3)	(4)	(5)	(6)
大数据试验区		0.122 5*** (0.022 2)	0.109 6*** (0.019 3)	0.174 4*** (0.020 3)	0.127 2*** (0.027 9)	0.508 2*** (0.121 5)
反事实政策	0.027 4 (0.022 2)					
Adjusted R ²	0.753 2	0.745 8	0.760 0	0.779 0	0.741 5	
观测值	25 103	20 024	21 203	25 103	12 251	25 026

(4) PSM-DID 检验。本文采用倾向得分匹配方法(PSM)来克服样本自选择可能导致的“筛选效应”内生性问题,即采用 1:1 的最近邻匹配法进行 PSM-DID 检验。PSM 模型采用前文所有控制变量作为预测变量,平衡性检验结果显示,预测变量的配对偏差均小于 5%,且实验组与控制组至少在 10%的水平上无显著差异,表明匹配效果较好。经过 1:1 最近邻匹配后,获得 12 251 个观测值,重新进行回归检验,结果见表 4 的(5)列,政策效应仍然在 1%水平上显著。进一步按照 1:2 和 1:3 的比例进行匹配,估计结果依然稳健(分析结果略,备索)。

(5) 工具变量法。尽管本文核心解释变量属于外生性较强的政策变量,但为排除可能的内生性问题,进一步采用工具变量法进行稳健性检验。参照赵晓阳和衣长军(2022)的方法^[25],结合国家统计局发布的《数字经济及其核心产业统计分类(2021)》,以企业所在城市到沿海港口的距离(取自然对数)与每年全国数字企业存量(取自然对数)的交互项构造工具变量,进行工具变量法(2SLS)检验。第一阶段的检验结果显示(具体结果略),F 统计量远大于经验法则的临界值,工具变量与核心解释变量满足相关

性要求。K-Paaprk Wald F 统计量和 K-Paaprk LM 统计量通过了弱工具变量检验和可识别性检验。第二阶段的估计结果见表 4 的(6)列,“大数据试验区”的估计系数还是显著为正。上述检验表明本文的分析结果具有较好的稳健性。

3. 影响机制分析

(1)信息改善效应。本文采用两种方法来衡量企业的信息不对称程度:一是“信息不对称 1”,借鉴 Dechow 等(1995)的方法^[26],利用修正的 Jones 模型计算可操控应计利润,用其绝对值来衡量外界与企业间的信息不对称程度,该指标值越小则企业的信息不对称程度越低。二是“信息不对称 2”,借鉴李莉等(2014)的研究^[27],以非流动比率来测度企业的信息不对称程度,非流动比率越大,意味着企业股票的流动性越低,则企业信息不对称程度越高。分别以上述两个变量为被解释变量的检验结果见表 5 的 Panel A,“大数据试验区”的估计系数均显著为负,表明国家大数据综合试验区的设立显著降低了企业内外部之间的信息不对称程度,产生了显著的信息改善效应。

(2)治理改进效应。本文从外部交易成本和内部控制质量两个方面来刻画企业的治理状况:一是借鉴赵璨等(2020)的做法^[28],采用固定资产投资、在建工程、无形资产及长期待摊费用合计占总资产之比来度量企业的资产专用性,并以此作为“外部交易成本”的代理变量,资产专用性越强,变现时价格损失越大,外部交易成本就越高。二是采用迪博企业内部控制指数来衡量企业的“内部控制质量”,指数越大内部控制质量越高。分别以上述两个变量为被解释变量的检验结果见表 5 的 Panel B,“大数据试验区”对“外部交易成本”和“内部控制质量”的估计系数分别显著为负和显著为正,表明国家大数据综合试验区的设立能显著降低企业的外部交易成本并提高企业的内部控制质量,产生了显著的治理改进效应。

(3)资源获取效应,本文主要从缓解融资约束的角度来考察企业的资源获取情况,选用了权益融资成本和债务融资成本两个指标:一是借鉴王翌秋和谢萌(2022)的做法^[29],采用 PEG 模型估计企业的“权益融资成本”;二是用利息支出与短期债务加长期债务之和的比值来衡量企业的“债务融资成本”,其中,短期负债为资产负债表中的短期借款,长期负债包括 1 年内到期的长期借款、应付债券、长期应付款、其他长期负债项。分别以上述两个变量为被解释变量的检验结果见表 5 的 Panel C,“大数据试验区”对“权益融资成本”和“债务融资成本”的估计系数均显著为负,表明国家大数据综合试验区的设立能显著降低企业的融资成本,产生了显著的融资约束缓解效应。由此,假说 H2 得到验证。

表 5 影响机制检验结果

变量	Panel A:信息改善效应		Panel B:治理改进效应		Panel C:资源获取效应	
	信息不对称 1	信息不对称 2	外部交易成本	内部控制质量	权益融资成本	债务融资成本
大数据试验区	-0.127 3*** (0.042 0)	-0.312 6*** (0.100 3)	-0.005 5*** (0.001 1)	0.008 5** (0.003 9)	-0.008 9** (0.004 3)	-0.026 8*** (0.005 3)
Adjusted R ²	0.737 1	0.582 2	0.701 4	0.395 6	0.692 1	0.577 2
观测值	21 389	22 580	25 045	22 341	22 338	22 338

4. 异质性分析

(1)企业异质性。本文分别对处于不同生命周期的企业和风险承担能力不同的企业进行分组检验。采用现金流组合法将样本企业划分为“成长期企业”“成熟期企业”“衰退期企业”3 个子样本,分别进行

模型检验,回归结果见表6的Panel A。“大数据试验区”对“数字技术创新”的估计系数在“成长期企业”和“成熟期企业”子样本中均显著为正,而在“衰退期企业”子样本中不显著,表明国家大数据综合试验区的设立能显著促进处于成长期和成熟期的企业数字技术创新,但对衰退期企业的影响不显著。借鉴李小荣和张瑞君(2014)的做法^[30],选取企业近三年滚动的ROA波动程度(近三年经行业调整的ROA标准差)来度量企业的风险承担能力,并将样本企业划分为“低风险承担企业”和“高风险承担企业”2个子样本,分别进行模型检验,回归结果见表6的Panel B。在“高风险承担企业”子样本中,“大数据试验区”对“数字技术创新”的估计系数显著为正,而在“低风险承担企业”子样本中,“大数据试验区”对“数字技术创新”的估计系数不显著,表明国家大数据综合试验区的设立能显著促进风险承担能力较强企业的数字技术创新,但对风险承担能力较低企业的影响不显著。

(2)行业异质性。根据2018年中国投入产出表中数字经济核心产业对各行业的直接消耗系数计算行业的ICT强度(与ICT的关联程度),并以均值进行高ICT行业与低ICT行业的划分,进而将属于高ICT行业的样本企业归为“高ICT行业”子样本,将属于低ICT行业的样本企业归为“低ICT行业”子样本,分别进行模型检验,回归结果见表6的Panel C。在“高ICT行业”子样本中,“大数据试验区”对“数字技术创新”的估计系数显著为正,而在“低ICT行业”子样本中,“大数据试验区”对“数字技术创新”的估计系数不显著,表明国家大数据综合试验区的设立能显著促进与ICT关联度较高行业的企业数字技术创新,但对与ICT关联度较低行业的企业影响不显著。由此,假说H3得到验证。

表6 异质性分析结果

变 量	Panel A:生命周期异质性			Panel B:风险承担能力异质性		Panel C:行业异质性	
	成长期企业	成熟期企业	衰退期企业	低风险承担企业	高风险承担企业	高 ICT 行业	低 ICT 行业
大数据试验区	0.132 9*** (0.034 3)	0.114 6*** (0.035 4)	0.067 7 (0.052 4)	0.052 3 (0.043 1)	0.157 3*** (0.028 0)	0.183 9*** (0.028 1)	0.018 9 (0.026 3)
Adjusted R ²	0.777 9	0.751 3	0.702 1	0.787 6	0.734 9	0.787 4	0.691 6
观测值	10 703	8 990	4 162	10 114	14 623	14 220	10 841

五、结论与启示

在数字经济时代,企业的数字技术创新逐渐成为创新驱动发展的重要支撑。作为一种新型生产要素,数据要素不仅是基础性和战略性资源,而且也是推动数字技术创新的重要因素,因此,地区大数据发展应该具有促进企业数字技术创新的基本功能。科学评估设立国家大数据综合试验区对企业数字技术创新产生的政策效应,有助于正确认识数字经济背景下大数据发展对企业高质量发展的深刻影响。本文将国家大数据综合试验区的设立作为一项准自然实验,以2007—2020年沪深两市A股上市公司为研究样本,运用Python软件对企业数字技术专利进行识别,进而采用多期双重差分模型实证考察该政策对企业数字技术创新的影响及其作用机制,主要结论如下:(1)设立国家大数据综合试验区显著促进了试验区内企业数字技术创新水平的提升,该结论在经过控制宏观经济因素及其他政策干扰、替换被解释变量、安慰剂检验、PSM-DID检验、工具变量法等一系列稳健性检验后依然成立;(2)设立国家大数据综合试验区对试验区内企业产生了显著的信息改善效应、治理改进效应和资源获取效应,从而有效促进了企业数字技术创新;(3)设立国家大数据综合试验区,对于处于成长期和成熟期的企业、风险承担能力强的企业、与ICT关联度高行业的企业产生了显著的数字技术创新促进效应,而对于处于衰退期的企业、风

险承担能力弱的企业、与 ICT 关联度低行业的企业的技术创新水平没有显著影响。

本文为地区大数据发展赋能企业数字技术创新提供了经验证据,同时,基于以上结论,可得出以下启示:第一,各地区应积极推动大数据发展,并充分发挥大数据在促进技术进步和企业发展等方面的积极作用。目前我国在大数据发展和应用方面已具备一定基础,拥有市场优势和发展潜力,但也存在政府数据开放共享不足、产业基础薄弱、缺乏顶层设计和统筹规划、法律法规建设滞后、创新应用领域不广等问题。因此,要通过数据要素市场和大数据基础设施建设促进大数据和大数据产业的发展;在完善数据要素市场和数字基础设施、推动大数据产业发展的同时,要坚持创新驱动发展,在微观层面帮助企业精准破解在数字技术创新过程中的掣肘痛点与难点。第二,要充分发挥大数据综合试验区的示范引领作用与辐射带动作用,形成以数字技术进步推动大数据发展、以大数据发展促进数字技术创新的良性循环。第三,要科学认识地区大数据发展赋能企业数字技术创新的微观机制和异质性特征,更好发挥政府引导资源配置的作用。地方政府应当建立健全数据要素市场和数字基础设施建设的监督和反馈机制,并针对不同类型企业制定差异化和个性化的数字技术创新帮扶政策。

参考文献:

- [1] 刘洋,董久钰,魏江. 数字创新管理:理论框架与未来研究[J]. 管理世界,2020,36(7):198-217+219.
- [2] 谢谦,郭杨. 数字技术、创新要素结构优化与企业全要素生产率[J]. 北京师范大学学报(社会科学版),2022(6):134-144.
- [3] 罗佳,张蛟蛟,李科. 数字技术创新如何驱动制造业企业全要素生产率?——来自上市公司专利数据的证据[J]. 财经研究,2023,49(2):95-109+124.
- [4] 黄勃,李海彤,刘俊岐,等. 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,2023,58(3)97-115.
- [5] 孙哲远. 数字经济发展如何影响制造业企业“脱虚向实”?——来自国家级大数据综合试验区的证据[J]. 现代经济探讨,2022(7):90-100.
- [6] 陈文,常琦. 大数据赋能了企业绿色创新吗——基于国家级大数据综合试验区的准自然实验[J]. 财经科学,2022(8):76-92.
- [7] 王晓红,李娜. 数字技术发展、产学研合作与企业创新能力——基于国家级大数据综合试验区的分析[J]. 科技管理研究,2022,42(17):1-8.
- [8] 侯林岐,程广斌,王雅莉. 国家级大数据综合试验区如何赋能企业数字化转型[J/OL]. 科技进步与对策:1-11(2022-09-28). <http://kns.cnki.net/kcms/detail/42.1224.G3.20220926.1210.002.html>.
- [9] 邱子迅,周亚虹. 数字经济发展与地区全要素生产率——基于国家级大数据综合试验区的分析[J]. 财经研究,2021,47(7):4-17.
- [10] LIU D, CHEN S, CHOU T. Resource fit in digital transformation lessons learned from the CBC Bank global e-banking project[J]. Management Decision, 2011, 49(910):1728-1742.
- [11] 吴非,胡慧芷,林慧妍,等. 企业数字化转型与资本市场表现——来自股票流动性的经验证据[J]. 管理世界,2021,37(7):130-144+10.
- [12] DALENOGARE L S, BENITEA G B, AYALA N F, et al. The expected contribution of Industry 4.0 technologies for industrial performance[J]. International Journal of Production Economics, 2018, 204:383-394.
- [13] 申明浩,谭伟杰. 数字化与企业绿色创新表现——基于增量与提质的双重效应识别[J]. 南方经济,2022(9):118-138.
- [14] GOLDFARB A, TUCKER C. Digital Economics[J]. Journal of Economic Literature, 2019, 57(1):3-43.
- [15] COHEN L, DIETHER K, MALLOY C. Misvaluing innovation[J]. The Review of Financial Studies, 2013, 26(3):635-666.

- [16] CLEMONS E K, REDDI S P, ROW M C. The impact of information technology on the organization of economic activity: The “Move to the Middle” hypothesis[J]. *Journal of Management Information Systems*, 1993, 10(2):9-35.
- [17] 解维敏,方红星. 金融发展、融资约束与企业研发投入[J]. *金融研究*, 2011(5):171-183.
- [18] FUSTER A, PLOSSER M, SCHNABL P, et al. The role of technology in mortgage lending[J]. *The Review of Financial Studies*, 2019, 32(5):1854-1899.
- [19] 黄静如,刘永模. 媒体关注对企业债务融资成本的影响研究——基于会计稳健性的中介效应检验[J]. *投资研究*, 2020, 39(2):113-133.
- [20] 李云鹤,李湛,唐松莲. 企业生命周期、公司治理与公司资本配置效率[J]. *南开管理评论*, 2011, 14(3):110-121.
- [21] 童锦治,刘诗源,林志帆. 财政补贴、生命周期和企业研发创新[J]. *财政研究*, 2018(4):33-47.
- [22] CHEN M A, WU Q, YANG B. How valuable is fintech innovation? [J]. *The Review of Financial Studies*, 2019, 32(5):2062-2106.
- [23] YANG C H. How artificial intelligence technology affects productivity and employment: Firm-level evidence from Taiwan [J]. *Research Policy*, 2022, 51(6):104536.
- [24] 向海凌,陈玉茹,吴非. QFII 持股促进企业数字化转型的机制研究——影响路径、企业异质性与提质补短效应[J]. *西部论坛*, 2022, 32(6):1-19.
- [25] 赵晓阳,衣长军. 数字经济发展提高了城市创业活跃度吗? [J]. *现代财经(天津财经大学学报)*, 2022, 42(11):19-31.
- [26] DECHOW P M, SLOAN R G, SWEENEY A P. Detecting earnings management[J]. *The Accounting Review*, 1995, 70(2):193-225.
- [27] 李莉,闫斌,顾春霞. 知识产权保护、信息不对称与高科技企业资本结构[J]. *管理世界*, 2014(11):1-9.
- [28] 赵璨,曹伟,姚振晔,等. “互联网+”有利于降低企业成本粘性吗? [J]. *财经研究*, 2020, 46(4):33-47.
- [29] 王翌秋,谢萌. ESG 信息披露对企业融资成本的影响——基于中国 A 股上市公司的经验证据[J]. *南开经济研究*, 2022, 233(11):75-94.
- [30] 李小荣,张瑞君. 股权激励影响风险承担:代理成本还是风险规避? [J]. *会计研究*, 2014(1):57-63+95.

Research on the Impact of Big-data Development on Enterprise Digital Technology Innovation: Quasi-natural Experiment Based on the National Big Data Comprehensive Pilot Zone

DAI Yan-juan, SHEN Wei-peng, TAN Wei-jie

(School of Economics and Trade, Guangdong University of Foreign Studies, Guangzhou 510006, Guangdong, China)

Abstract: As a new factor of production in the era of the digital economy, the data factor is different from the traditional factors of production such as labor, capital, and technology. It is not only the basic strategic resources but also the driving force of factors occupying new advantages in the international digital competition in the new era. In the context of the development of the digital economy, formulating relevant policies and measures, effectively promoting industrial innovation changes, and enterprises increasing their investment in digital technology innovation projects have important theoretical and practical significance. How to effectively promote enterprise digital technology innovation? The construction and improvement of the data factor market

may be a breakthrough. However, the empirical analysis of the effect of big data development on the innovation of digital technology of enterprises is relatively scarce.

In this paper, Python software is used to identify enterprise digital innovation patents, and based on the quasi-natural experiment of the national big data comprehensive pilot zone, the A-share listed companies in Shanghai and Shenzhen from 2007 to 2020 are taken as research samples, and the multi-period differential method is used to investigate the impact of big data on enterprise digital innovation and its mechanism. The analysis results show that the construction of the big data pilot area significantly improves the digital innovation performance of enterprises, and this conclusion is still valid on the basis of parallel trends, PSM-DID, placebo tests, and a series of robustness tests. In the big data pilot areas, the information effect, governance effect, resource effect, and other mechanisms have contributed significantly to the digital innovation of enterprises. In addition, the construction of a data factor market reduces the degree of information constraints, internal and external governance costs, and financing costs of enterprises, and helps improve enterprises' digital innovation ability. Last but not least, heterogeneity analysis shows that the digital technology innovation effect of building enabling enterprises in big data pilot zones is not obvious in declining enterprises, enterprises with low risk-bearing capacity, and industries with low ICT intensity.

Compared with the existing literature, from the perspective of research, this paper puts the macro-economic policies supporting the development of big data and micro-enterprise digital innovation activities in the same analytical framework, explores the dividend effect of big data development on enterprise digital innovation, and enriches the research literature of big data development policy evaluation. In terms of the research content, taking the establishment of the national big data comprehensive pilot zone as the exogenous impact, the paper discusses that the digital innovation of enterprises empowered by the development strategy of big data mainly depends on the "information effect", "governance effect" and "resource effect", and identifies the regional, industrial and enterprise heterogeneity influencing characteristics in the process of digital innovation of enterprises empowered by big data, deepening the existing research. In terms of research data, Python software is used to identify the text content of digital economy-related policies, and the obtained keyword frequency information is matched with the patent classification number to obtain the IPC classification number of digital innovation patents, so as to identify the enterprise digital innovation evaluation index more accurately. On this basis, the digital innovation effect of the big data pilot zone policy is further investigated to provide micro evidence for the policy evaluation of the big data comprehensive pilot zone. This study not only reveals the micro policy dividends of big data pilot zone construction, but also has important experience and enlightenment for formulating relevant policies of data factor market construction and promoting high-quality development of the real economy.

Key words: data factor; big data comprehensive pilot zone; digital technology innovation; digital economy; digital infrastructure; big data industries

CLC number: F273. 1; F127

Document code: A

Article ID: 1674-8131(2023)02-0016-13

(编辑:刘仁芳)