

大数据发展、人工智能应用与 全要素生产率*

——来自国家级大数据综合试验区的证据

张龙鹏 唐 芳

[摘要] 本文基于中国A股上市公司数据,以设立国家级大数据综合试验区衡量大数据发展,使用双重差分法研究了大数据发展对企业全要素生产率的影响,并讨论了人工智能应用在其中所发挥的作用机制。实证结果显示,大数据发展能够显著提升企业全要素生产率,促进人工智能应用,此为重要影响机制之一;大数据发展的全要素生产率促进效应对于五大城市群企业、中小规模企业、制造业企业更为显著;国家级大数据综合试验区与宽带中国示范城市和国家信息消费城市建设具有同向的政策联动效应。进一步地,从空间视角来看,大数据发展存在一定程度的行政隔离,数据要素跨省域流动受阻,进而削弱了企业全要素生产率的提升效应。本研究深化了大数据发展与全要素生产率认识,对推动大数据发展、助力企业数字成长和提升生产效率提供启示参考。

[关键词] 大数据 人工智能 全要素生产率 国家级大数据综合试验区

[中图分类号] F062.4 [文献标志码] A [文章编号] 1009-8461(2025)04-081-23

一、引言

近年来,由于供给端的资本、劳动力等生产要素发挥作用受限,经济贡献度下降,中国经济增速放缓。在此背景下,党的二十大报告强调要着力提高全要素生产率以推动经济高质量发展。提升全要素生产率作为新发展理念的重要体现和高质量发展的动力源泉,被证明是实现经济增长

* 作者简介:张龙鹏,电子科技大学深圳高等研究院、公共管理学院副教授;唐芳,电子科技大学公共管理学院硕士研究生。

基金项目:国家社会科学基金青年项目“全球价值链视角下人工智能产业融合的效应与路径研究”(20CJY009);深圳市哲学社会科学规划课题“深圳公共数据要素可信流通路径研究”(SZ2024C008)。

的一条必选路径(刘建翠, 2022)。全要素生产率包含技术进步、效率提升等多层含义, 其在研究中常被用作表征经济发展质量的关键指标。

在数字经济时代, 大数据、人工智能等新一代数字技术正成为经济增长的关键动力, 数据要素对经济发展的贡献率持续上升, 因此, 研究提升全要素生产率的路径, 需结合数字经济的时代背景进行探讨。数据要素是数字经济不断壮大的关键生产要素、国家基础性战略性资源, 是推动经济社会高质量发展的重要引擎。我国政府十分重视发展大数据, 对大数据在全国和区域性的发展作出重要部署, 陆续发布了《“十四五”大数据产业发展规划》《关于构建更加完善的要素市场化配置体制机制的意见》《促进大数据发展行动纲要》等重要文件。可见, 如何通过善用大数据提升全要素生产率推动经济高质量发展是一个重要课题。

学术界已对关于大数据促进企业生产率攀升的研究进行了广泛讨论, 其中主要有两类: 一是说明大数据能缩减企业决策成本、提升决策精度, 推进企业生产效率、财务管理效率、全要素生产率等的提升, 有利于企业成长(Chen et al., 2021; 张益豪和郭晓辉, 2023); 二是研究作用机制, 主要涉及研发创新、产业结构升级、经济集聚等角度(殷红等, 2020; 曹正旭等, 2020)。进一步地, 伴随着海量资源挖掘和数据流通调动, 人工智能作为数字经济的新引擎, 正在给生产流通、供应链和流程管理等带来可观效益, 其作用不可忽视。虽然许多学者从机器人的劳动替代效应、企业全价值链等视角研究了人工智能对全要素生产率的影响(Koch et al., 2021; 张龙鹏和钟易霖, 2023), 但把人工智能作为大数据发展与全要素生产率提升的影响因素的机制研究较少。总的来说, 已有文献表明大数据、人工智能对企业全要素生产率的提升都具有推动作用, 但是较少揭示大数据提升全要素生产率过程中人工智能应用引起的获益差异性。因此, 本文将主要回答两个问题: 大数据发展是否能提升企业全要素生产率? 人工智能是否在作用机制中扮演了关键角色?

回顾已有文献, 我们还注意到大数据发展的经济效应研究包含区域创新绩效等方面(李雪等, 2021); 也有学者对大数据发展的空间溢出性作了探索, 证明大数据能够促进资源要素在城市间流动, 形成空间辐射效应(郭炳南等, 2022)。但是, 数据流动范围一定是存在限度的。在实际的行政空间里, 不同地区的大数据发展可能在数据供需匹配、数据共享制度建设、异议处理措施设置等方面缺乏共通性, 阻碍着大数据跨地区流动。因此, 基于溢出效应, 我们还需要探究一个问题: 大数据跨行政区域流通是否存在流通阻碍? 综上, 本文为了探究大数据区域发展更深层次的内涵, 在提升效应的基础上进一步实证大数据发展的行政隔离效应。

出于数据隐私保护和大数据监管等要求, 数据流通不可避免地具有较大的不透明性、难验证性和未知性等特征, 难以进入大数据这个“黑匣子”的内部(Davenport et al., 2012)。因此, 本文衡量大数据发展也基于上述的“黑匣子”背景, 从区域大数据政策视角探讨其整体发展效果。2016年2月、10月, 国家发展改革委、工业和信息化部、中央网信办发函同意在贵州、京津冀等八个区域推进国家级大数据综合试验区建设(以下简称“试验区”)^{①②}。设立试验区意在从大数

①《七个地方和区域获批国家大数据综合试验区》, 中华人民共和国工业和信息化部, 2016-10-10, https://www.miit.gov.cn/ztzl/lstz/tddsycyfz/sdsf/art/2020/art_59cc06f2e0604858a60ab35c6b0319de.html。

②八大国家级大数据综合试验区分别为: 国家大数据(贵州)综合试验区, 京津冀、珠三角、上海、河南省、重庆、沈阳和内蒙古国家大数据综合试验区。具体城市为: 贵阳、北京、天津、张家口、廊坊、承德、秦皇岛、石家庄、广州、深圳、佛山、珠海、惠州、上海、郑州、洛阳、重庆、沈阳和内蒙古。

据制度创新、公共数据开放共享和大数据创新应用等方面进行试验探索，推动我国大数据创新发展，形成可复制、可推广的大数据区域发展经验。试验区的设立、推广作为一项重大试点改革，是循时代要求、抓住发展机遇的关键一招。随着国家级大数据综合试验区的深入建设与大数据战略行动统筹推进，大数据发展的区域性成果得到充分体现。一方面，该政策干预基本不受企业层面的影响，作为外生冲击较好地解决了内生性问题，符合外生性、整体性、代表性等要求。另一方面，试验区形塑了一个数据流动空间，让研究不被局限在表面的因果关系，有利于进一步认识大数据发展在现实与虚拟双重空间融合的数据驱动力和行政阻力。因此，本文以2007—2019年中国A股上市公司为样本，利用设立国家级大数据综合试验区的准自然实验，通过双重差分模型考察大数据发展对企业全要素生产率的影响、机制、异质性等内容，并对政策联动和行政隔离展开分析。

二、文献综述

（一）文献回顾

1. 大数据的概念、特点与趋势

大数据是难以在一定时间范围内用传统工具进行储存、分析和可视化等进一步操作的海量数据集合，具有速度快、结构复杂、种类繁多、量级庞大等特点（Sagiroglu & Sinanc, 2013）。大数据有许多分类标准，例如基于统一的空间和时间参照系统，形成了空间、时间和时空大数据。如今数字资源深度融合、科学理论持续更新，对大数据的研究也很丰富，既有大数据系统、分析技术、安全与隐私保护等传统研究内容，也有大数据平台治理等与时俱进的研究（程学旗等，2014；刘凌和罗戎，2017）。

2. 大数据发展的社会经济绩效

在数字化背景下，大数据得到广泛使用，形成了范围经济、规模经济和长尾效应的经济环境，不断完善微观价格机制，提升经济的均衡水平（荆文君和孙宝文，2019）。学者们关注了大数据发展对资源利用效率、市场效率等经济效率的影响，大数据以更低的成本实现资源需求侧管理，提升了城市运转中对空间范围内各类资源使用的精度和效度（Martin & Nagel, 2022）。而且，大数据通过推动产业融合、要素跨区域双向流动，能提升要素配置效率，促进城乡区域融合发展，具有显著的空间溢出效应（黄永春等，2022）。上述研究普遍认为，大数据作为重要生产要素，能推动经济高质量发展。

当然，数据只有流动起来并被加以合理利用，其技术与经济特性才能充分发挥，进而实现价值。从微观数据来看，已有研究围绕数据价值展开了讨论，主要包含政府与企业数据价值。对政府数据流动价值的研究主要集中在政府数据开放上。有学者从政府数据开放的角度研究了其必要性，以及如何开放。政府数据开放是推动经济创新发展和促进数据文化形成的新途径（司林波等，2017）。地方政府数据开放、数据适度开放能够通过优化营商环境、促进高质量创新和减少企业创新成本等方面提升企业全要素生产率（彭远怀，2023；张莉和林安然，2023）。同样地，企业要增强对数据的洞察力，应用大数据技术开展管理和生产活动，通过大数据技术分析业务，

来促进企业经营绩效的提升 (Wamba et al., 2017)。企业通过获取大数据资源, 匹配并结合企业能力, 能实现企业产品创新绩效 (谢康等, 2020)。数据与企业已有的技术资源充分融合, 能够优化创新要素结构, 实现企业全要素生产率的提升 (谢谦和郭杨, 2022), 即数据要素在企业中的流动和充分利用对企业生产率提升可能具有积极作用。

3. 数字化发展对全要素生产率的影响

已有文献主要关注了数字基础设施建设和数字化转型对全要素生产率的影响。一是数字基础设施建设方面。数字基础设施建设是大数据发展的底层支撑, 涵盖了云计算、数据中心、物联网等丰富内容, 能强化组织的数字化能力, 并通过技术驱动、劳动力交易平台创新、大数据发展政策创新等手段促进经济增长 (Popescu et al., 2018; Zhang, 2021)。以强化数字化能力为导向的数字基础设施建设, 推动着科技与经济深度融合, 同时强化了数字的应用、营利和创新等方面, 引领着技术效率提升, 显著促进全要素生产率的增长 (Mitra et al., 2016; 范合君和吴婷, 2022)。二是数字化转型方面。随着新一代信息技术与传统制造技术深度融合, 企业的数字化转型、数字能力提升引起广泛关注。现有文献指出了数字化转型既能通过优化人力资本结构、加速企业间知识溢出、缓解融资约束等来提升企业全要素生产率 (赵宸宇等, 2021; 涂心语和严晓玲, 2022; 花俊国等, 2022), 也能通过革新企业间互动模式、集成生产所需资源等提升企业绩效 (Aversa et al., 2017), 证明了数字化转型是数字经济时代提升企业生产效率的强大驱动力。部分文献还从企业数字技术资源整合、技术成本控制、企业信息系统质量提升、商业模式创新等方面, 详述了企业通过提升数字技术能力可以提升生产率 (Bharadwaj, 2000; Dedrick et al., 2013)。进一步地, 大量文献还从数字化转型助力企业实现差异化和成本领先、提升企业的竞争地位等方面, 探讨其对资源配置效率、企业绩效等的正面影响 (吕可夫等, 2023; 杨德明和刘泳文, 2018)。

(二) 文献评述

回顾以上三方面文献, 可以发现, 首先, 研究大数据发展及其经济效应具有极大的必要性, 且已有研究回应了政府的大数据政策部署。其次, 如何提升企业全要素生产率的研究较多, 这些研究肯定了数字基础设施建设、数字化转型对企业全要素生产率的促进作用。再次, 以企业、产业为样本的研究较多, 表明大数据发展和微观企业全要素生产率提升有深刻关联。最后, 既有研究的发现表明, 数据生产要素在优化要素结构上的关键作用, 并且对企业适应外部环境变化提供了保障。

对此, 我们还总结了三个可以进一步研究的空间。第一, 虽然以往研究集中地说明了大数据驱动企业成长的事实, 但是从窥探大数据局部特征衡量大数据的整体发展具有局限性。第二, 大数据不断优化人工智能应用, 虽然以往文献对技术创新的机制进行了探究, 却鲜有实证回应大数据在提升全要素生产率时人工智能这一数字技术影响的作用机制。第三, 虽已有文献关注到数字经济发展的区域性特点, 但是较少学者结合新地理经济和数字经济, 联合行政隔离因素来探讨大数据空间效应。

鉴于此, 本文的研究重点有三方面。第一, 本文将从大数据发展的整体性视角出发, 涵盖更广泛的大数据发展含义, 讨论区域大数据政策对企业微观层面的影响和人工智能的影响的作用机

制。第二，大数据发展需要互联网广泛普及、数字资源累积丰富、数字网络化流通等数字基础，那么多个大数据政策之间是否存在正向关联？以往的数字政策是否与设立试验区有政策联动效应？第三，城市之间的经济发展并不是相互独立的，实际的行政地理空间有没有阻碍不可见的数据流动，进而影响企业全要素生产率？综上，出于政策支持、数据空间流动和区域发展等特点，本文提出研究问题，并从理论和实证分析两方面提供答案。

本研究的边际贡献主要体现在如下方面。第一，本文丰富了以国家级大数据综合实验区为代表的大数据发展与微观企业全要素生产率的研究。第二，本文实证回应了企业应用人工智能是数字化转型、信息化应用等企业各类数字改革的重要体现，也是数据生产要素发挥资源配置作用的重要机制。第三，本文说明了在不同样本下，大数据发展提升企业全要素生产率有差异，位于核心国家级城市群的企业、制造业和中小型企业能够极佳地吸收大数据发展和人工智能应用所带来的发展机遇。第四，本文分析了发展大数据的政策联动效应，进一步认识了我国大数据发展政策的战略性和可持续性。第五，本文探讨了数据跨区域流动的行政隔离现象，从行政地理空间的角度说明了数据的隔离性。

三、理论分析与研究假设

大数据政策可以夯实与构建更加稳固的数据基础制度，从根本上奠定大数据高质量发展的基础，同时，政策的引导力量还能助力大数据和人工智能等数字技术创新。基于夯实基础制度和拔高技术层次，大数据持续地改变宏观经济格局和微观企业生产，实现了数据从网络空间与物理空间的联通，进一步强化了大数据服务于社会经济活动的空间经济特性。基于以上研究情景，本部分从数据基础制度、技术进步和空间大数据驱动力三个方面阐述大数据发展影响企业全要素生产率可能的理论逻辑。

第一，从构建数据基础制度层面来看，大数据政策区域性部署的关键目标在于构建起与数据要素生产力相适应的生产关系，巩固并建设更加系统的数据基础制度，遵循数据市场的基本逻辑，以政策力量有效破除阻碍数据生产、分配、流通、利用等的障碍，进而持续地推进数字化发展和实现数据要素有效的市场化配置。进一步地，数据要素有效配置的核心在于数据与企业生产研发、人才培养、系统管理等方面的深度融合，其不仅能够提升企业大数据能力和技术能力，优化业务服务质量，还能推动强化企业大数据管理能力，从节约成本、缓解数字约束等层面提升企业全要素生产率（谢贤君和郁俊莉，2023）。此外，政府主导的大数据政策通过建立健全适配数据要素特性、遵循市场运行规律的基础制度框架，有效增强了企业外部数据资源的可获得性与多样性，充分释放数据要素在企业生产运营中的价值潜能，最终驱动全要素生产率持续提升。^①

第二，从大数据驱动技术进步的角度来看，其能提升企业生产率。首先，创新是引领发展

^① 国家发展改革委：《加快构建中国特色数据基础制度体系 促进全体人民共享数字经济发展红利》，中华人民共和国国家发展和改革委员会官网，2023-01-01，https://www.ndrc.gov.cn/fzggw/wld/hlf/lddt/202301/t20230101_1345222.html?code=&state=123。

的第一动力，也是企业生产进步的动力源泉。大数据技术的应用显著增强了企业创新活动的靶向性，通过降低技术研发的不确定性与试错成本，有效提升了突破性创新的发生概率与质量水平（Bhattacharya et al., 2017）。其次，大数据的精确性和便利性在企业技术创新上得到了充分体现。依托实时数据流的采集解析与智能算法支撑，大数据技术赋能使企业能够动态捕捉需求端的异质性特征，进而通过模块化定制与敏捷开发实现产品服务创新，为客户提供更多个性化的服务与产品，实现企业价值增值（陈晓红，2018）。此外，企业间研发网络的形成有赖于大数据的发展，技术、知识等沿着研发网络流动，这些要素所形成的技术溢出和重组创新放大了数据价值创造的辐射范围，能够显著促进全要素生产率的提升（Tientao et al., 2016）。

第三，从空间经济学的观点来看，空间经济学对生产要素在城市空间中如何发挥作用、如何分布尤其关心。显性交通运输载体与隐性数据、信息及知识要素的传输与交互，共同印证了地理空间单元间普遍存在的物质交互与经济关联性。相对于实体性交换，数据流动的空间强制性相对较小。并且，数据的空间流动能够缓解创新要素的空间错配，在遵循地理邻近性原则的基础上，合理调节其他要素的空间流动，并对邻近地区产生正向空间溢出效应（彭影和李士梅，2023）。空间经济的发展与内生的要素流动、信息传递、空间品质等紧密联系，空间大数据与经济系统的深度交互遵循地理学第一定律的空间关联特征，呈现明显的空间溢出特性（沈能，2013）。

根据上述分析，本文提出假设 1。

假设 1：大数据发展能够显著提升企业全要素生产率。

本文进一步从两个方面说明人工智能的作用机制。基于已有研究，从企业成长、产业升级和空间演化三个层次分析大数据对人工智能的影响。第一，在企业成长方面。大数据发展释放出大量的数据要素，促进着人工智能的进一步发展，越来越多企业认识到应用人工智能的重要性。在动态变化的数据环境中，人工智能应用能增强大数据分析能力。应用人工智能也能帮助企业优化大数据管理，更好地适应大数据时代。第二，在产业升级上。以产业智能化升级为导向，人工智能被嵌入产业链上、中、下游。人工智能可以充分利用大数据技术收集关键信息，准确预测市场供需变化，进而实现产业升级和产业合理化（Zou & Xiong, 2023）。大数据以人工智能场景创新、技术转化落地为发展导向，让人工智能成为产业升级的有力推手。第三，在空间演化方面。通过收集并分析空间大数据，利用深度学习的感知技术、基于人工智能的决策算法能够不断地取得空间数字建设进展。大数据能驱动人工智能应用挖掘地理信息、改变传统地理资源分布，其不仅在数据分析层面深化了人工智能的空间性内涵，还在经济成果转化方面扩大了人工智能技术的使用范围。

人工智能应用可通过提升技术水平促进全要素生产率的攀升。由于机器学习、深度学习等关键技术的发展，人工智能技术具备了通用技术的特征（Cockburn et al., 2018）。根据产业融合理论，人工智能的通用性必然会使人工智能渗透进经济社会、生产生活各环节，形成技术与产业的深度融合。产业融合能解构、重组产业价值链，进而改变产业价值创造方式和增值能力（宋怡茹等，2017）。人工智能作为先进生产力，渗透在各种生产要素之中，融入劳动生产过程，带动着产业升级和转型，促进着经济增长。同时，基于全要素生产率是经济增长的有效源泉与动力的地位，许多学者证明了人工智能通过技术进步推动全要素生产率的提升。一方面，技术与产业的融

合即人工智能与产业融合发展具有显著的技术创新效应（张龙鹏和张双志，2020）。另一方面，由技术创新、技术引进、技术溢出等产生的技术创新效应显著地提升了全要素生产率（陶长琪和齐亚伟，2010）。

基于此，本文认为人工智能应用是一个重要作用机制。因此，提出假设2。

假设2：人工智能应用是大数据发展促进企业全要素生产率提升的重要机制。

四、研究设计

（一）数据说明

本文选取2007—2019年中国A股所有上市公司为样本，度量上市公司人工智能应用的数据来源于上市公司年报，其他变量度量数据来自国泰安（CSMAR）数据库。本文对数据进行以下预处理：一是为确保研究样本的合理性，剔除了金融保险类公司、ST、*ST、PT公司以及主要变量缺失的公司样本，总体样本量为31054个。二是由于试验区批复时间分为2016年2月与10月，在先前的很长一段时间内，各试验区的统筹准备工作已然进行，为了突出政策效应，考虑政策集中生效时间，本文认为所有的试验区设立冲击时间为2016年。三是为了剔除新冠疫情巨大的外生冲击，本文将样本时间截至2019年。

（二）变量定义及测度

1. 全要素生产率

本文的被解释变量为企业全要素生产率（TFP）。当前对企业全要素生产率测量的方法众多，其中Levinsohn-Petrin方法，即LP半参数法，因其处理灵活而广受青睐，其解决了指标难以寻找的问题而代之以中间品投入测量（Levinsohn & Petrin, 2003）。本文在通过LP法计算企业全要素生产率时，选取上市公司主营业务收入作为产出变量的代理变量，以企业职工数代替劳动投入，用固定资产原值、工程物资与在建工程之和测度资本投入，以购买商品、接受劳务支付的现金表示中间品。最后进行对数化处理，以TFP_LP表示。

2. 大数据发展

本文的核心解释变量是大数据发展。鉴于大数据发展是一个抽象概念，为涵盖更广泛的含义，本文以国家级大数据综合试验区的设立表示大数据发展，以实证分析大数据发展与全要素生产率的关系。综合来看，相对于微观经济主体，政策是一个外生性变量，因此用其分析因果关系更客观；并且以设立试验区衡量大数据发展，本文也从政策设计层面展现了国家发展大数据直接推动社会经济高质量发展的努力。

本文使用双重差分法（difference-in-differences, DID）评估设立国家级大数据综合试验区这一政策冲击对企业全要素生产率影响效应。首先，本文构建实验组与对照组虚拟变量（treat），实验组为设立试验区城市所含上市公司，取值为1；对照组为未设立试验区城市的上市公司，取值为0。其次，本文设计了政策时间虚拟变量（post），2016年至2019年定义为1，2015年及之前定义为0。最后，本文将以上两个虚拟变量交互生成解释变量DID，代表在当年是否设立试验

区，设立为 1，未设立为 0。

3. 人工智能应用

本文的作用机制变量为人工智能应用（AI）。目前，针对人工智能的测量在不同领域测度方式各异，如工业产品出口中以工业机器人的使用度量。但有学者认为采用机器人数据刻度的局限性较大，无法透视人工智能技术的演变规律，而宜采用上市公司年报文本数据的“人工智能”相关词频表示（吴非和徐斯旸，2022），即基于有关上市公司人工智能应用的研究，统计上市公司年报文本中出现与人工智能相关的关键词的次数作为上市公司人工智能应用程度的代理变量，词频高低代表上市公司当年的人工智能应用程度高低。这些关键词已经能够在很大程度上反映企业人工智能应用情况。例如，“智能”关键词后面通常会紧跟“制造”“生产”“仓储”“物流”“管理”“终端”等关键词，进而表明人工智能在这些企业活动领域的应用。因此，基于张龙鹏和张双志（2020）的研究，本文从智能、云、数据、物联、机器学习五个维度提取与人工智能相关的关键词，从广义层面度量企业的人工智能应用程度以检验稳健性，记作 AI1。

4. 控制变量

为了缓解遗漏重要解释变量所带来的内生性问题，本文控制了以下变量：

①企业年龄（Age），用当年年份减去企业注册年份得到，并将数值加 1 后取自然对数处理；②企业规模（Size），企业规模为平均每个员工所拥有的资产额，取自然对数处理；③产权性质（SOE），分为国有企业与非国有企业，若上市公司的最终控制人为国有，则产权性质赋值为 1，否则为 0；④财务杠杆（Lev），该指标以总负债占总资产的比重衡量；⑤大股东持股比例（Top1），用上市公司第一大股东持股比例衡量。

变量的描述性统计见表 1。在上市公司样本中，使用 LP 法测算的企业全要素生产率（TFP）的极值是 12.518，说明 TFP 的整体波动范围较大。两种人工智能应用（AI 与 AI1）测量指标的标准差分别为 38.062 和 59.021，说明各企业人工智能应用的差距较大。

表1 变量的描述性统计

变量	样本量（个）	均值	标准差	最小值	最大值
全要素生产率（TFP_LP）	31016	13.308	1.012	5.731	18.249
人工智能应用（AI）	31054	12.491	38.062	0	826
人工智能应用（AI1）	29992	34.445	59.021	0	1064
企业年龄（Age）	31054	2.834	0.347	0	4.174
企业规模（Size）	31022	14.438	1.022	9.719	22.334
产权性质（SOE）	30609	0.393	.488	0	1
财务杠杆（Lev）	31054	0.429	.211	0.007	2.128
大股东持股比例（Top1）	31053	35.325	15.255	0.29	100

（三）基准回归的模型设定

国家级大数据综合试验区的设立，目的在于实现跨区域协同、区域内建示范以及夯实大数据基础设施统筹发展，对于企业全要素生产率而言具有相对外生性。八大试验区于2016年被集中批复，被纳入实验组的上市公司分布在研究区间的相同年度，因此为验证试验区设立对企业全要素生产率的影响，本文将国家级大数据综合试验区的设立视为准自然实验，构建双重差分模型来检验假设1，设定模型如下。

$$TFP_LP_{it} = \beta_0 + \beta_1 Treat_i \times Post_t + Controls_{it} + Firm_i + Year_t + \mu_{it} \quad (1)$$

（1）式中， i 表示上市公司， t 为年份， TFP_LP 表示LP法测算的企业全要素生产率，试验区设立的虚拟变量 $Treat_i$ 与事件冲击的虚拟变量 $Post_t$ 交互生成政策变量（DID）， $Controls$ 表示控制变量集， $Firm_i$ 表示不随时间变动的企业个体效应， $Year_t$ 则是不随个体变动的的时间效应， μ_{it} 为误差项。

五、实证结果分析

（一）基准回归

基于普通最小二乘法，表2显示了基准回归结果。第（1）（2）列均加入了企业固定效应与年份固定效应，第（2）列在前者的基础上加入了控制变量。第（2）列结果表明，设立国家级大数据综合试验区对上市公司全要素生产率的回归系数为0.067，且在1%的水平上显著，说明国家级大数据综合试验区的设立对上市公司全要素生产率存在显著正向影响，即大数据发展能显著提升企业全要素生产率，假设1得证。

表2 基准回归

变量	TPF_LP	
	(1)	(2)
<i>DID</i>	0.066*** (0.022)	0.067*** (0.021)
<i>Age</i>		0.201*** (0.074)
<i>Size</i>		0.170*** (0.014)
<i>SOE</i>		0.058 (0.046)
<i>Lev</i>		0.419*** (0.057)
<i>Top1</i>		0.000 (0.001)

(续表)

变量	(1)	(2)
	<i>TPF_LP</i>	
<i>Constant</i>	12.920*** (0.015)	9.866*** (0.275)
年份固定效应	是	是
企业固定效应	是	是
观测值(个)	31016	30570
拟合优度	0.197	0.237

注：1) *** 代表 1% 的显著性水平；2) 括号内为估计系数的稳健标准差（下同）。

（二）倾向得分匹配与双重差分估计（PSM-DID）

虽然设立试验区相对企业全要素生产率充分外生，但是如果试验区与非试验区的上市公司在某些不可观测因素上存在显著差异，并随着时间变化而呈现出不同趋势，那么 OLS 不能缓解内生选择产生的影响。也就是说，企业进行数字化改革的可能性或接受大数据发展的程度和效应与企业随时间变动的因素相关时，长期趋势相同的要求很难满足。为了处理这一内生性问题，本文参考了王庶和岳希明（2017）的处理方式，进行随机试验，采用倾向得分匹配与双重差分估计结合的方法估计试验区设立的事前全要素生产率变动趋势。

为了估计各企业全要素生产率受大数据发展影响的可能性，本文首先给试验区内的上市公司匹配非试验区反事实个体，得到倾向得分。图 1 表示处理组和控制组受大数据发展影响的倾向得分共同支撑域占比较大，可以进一步为处理组匹配反事实个体。图 2 则是两个组别的关键变量在匹配前后的比较：匹配前，处理组与控制组的协变量均存在显著差异；匹配后，组别协变量间的差异已不显著。这说明倾向匹配法较好地降低了处理组和控制组在试验区设立前可观测特征的差异。图 3 和图 4 则是采用最近邻匹配法绘制了两组样本匹配前后的核密度分布图，匹配之后的处理组和控制组的拟合效果吻合，分布形态更加接近。表 3 是倾向得分匹配之后对样本数据的整体平衡性检验结果，显示通过检验。

PSM-DID 的结果如表 4 所示，相比于非试验区企业，试验区企业会因大数据发展相对增加全要素生产率提升概率 6.7 个百分点；即使采取不同邻近数，只有稳健标准差有所波动，估计结果的大小、符号和显著性与基准回归结果一致。PSM-DID 分析结果表明了平衡条件检验通过，且基准结果可靠。

表3 倾向得分匹配中平衡条件的整体检验

	Ps R2	LR chi2	$P > \chi^2$	MeanBias	MedBias	<i>B</i>	<i>R</i>	%Var
匹配前	0.018	718.52	0.000	12.6	13.1	31.7*	1.45	100
匹配后	0.000	4.91	0.427	1.0	1.2	3.0	0.91	50

表4 大数据发展对企业全要素生产率的影响——基于PSM-DID的估计

指标	基准结果	近邻匹配稳健性
<i>DID</i>	0.067*** (0.021)	0.067*** (3.217)
控制变量	是	是
年份固定效应	是	是
企业固定效应	是	是
观测值(个)	30570	30606
拟合优度	0.237	0.237

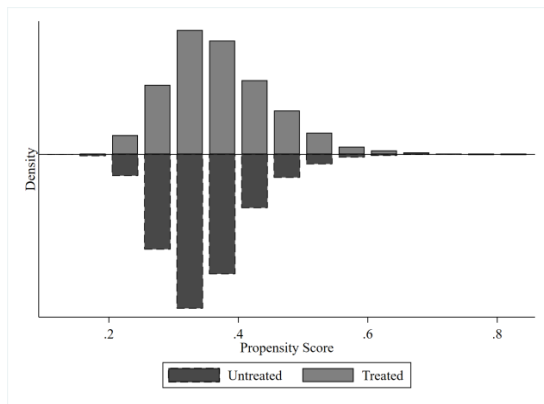


图1 处理组和控制组倾向得分比较

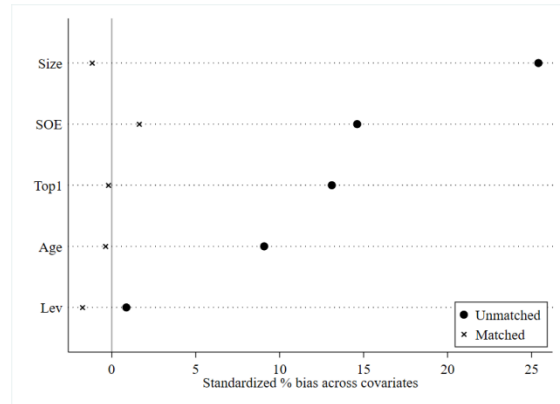


图2 协变量在处理组与控制组匹配前后的平衡性检验

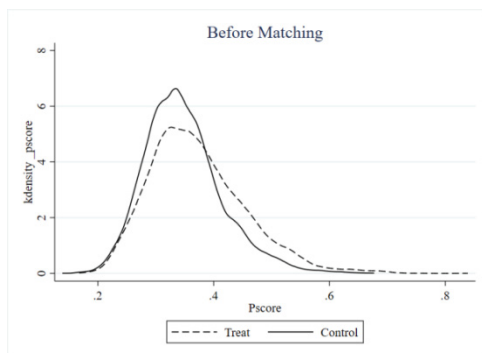


图3 匹配前的核密度函数图

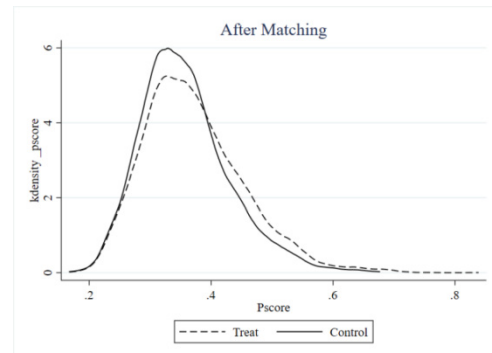


图4 匹配后的核密度函数图

(三) 稳健性检验

1. 安慰剂检验

虽然我们采用了 OLS 以及倾向得分匹配后的双重差分来进行估计，但还需要排除随时间变化的企业特征影响。首先，本文把置换检验作为基本思想进行安慰剂检验，其本质是通过置换样本顺序，重新计算统计检验量并构造经验分布，利用样本数据的随机排列进行统计推断，以检验其稳健性。图 5 结果表明，虚拟估计系数分布在零附近，服从正态分布，且真实估计系数 0.067 未被虚拟区域覆盖，符合安慰剂检验预期。此外，我们还进行了虚拟生成企业样本的空间安慰剂检验（见图 6），还有先随机抽取“伪处理样本”后，再随机抽取统一的“伪冲击时间”的混合安慰剂检验（见图 7、图 8），以上操作均是对虚拟过程重复估计 500 次。图 6、图 7、图 8 中的处理效应估计值（0.067）均是一个极端值，检验结果与图 5 的保持一致，安慰剂检验通过。

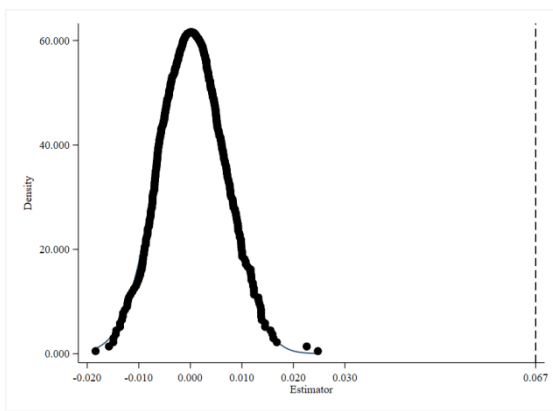


图5 置换检验

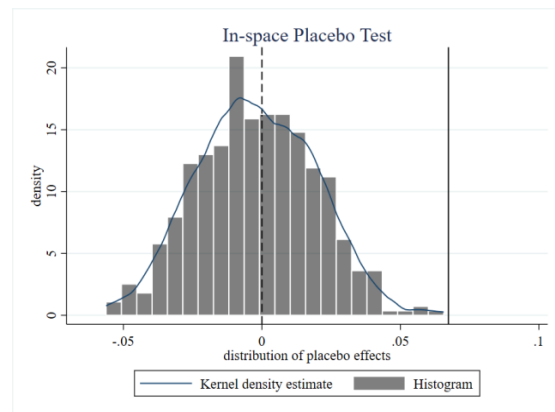


图6 空间安慰剂检验

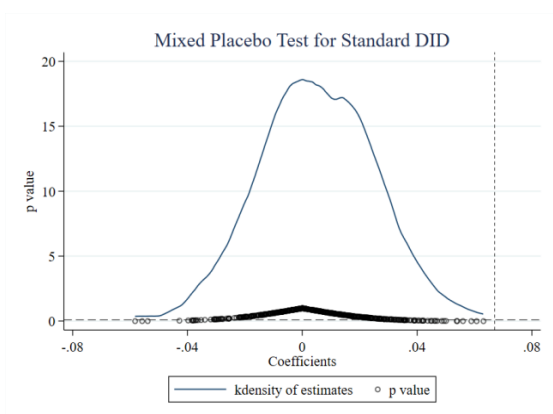


图7 混合安慰剂检验1

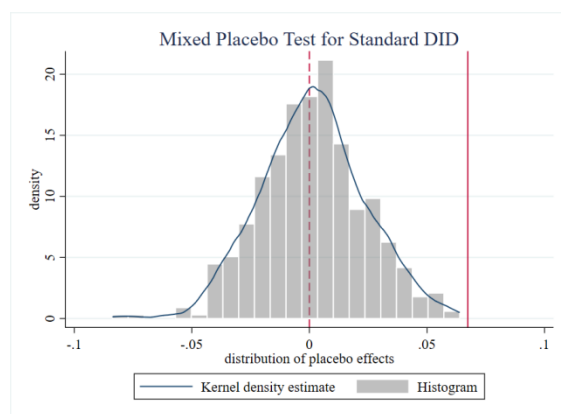


图8 混合安慰剂检验2

2. 改变政策冲击时间的检验

为进一步检验实证结果的稳健性，采用改变政策冲击时间的方式进行反事实检验。因此，将

设立时间滞后一期，以伪政策冲击来探索回归结果是否显著，即考察未设立试验区时，企业全要素生产率变化是否显著。表5的第(1)列表明，将政策实施时间滞后一期的回归结果不显著，符合反事实假定。

3. 工具变量回归

为了排除由于双向因果导致的内生性问题，本文需要进一步进行工具变量回归以重新审视大数据发展对企业全要素生产率的影响。由于本研究使用的数据横截面数量和时间序列均较长，我们参考了Norkute et al. (2021)和Cui et al. (2022)开发的一种通用工具变量方法，称作“2SIV”，用来估计具有未观测共同因子的面板回归模型。相对于常见的2SLS，该方法的优势主要在于：第一，该算法不受偶然参数偏差的影响，尤其是不受高阶偏差的干预，有效解决了样本中可能会出现的大尺度失真；第二，允许估计异质斜率系数的面板回归模型；第三，只要外部工具变量是可用的，便允许灵活地指定工具变量。

此处构建如式(2)所示的2SIV估计模型，以进行工具回归分析。

$$TFP_LP_{it} = \beta_0 + \alpha TFP_LP_{it-1} + \beta_1 Treat_i \times Post_t + \gamma Controls_{it} + Firm_i + Year_t + \mu_{it} \quad (2)$$

式(2)中， TFP_LP_{it-1} 为被解释变量企业全要素生产率的滞后一期项； μ_{it} 是一个复合误差项，包含了获取特定个体、特定时点的效应等误差。模型中的其余变量的指标设计与基准回归模型(1)一致。

基于该模型，我们采用2SIV进行工具变量回归，通过生成去因子化协变量的不同滞后项作为工具变量。由于本研究选取了6个自变量，所以一共生成了 $6 \times 3 = 18$ 个工具变量。结果如表5第(2)列所示， $L1.TPF_LP$ 回归系数为0.266，核心回归系数为0.045，都在1%的水平上显著。同时，Hansen检验统计量的 $P=0.000$ ，接受原假设，模型拒绝过度识别假设，即说明由不同滞后项生成的工具变量具有外生性，工具变量有效。

4. 替换被解释变量的检验

前文指出以LP法测算的全要素生产率，是对Olley & Pakes (1992)估计方法的改进，但是现代研究并未摒弃使用OP估计法，其仍是重要的全要素生产率估计方法。因此，本文采用OP方法对前文基准回归结果进行稳健性检验。表5的第(3)列显示了相应的估计结果，设立国家级大数据综合试验区的估计系数显著为正，表明重新使用OP方法估计全要素生产率，基准回归结果仍然成立。

表5 稳健性检验

变量	(1)	(2)	(3)
	滞后一期	2SIV 工具变量回归 TPF_LP	替换被解释变量 TPF_OP
DID	0.033 (0.022)	0.045*** (0.017)	0.040** (0.018)
$L1.TPF_LP$		0.266*** (0.053)	

(续表)

变量	(1)	(2)	(3)
	滞后一期	2SIV 工具变量回归	替换被解释变量
		<i>TPF_LP</i>	<i>TPF_OP</i>
控制变量	是	是	是
年份固定效应	是	是	是
企业固定效应	是	是	是
观测值	30606	20676	30606
拟合优度	0.258		0.338

注:(1)**、*** 分别代表 5% 和 1% 的显著性水平,括号内为估计系数的稳健标准差。(2) 过度识别检验结果为 $\chi^2(11)=107.8234$, $\text{Prob} > \chi^2 = 0.0000$ 。

(四) 作用机制检验

理论分析表明,人工智能应用是试验区设立促进企业全要素生产率提升的一个重要机制。在表 6 第(1)列中,国家级大数据综合试验区的设立对人工智能应用的回归系数为 4.487,并通过 1% 显著性水平的检验;第(2)列中,人工智能应用对企业全要素生产率的估计系数为 0.002,并在 1% 的水平上显著。进一步地,本文替换人工智能应用的度量方式,结果如第(3)列所示,试验区设立的估计系数为 14.472,并在 1% 水平上显著;第(4)列中,人工智能应用在 1% 的水平上显著促进了企业全要素生产率的提升,通过稳健性检验。因此,实证结果表明国家级大数据综合试验区的设立能够通过促进人工智能应用提升企业全要素生产率,假设 2 成立。

表6 作用机制检验

变量	(1)	(2)	(3)	(4)
	<i>AI</i>	<i>TFP_LP</i>	<i>AII</i>	<i>TFP_LP</i>
<i>DID</i>	4.487*** (1.235)		14.472*** (2.320)	
<i>AI</i>		0.002*** (0.000)		
<i>AII</i>				0.001*** (0.000)
控制变量	是	是	是	是
年份固定效应	是	是	是	是
企业固定效应	是	是	是	是
城市—年份固定效应	否	否	否	否
行业—年份固定效应	否	否	否	否
观测值(个)	30176	30172	29126	29122
拟合优度	0.845	0.889	0.814	0.892

六、异质性检验

基于上市公司研究样本，基准回归以及稳健性检验表明国家级大数据综合试验区的设立显著提升了上市公司全要素生产率；影响机制检验则表明人工智能应用是试验区设立促进全要素生产率提升的关键路径。在此部分，本文将对异质性予以检验。

（一）注册地异质性

大数据发展对不同注册地的企业可能存在差异。有学者结合城市群发展研究企业布局问题，考察了大数据发展中工业智能化对企业分布产生选择性偏向，不同城市的企业地理格局会呈现不同程度的政策效应，中心层次的国家城市群往往会集聚企业发展所需的重要数字资源（王林辉等，2022）。企业的区位因素和空间地理位置在大数据发展促进企业全要素生产率的过程中扮演重要角色。目前我国已规划布局19个国家级城市群，其中京津冀、长三角、粤港澳、成渝与长江中游城市群在所有城市群发展中，区域发展收效显著。上海社会科学院信息研究所与社会科学文献出版社共同发布的《全球信息社会蓝皮书·全球信息社会发展报告（2021）》指出，人工智能的国家竞争，本质是城市或城市群竞争。2022年7月，科技部等六部门印发《关于加快场景创新以人工智能高水平应用促进经济高质量发展的指导意见》，鼓励京津冀、长三角、粤港澳大湾区等城市群探索建立人工智能场景创新共同体，发挥中心城市科技辐射带动作用。由此，本文将进一步考察在国家级大数据综合试验区范围内的上市公司，受到国家级城市群的辐射效应而产生的异质性。

在表7中，第（1）（3）列分别为企业注册地在五大城市群与非五大城市群的异质性回归结果。该结果显示：五大城市群内的样本企业，其核心解释变量的回归系数为0.425，并在1%的水平上显著，而非五大城市群的样本企业回归结果不显著。因此可以认为，与非五大城市群的结果相比较，试验区内的上市公司被“试验区+城市群”政策红利惠及，资源聚集而显示出强劲的全要素生产率提升效应，产生了空间红利。此外，第（2）和（4）列进一步分别验证人工智能应用的作用机制，五大城市群内的样本均通过1%显著性检验，作用机制也成立。

（二）企业规模异质性

大数据发展在大规模企业和中小规模企业中提升全要素生产率的水平可能存在差异。大企业以其丰富的数据更大可能地提升其市场价值，其雄厚的技术优势可以形成信息与通信技术的企业生产率效应，而中小规模企业的生产经营和对外信息不对称更严重。我们有必要实证考察是大规模企业还是中小规模企业在大数据发展中获取更多机遇。

根据2019年各上市公司的员工人数，按照样本企业是否高于样本企业平均员工数量，划分了大企业组与中小企业组。表7列（5）至列（8）是企业规模所产生的异质性效应检验结果。第（5）列的大企业组结果不显著；第（7）列的中小企业组中的估计系数为0.084，且在1%的水平上显著；第（8）列显示中小企业组通过机制检验。以上表明，中小企业组的估计结果显著正相关，但是大企业组不显著，即大数据发展对提升中小企业全要素生产率的作用更显著。大数据发展对于规模更小的企业而言，为其增添了获取与预测信息的精度。在市场经济中，规模小的企

业更能从大数据发展中实现企业价值的提升，因此，大数据和人工智能技术应用于中小规模企业时显现出强烈的企业全要素生产率提升效应。

（三）行业异质性

企业的竞争力维度随着时代发展而不断创新，制造业追求产品升级、提高综合实力的关键途径之一是向服务化升级转型，制造业和服务业成为关注点。国务院更是在《关于加快推动制造业服务业高质量发展的意见》中指出，要利用新一代信息技术促进企业变革。数据要素禀赋差异在制造业和服务业间有差别。制造业往往通过大数据实现信息化改革，生产更精确、先进和优质的产品，进而反映在产业效益上；服务业也抓住大数据契机，拓展金融服务、数字创意等新兴服务产业，这体现为产业模式创新。因此，本文依照《上市公司行业分类指引（2012年修订）》，关注了制造业和服务业在国家级大数据综合试验区设立情境下的企业全要素生产率提升情况。^①

在表7中，第（9）（11）列显示了试验区设立影响上市公司全要素生产率的行业异质性效应检验结果。在制造业样本中，试验区设立的估计系数为0.064，并在1%的水平上显著，且（10）（12）列结果表示作用机制成立；而服务业样本回归结果不显著。主要原因在于，大数据发展有针对性地增强了制造业的先进制造能力，提升了制造企业全要素生产率水平，其数字政策吸收和转化效果显著高于服务业。

异质性检验表明，大数据发展更能提升注册地在五大城市群内的企业、中小规模企业、制造业企业的全要素生产率。

表7 异质性检验

变量	五大城市群		非五大城市群		大企业组		中小企业组		制造业		服务业	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	异质性	机制	异质性	机制	异质性	机制	异质性	机制	异质性	机制	异质性	机制
	<i>TFP_LP</i>	<i>AI</i>	<i>TFP_LP</i>	<i>AI</i>	<i>TFP_LP</i>	<i>AI</i>	<i>TFP_LP</i>	<i>AI</i>	<i>TFP_LP</i>	<i>AI</i>	<i>TFP_LP</i>	<i>AI</i>
DID	0.425*** (0.091)	13.807*** (3.968)	0.374 (0.347)	7.738 (12.638)	-0.009 (0.033)	1.637 (2.444)	0.084*** (0.026)	5.221*** (1.491)	0.064*** (0.023)	6.915*** (1.825)	0.029 (0.043)	2.769 (2.220)
观测值(个)	9874	9876	1016	1017	7191	7191	21604	21608	18499	18500	9405	9410
拟合优度	0.267	0.163	0.254	0.166	0.412	0.173	0.217	0.134	0.271	0.170	0.243	0.108
控制变量	是	是	是	是	是	是	是	是	是	是	是	是
年份固定效应	是	是	是	是	是	是	是	是	是	是	是	是
企业固定效应	是	是	是	是	是	是	是	是	是	是	是	是

^① 中国证券监督管理委员会：《上市公司行业分类指引（2012年修订）》。

七、进一步分析

（一）政策联动效应

大数据发展离不开完善的数字基础设施、广阔的数字消费市场。2013年被称为中国“大数据元年”，我国互联网行业进入移动端时代，网民数量大规模增长，企业的业务触角开始向大数据领域蔓延，拉动了数字经济快速成长。大数据的收集、存储与分析依赖数字基础设施的不断完善。2013年8月所启动的“宽带中国”战略为大数据向更高程度发展打下了坚实基础。^①宽带网络是各国部署信息化的优先战略行动，是中国推动科技和产业发展的基础性行动。宽带网络普及带来海量数据处理需求，不断推进数字化、网络化和智能化深入发展，极大地刺激着企业采用大数据技术，以焕发更强的数字创新活力，提升了全要素生产率（刘传明和马青山，2020）。同时，工信部在2013年年末公布了首批国家信息消费试点城市，旨在进一步扩大与升级数字消费。^②信息消费市场积累了大量的大数据资源，推动了信息消费相关产业升级和市场发展，进而实现经济高质量发展。基于上述分析，本文拟讨论国家大数据综合试验区与“宽带中国”战略、国家信息消费城市试点政策的联动对企业全要素生产率的影响。

本文构建式（3）所示的三重差分模型（DDD）以实证分析国家大数据综合试验区政策与其他政策的联动效应。

$$TFP_LP_{it} = \alpha + \beta DID_{it} \times RP_{it} + \gamma Controls_{it} + Firm_t + Year_t + \mu_{it} \quad (3)$$

在式（1）的基础上，式（3）加入了“宽带中国”战略与国家信息消费试点政策的哑变量 RP ，用来识别所属城市是否为“宽带中国”战略试点城市和国家信息消费试点城市，如果是， RP 定义为 1，否则为 0。模型中的其他变量的指标设计与基准回归模型一致。

表 8 为分析政策联动效应的估计结果。第（1）列和第（2）列为国家大数据综合试验区政策与“宽带中国”战略的估计结果， $DID \times RP$ 的核心估计系数显著为正，表明人工智能应用机制检验通过；第（3）列和第（4）列是国家大数据综合试验区政策与国家信息消费试点政策的估计结果， $DID \times RP$ 的估计系数也显著为正，表明机制检验通过。表 8 的回归结果表明，“宽带中国”战略所带来的数字基础设施的完善、国家信息消费试点政策所带来的数字消费市场的扩大都能够与国家大数据综合试验区政策形成政策联动效应，通过强化企业人工智能应用，进而提升了企业全要素生产率。

^①《国务院关于印发“宽带中国”战略及实施方案的通知 国发〔2013〕31号》，中国政府网，2013-08-17，https://www.gov.cn/zwqk/2013-08/17/content_2468348.htm。

^②《首批国家信息消费试点市（县、区）名单公示》，中华人民共和国工业和信息化部官网，2013-12-31，https://www.miit.gov.cn/zwqk/wjgs/art/2020/art_dfd24dee76c3460c928c653a809d53e9.html。

表8 政策联动效应

变量	(1)	(2)	(3)	(4)
	宽带中国		信息消费	
	<i>TFP_LP</i>	<i>AI</i>	<i>TFP_LP</i>	<i>AI</i>
<i>DID</i> × <i>RP</i>	0.048** (0.024)	2.725* (1.433)	0.065*** (0.024)	3.272** (1.405)
观测值	20881	20884	20266	20270
拟合优度	0.236	0.062	0.235	0.063
控制变量	是	是	是	是
企业固定效应	是	是	是	是
年份固定效应	是	是	是	是

(二) 行政隔离效应

在数据要素流通体系尚不完善、参与主体缺乏规范化监管、数据共享难等形势下，打通数据流通壁垒、治理“数据孤岛”成为大数据发展的时代要求。不同数据系统间存在利益竞争，随着互联网环境状态更新或者技术发展阶段跃进，就会形成不同级别的数字隔离或数字集成。此外，数据事务本身也具有一定的隔离性，信息和数据流通阻力通常涉及企业、政府等部门。既有的制度路径使得数据难以做到比较完善地平衡利益，数据的互联互通和统筹管理面临跨区域协调能力不足、资源调度困难和融合对接失序等问题。数据在空间上的网络张力需要进一步被拉开，才能实现公共数据和社会数据的融合应用。

破除共享大数据发展的阻碍，塑造数据流动的有利空间可以成为一个突破口。从空间溢出效应的基本特征来看，省域层面的行政阻力应被首要考察。根据大数据空间溢出效应的结果，本部分以非试验区与近邻试验区所含城市是否在同一省份为分组标准，划分了同省份组和跨省份组，来探讨大数据发展的行政隔离现象。其中，同省份组指非试验区与近邻试验区在同一个省份；跨省份组表示非试验区与近邻试验区在不同省份。表9第(1)(2)列和(3)(4)列分别为同省份组和跨省份组，核心解释变量回归系数分别是0.079和0.065，分别通过了5%和1%的显著性检验；并且两组机制检验均通过，都在1%的显著性水平上正相关。通过对比发现，同省份组估计系数大于跨省份组，且同省份组的拟合优度更高。也就是说，当非试验区城市与近邻试验区在同一省级行政区时，大数据发展促进企业全要素生产率的效应强于不处于同一省级行政区。以上结果表明，大数据发展存在一定程度的行政隔离，阻碍着数据要素跨省域流动，进而制约了企业全要素生产率的提升。

表9 大数据发展的行政隔离效应

变量	(1)	(2)	(3)	(4)
	同省组		跨省组	
	<i>TFP_LP</i>	<i>AI</i>	<i>TFP_LP</i>	<i>AI</i>
<i>DID</i>	0.079** (0.039)	6.406*** (1.519)	0.065*** (0.021)	4.197*** (1.265)
控制变量	是	是	是	是
年份固定效应	是	是	是	是
企业固定效应	是	是	是	是
观测值(个)	13213	13216	28240	28246
拟合优度	0.254	0.019	0.238	0.031

八、结论与启示

(一) 研究结论

本文使用2007年至2019年中国A股上市公司的面板数据，以设立国家级大数据综合试验区衡量大数据发展，实证研究了大数据发展对企业全要素生产率的影响效应和作用机制，为后续的政策实施和企业发展提供一定的参考价值。研究表明，第一，大数据发展显著地促进了企业全要素生产率的提升，通过了PSM-DID、安慰剂检验、改变政策冲击时间、两阶段工具变量回归和替换被解释变量的检验；第二，机制研究表明，这一促进效应是通过人工智能应用实现的；第三，异质性检验表明，大数据发展更能提升注册地在五大城市群内的企业、中小规模企业、制造类企业的全要素生产率；第四，设立国家级大数据综合试验区与宽带中国示范城市和国家信息消费城市建设具有同向的政策联动效应；第五，大数据发展存在一定程度的行政隔离，数据要素跨省域流动受阻，进而削弱了企业全要素生产率的提升效应。

(二) 政策启示

第一，为充分发挥大数据对企业全要素生产率的赋能作用，企业需在市场竞争中加强关注有价值的大数据机遇和人工智能技术机会。首先，相关部门应洞察市场的大数据功能需求，推动企业建立数据采集、治理与共享的全方位管理体系，强化高数据价值的结构化处理与多源异构数据融合的能力。其次，相关部门引导企业依托大数据分析精准定位研发、生产、管理等环节的效率痛点，深化人工智能场景渗透，建立大数据、人工智能有效对接企业的业务框架。最后，相关部门应利用企业内外部的数字生态不断优化企业人工智能应用，一方面破除因大数据资源使用不当的低效率瓶颈，补齐短板、直击发展痛点；另一方面要努力在数据要素使用、人工智能技术应用和提升企业全要素生产率之间锻造可持续性，实现技术赋能与企业效率变革的良性循环。

第二，国家级大数据综合试验区的设立虽然促进了制造业全要素生产率的提升，却对服务业全要素生产率提升的作用不显著。首先，市场参与主体要重塑服务业要素禀赋结构，重点建设服

务业数据中台，强化客户行为、服务交互等非结构化数据的采集与价值挖掘能力，推动数据要素与知识资本深度融合，破解服务业高度依赖人力和低标准化的禀赋约束。其次，市场参与主体要改造服务业人工智能应用场景，通过开发人工智能工具包优化实现服务流程标准化与个性化需求的动态平衡，应探索人工智能在服务业中的协同使用，逐渐治愈服务业的“鲍莫尔病”。最后，市场参与主体依托试验区的数据政策优势，打造数智化的服务价值链，同步构建服务业人工智能应用框架，深化推动产业融合发展，进而提升全要素生产率。

第三，持续完善大数据发展政策的实施细节，减少由于行政空间产生的数据流动阻隔。各地方政府在统筹区域数字发展战略或者其他区域政策的过程中，要综合考虑自身区域和外部区域的溢出效应对本地区发展可能产生的影响，建立跨区域数据治理协同机制。一方面，政府应积极联动近邻城市，打破行政阻碍，开发大数据合作模式，探索数据跨区域高效流通的大数据发展机制，打通数据跨区域流通的“最后一公里”。另一方面，政府应实施差异化的大数据配置策略，对现有的大数据发展结构进行评估，依托区域内生优势和外生前景，渐进地实现企业全要素生产率的攀升。在大数据政策试验区域，政府可考虑培育数据驱动的产业链协同生态，加强大数据平台与企业产品制造以及服务环节的互联互通；针对大数据发展相对落后的地区，政府可以通过数据托管的模式实现技术赋能，进而激活本地特色产业发展潜能。此外，政府可以有针对性地建立企业全要素生产率动态监测平台，结合数据要素投入强度、人工智能应用等核心指标，分行业制定生产率提升路径图，实现大数据资源、创新技术和产业升级之间更为精准的匹配。

参考文献

- 曹正旭、董会忠、韩沅刚，2020：《工业集聚对全要素生产率影响机理及区域异质性研究》，《软科学》第9期。
- 陈晓红，2018：《数字经济时代的技术融合与应用创新趋势分析》，《中南大学学报（社会科学版）》第5期。
- 程学旗、靳小龙、王元卓、郭嘉丰、张铁赢、李国杰，2014：《大数据系统和分析技术综述》，《软件学报》第9期。
- 范合君、吴婷，2022：《新型数字基础设施、数字化能力与全要素生产率》，《经济与管理研究》第1期。
- 郭炳南、王宇、张浩，2022：《数字经济发展改善了城市空气质量吗——基于国家级大数据综合试验区的准自然实验》，《广东财经大学学报》第1期。
- 花俊国、刘畅、朱迪，2022：《数字化转型、融资约束与企业全要素生产率》，《南方金融》第7期。
- 黄永春、宫尚俊、邹晨、贾琳、徐子飞，2022：《数字经济、要素配置效率与城乡融合发展》，《中国人口·资源与环境》第10期。
- 荆文君、孙宝文，2019：《数字经济促进经济高质量发展：一个理论分析框架》，《经济学家》第2期。
- 李雪、吴福象、竺李乐，2021：《数字经济与区域创新绩效》，《山西财经大学学报》第5期。
- 刘传明、马青山，2020：《网络基础设施建设对全要素生产率增长的影响研究——基于“宽带中国”试点政策的准自然实验》，《中国人口科学》第3期。
- 刘建翠，2022：《中国的全要素生产率研究：回顾与展望》，《技术经济》第1期。

- 刘凌、罗戎, 2017:《大数据视角下政府数据开放与个人隐私保护研究》,《情报科学》第 2 期。
- 吕可夫、于明洋、阮永平, 2023:《企业数字化转型与资源配置效率》,《科研管理》第 8 期。
- 彭影、李士梅, 2023:《创新要素流动与城市绿色创新发展——数据要素流动环境的空间调节作用》,《科技进步与对策》第 1 期。
- 彭远怀, 2023:《政府数据开放的价值创造作用:企业全要素生产率视角》,《数量经济技术经济研究》第 9 期。
- 沈能, 2013:《局域知识溢出和生产性服务业空间集聚——基于中国城市数据的空间计量分析》,《科学学与科学技术管理》第 5 期。
- 司林波、刘畅、孟卫东, 2017:《政府数据开放的价值及面临的问题与路径选择》,《图书馆学研究》第 14 期。
- 宋怡茹、魏龙、潘安, 2017:《价值链重构与核心价值区转移研究——产业融合方式与效果的比较》,《科学学研究》第 8 期。
- 陶长琪、齐亚伟, 2010:《中国全要素生产率的空间差异及其成因分析》,《数量经济技术经济研究》第 1 期。
- 涂心语、严晓玲, 2022:《数字化转型、知识溢出与企业全要素生产率——来自制造业上市公司的经验证据》,《产业经济研究》第 2 期。
- 王林辉、姜昊、董直庆, 2022:《工业智能化会重塑企业地理格局吗》,《中国工业经济》第 2 期。
- 王庶、岳希明, 2017:《退耕还林、非农就业与农民增收——基于 21 省面板数据的双重差分分析》,《经济研究》第 4 期。
- 吴非、徐斯旸, 2022:《人工智能技术应用与上市企业市场价值》,《现代经济探讨》第 11 期。
- 谢康、夏正豪、肖静华, 2020:《大数据成为现实生产要素的企业实现机制:产品创新视角》,《中国工业经济》第 5 期。
- 谢谦、郭杨, 2022:《数字技术、创新要素结构优化与企业全要素生产率》,《北京师范大学学报(社会科学版)》第 6 期。
- 谢贤君、郁俊莉, 2023:《大数据如何影响企业全要素生产率——来自〈促进大数据发展行动纲要〉实施的准自然试验》,《当代经济管理》第 8 期。
- 杨德明、刘泳文, 2018:《“互联网+”为什么加出了业绩》,《中国工业经济》第 5 期。
- 殷红、张龙、叶祥松, 2020:《中国产业结构调整对全要素生产率的时变效应》,《世界经济》第 1 期。
- 张莉、林安然, 2023:《中国地方政府数据开放与经济发展质量》,《公共行政评论》第 2 期。
- 张龙鹏、张双志, 2020:《技术赋能:人工智能与产业融合发展的技术创新效应》,《财经科学》第 6 期。
- 张龙鹏、钟易霖, 2023:《价值链视角下人工智能应用对全要素生产率的影响——基于中国 A 股上市公司的实证研究》,《经济体制改革》第 4 期。
- 张益豪、郭晓辉, 2023:《大数据发展与企业全要素生产率——基于国家级大数据综合试验区的实证分析》,《产业经济研究》第 2 期。
- 赵宸宇、王文春、李雪松, 2021:《数字化转型如何影响企业全要素生产率》,《财贸经济》第 7 期。
- Aversa, P., Haefliger, S., and Reza, D. G., 2017, “Building a Winning Business Model Portfolio”, *MIT Sloan Management Review*, 58(4), 49–54.
- Bharadwaj, A. S., 2000, “A Resource-Based Perspective on Information Technology Capability and Firm

Performance: An Empirical Investigation” , *MIS Quarterly*, 24(1), 169–196.

Bhattacharya, U., Hsu, P. H., Tian, X., and Xu, Y., 2017, “What Affects Innovation More: Policy or Policy Uncertainty?” , *Journal of Financial and Quantitative Analysis*, 52(5), 1869–1901.

Chen, H., Wang, S., and Song, M., 2021, “Global Environmental Value Chain Embeddedness and Enterprise Production Efficiency Improvement” , *Structural Change and Economic Dynamics*, 58, 278–290.

Cockburn, I. M., Henderson, R., and Stern, S., 2018, “The Impact of Artificial Intelligence on Innovation: An Exploratory Analysis” , *Chicago: University of Chicago Press*.

Cui, G., Norkutė, M., Sarafidis, V., and Yamagata, T., 2022, “Two–Stage Instrumental Variable Estimation of Linear Panel Data Models with Interactive Effects” , *The Econometrics Journal*, 25(2), 340–361.

Davenport, T. H., Barth, P., and Bean, R., 2021, “How ‘Big Data’ is Different” , *MIT Sloan Management Review*, 54(1), 43–46.

Dedrick, J., Kraemer, K. L., and Shih, E., 2013, “Information Technology and Productivity in Developed and Developing Countries” , *Journal of Management Information Systems*, 30(1), 97–122.

Koch, M., Manuylov, I., and Smolka, M., 2021, “Robots and Firms” , *The Economic Journal*, 131(638), 2553–2584.

Levinsohn J, Petrin A., 2003, “Estimating Production Functions Using Inputs to Control for Unobservables” , *The Review of Economic Studies*, 70(2), 317–341.

Martin, I. W., and Nagel, S., 2022, “Market Efficiency in the Age of Big Data” , *Journal of Financial Economics*, 145(1), 154–177.

Mitra, A., Sharma, C., and Véganzonès–Varoudakis, M. A., 2016, “Infrastructure, Information & Communication Technology and Firms’ Productive Performance of the Indian Manufacturing” , *Journal of Policy Modeling*, 38(2), 353–371.

Norkutė, M., Sarafidis, V., Yamagata, T., and Cui, G., 2021, “Instrumental Variable Estimation of Dynamic Linear Panel Data Models with Defactored Regressors and a Multifactor Error Structure” , *Journal of Econometrics*, 220(2), 416–446.

Olley S, and Pakes A. 1992, “The Dynamics of Productivity in the Telecommunications Equipment Industry” , *NBER Working Paper*, (3977).

Popescu, G. H., Petrescu, I. E., and Sabie, O. M., 2018, “Algorithmic Labor in the Platform Economy: Digital Infrastructures, Job Quality, and Workplace Surveillance” , *Economics, Management and Financial Markets*, 13(3), 74–79.

Sagiroglu, S., and Sinanc, D., 2013, “Big data: A Review” , *2013 International Conference on Collaboration Technologies and Systems (CTS)*, May

Tientao, A., Legros, D., and Pichery, M. C., 2016, “Technology Spillover and TFP Growth: A Spatial Durbin Model” , *International Economics*, 145, 21–31.

Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J. F., Dubey, R., and Childe, S. J., 2017, “Big Data Analytics and Firm Performance: Effects of Dynamic Capabilities” , *Journal of Business Research*, 70, 356–365.

Zhang, X., 2021, “Broadband and Economic Growth in China: An Empirical Study During the COVID–19 Pandemic

Period” , *Telematics and Informatics*, 58,101533.

Zou, W., and Xiong, Y., 2023, “Does Artificial Intelligence Promote Industrial Upgrading? Evidence from China” , *Economic Research—Ekonomiska Istraživanja*, 36(1), 1666–1687.

Big Data Development, Artificial Intelligence Applications, and Total Factor Productivity: Evidence from the National Big Data Comprehensive Experimental Zone

ZHANG Longpeng TANG Fang

Abstract : Treating the establishment of the National Big Data Comprehensive Pilot Zone as an exogenous shock in terms of big data development, this paper uses difference-in-differences method to investigate the impact of big data development on firms’ total factor productivity as well as the mechanisms played by artificial intelligence applications based on the data of Chinese A-share listed companies from 2007 to 2019. The empirical results show that the development of big data can significantly improve the total factor productivity of enterprises, and promoting the application of artificial intelligence is an important influence mechanism. Among which, the total factor productivity promotion effect of big data development is more significant for enterprises in the five major urban agglomerations, small and medium-sized enterprises, and manufacturing enterprises. Besides, the National Big Data Comprehensive Pilot Zone has the same policy linkage effect as the construction of broadband China demonstration city and national information consumption city. Furthermore, there is a certain degree of administrative isolation in the development of big data, and the flow of data elements across provinces is blocked, which in turn weakens the improvement effect of total factor productivity of enterprises. The paper deepens the research on effect of the development of big data on firms’ total factor productivity, which provide enlightenment and reference for optimizing the development of big data, helping enterprises’ digital growth and improving production efficiency of China.

Keywords : Big Data; Artificial Intelligence; Total Factor Productivity; The National Big Data Comprehensive Pilot Zone

【责任编辑：余德淦】