

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

邱子迅, 周亚虹

(上海财经大学 经济学院, 上海 200433)

摘要: 数字经济发展的根本动力是作为通用技术的互联网和大数据技术的进步, 这种技术进步提高了信息生产和传播的效率, 推动了高质量发展。文章基于国家级大数据综合试验区, 利用 2013—2017 年我国地级市面板数据, 采用广义 DID 模型, 研究了数字经济发展对地区全要素生产率的影响。研究发现, 大数据试验区的建立显著提高了地区全要素生产率, 这种促进作用是由纯技术进步主导的。同时, 试验区的建立有助于增加地区和数字产业有关的创新, 提升地区制造业企业的智能化意识和研发水平。异质性分析表明, 试验区对生产率的促进作用在经济欠发达、创新水平低和产业结构高度高的地区更强。此外, 试验区的建立对劳动力净流出地区的全要素生产率的促进作用更大, 这为劳动力流失地区的发展提供了一定的参考。

关键词: 大数据; 全要素生产率; 国家级大数据综合试验区

中图分类号: F061.5 **文献标识码:** A **文章编号:** 1001-9952(2021)07-0004-14

DOI: 10.16538/j.cnki.jfe.20210119.103

一、引言

我国经济由高速增长阶段转向高质量发展阶段, 亟待转变发展方式, 优化经济结构, 转换增长动能。随着我国步入老龄化社会, 传统要素对经济增长的贡献逐渐减小。人口老龄化会引起劳动力有效供给不足, 而且随着城镇化的逐步推进, 土地、能源和原材料等各种资源也愈发紧张。因此, 传统的生产体系已经无法再满足经济持续发展的要求。根据测算, 如果仅仅依靠资本和劳动力的投入, 我国已无法保持过往的经济增长速度(马克·珀迪等, 2017)。

而数字经济迅速发展, 其特征是作为通用技术的大数据技术的进步, 大幅提高了信息生产与传播效率, 推动我国进入新工业革命时代。此外, 作为大数据的补充技术, 云技术和区块链技术使海量数据的计算存储、公开透明和集体维护成为可能。首先, 单位或个体可以利用大数据技术获得并分享有价值的信息, 在信息传递、增值和倍增的过程中提升人力资本质量, 激发创业激情。其次, 大数据技术使市场上各类信息更加透明, 企业能够利用大数据降低交易成本, 实现供需精准对接, 并在这个过程中加速内部革新和推进自动化。个体的创业和企业的自我革新将促进产业结构调整 and 升级, 催生一批新兴的先导产业, 如智慧城市、智慧安防、远程医疗、远程教育、无人驾驶、车联网等。最后, 政府使用相关技术改进政务服务, 将降低管理成本, 促进信息公

收稿日期: 2020-07-16

基金项目: 国家自然科学基金重点项目(71833004); 上海财经大学创新团队资助项目

作者简介: 邱子迅(1996—)(通讯作者), 男, 山东济南人, 上海财经大学经济学院博士研究生;

周亚虹(1965—), 男, 上海人, 上海财经大学经济学院教授, 博士生导师。

开,提升治理效率。因此,通过数字经济推动大众创新创业,加速产业结构调整,提升政府治理水平,增进人民福利,进而实现经济高质量发展,正成为全球发展的主要趋势。

全球主要工业化国家在过去几年相继推出大数据发展政策,推动数字经济发展。2011年,美国制定《大数据研究与发展计划》,标志着美国将大数据上升为国家战略。2013年,澳大利亚政府出台《公共服务大数据战略》,旨在利用大数据推动公共部门改革(张勇进和王璟璇,2014)。我国政府同样重视和鼓励大数据产业发展。2015年8月31日,国务院印发《促进大数据发展行动纲要》(以下简称《纲要》)。《纲要》指出,我国在大数据发展和应用方面已具备一定基础,拥有巨大的市场优势和发展潜力,但也存在数据开放共享不足、产业基础薄弱等问题亟待解决。为了加快大数据部署和深化大数据应用,通过不断的实践总结出一套可以在全国推广复制的大数据产业发展经验,实现数字发展、绿色发展、创新发展,《纲要》明确提出“开展区域试点,推进贵州等大数据综合试验区建设”。同年9月,贵州省启动首个国家级大数据综合试验区(以下简称试验区或大数据试验区)的建设。次年,第二批获批建设国家级大数据综合试验区的省份名单发布,包括北京、天津、河北、内蒙古、辽宁、河南、上海、重庆和广东。大数据试验区主要通过承担七大任务推进地区大数据产业发展,包括数据资源管理与共享开放、数据中心整合、数据资源应用、数据要素流通、大数据产业集聚、大数据国际合作以及大数据制度创新。

数字经济的快速发展引起了国内外学者的广泛关注,国际上对数字经济的研究经历了从信息经济和互联网经济到数字经济的探索过程(许宪春和张美慧,2020)。信息化和互联网对经济增长影响的研究较多,得到的结论也比较统一。通讯设备(Röller和Waverman,2001)、电信基础设施(郑世林等,2014)以及互联网(郭家堂和骆品亮,2016;韩先锋等,2019)等信息基建的普及能够显著促进地区经济增长。基于上市公司的研究认为,企业的计算机化(Brynjolfsson和Hitt,2000)和互联网化(杨德明和刘泳文,2018)有助于降低决策成本,提高生产效率。

随着大数据技术和人工智能的发展,学者关注的焦点逐渐由互联网等信息基础设施转向人工智能。很多研究发现,与作为基础设施的互联网和信息通讯技术设备不同,人工智能更多的是作为一种新型生产要素,其本身属于高科技机械设备,具有资本的性质,同时又可在一定程度上替代劳动力,具有劳动力的性质。大量研究通过加入机器智能和人类劳动互补的假设来改进新古典模型,发现机器智能的运用可以将经济增长率提升一个数量级甚至更多(Hanson,2001),而且对中低端劳动力会产生极强的替代作用(Acemoglu和Restrepo,2016)。国内学者从劳动力需求、资本回报和全要素生产率三条途径,研究了人工智能如何应对老龄化冲击,促进经济增长(陈彦斌等,2019)。

多数文献认同数字经济发展提升了经济生活效率。但就“数字鸿沟”这一问题,学界尚未形成共识。一方面,信息技术的使用需要一定的受教育程度,相对落后地区的居民使用信息技术的能力较差(Bonfadelli,2002),且现实中积累的资本会通过互联网接入转化为互联网资本,在网络效应下地区间和个体间的发展不平等会愈发严重(邱泽奇等,2016;王修华和赵亚雄,2020)。另一方面,信息通讯技术和金融科技可以有效降低个体获取市场信息的成本,缓解落后地区的信息不对称问题(Jensen,2007;许竹青等,2013)以及低收入群体的信息约束和信贷约束,使创业机会均等化,促进其生产经营(谢绚丽等,2018;张勋等,2019)。

通过文献梳理可以发现,相关研究依然有改进提升的空间。首先,虽然有文献通过微观理论模型和宏观数值模拟,考察了作为要素投入的大数据和人工智能对经济增长的影响,但是尚缺乏相关问题的实证研究,特别是与地区经济直接相关的中观层面的研究。其次,现有文献尚未对数字经济能否促进均衡发展这一问题达成一致,关于大数据和人工智能在促进均衡发展中的作

用讨论较少。最后,在我国有为政府和有效市场良性互动的环境下,政府出台了相关的区域产业政策,政策效果需要分析。本文试图回答以下问题:快速发展的数字经济能否承担起经济发展动能转换的重任?如果可以,其主导机制是什么?又能否助力后发地区实现跨越式发展?

本文使用2013—2017年我国地级市平衡面板数据,基于国家级大数据综合试验区这一准自然实验,评估了数字经济发展对地区全要素生产率的作用。研究发现,试验区的建立通过推动区域的纯技术进步促进了生产率提升,同时显著增加了区域的数字产业创新以及区域内制造业企业的研发投入和智能化水平。这种促进作用在经济欠发达、创新水平低和产业结构高度高的地区更强。此外,试验区对劳动力净流出地区的帮助更大,有助于实现跨越式发展。

本文的结构安排如下:第二部分介绍分析框架与作用机制,第三部分说明使用的计量模型和变量,第四部分介绍实证结果,最后进行总结并提出政策建议。

二、分析框架与作用机制

全要素生产率(TFP)一直是经济增长相关文献的重点研究对象,其定义是资本、劳动力等要素投入不变时经济增长的速度(Solow, 1957)。与GDP等经济增长数量指标相比,全要素生产率能够更好地反映经济增长质量,如人力资本质量、企业技术和管理模式以及地方产业结构等,因而是国家政策评估的重要标准。本文同样选取全要素生产率作为研究对象。Farrell(1957)将全要素生产率的变化分解为纯技术进步的变化和技术使用效率的变化。全要素生产率的提升可能是其中一个指标变化所主导的,也可能是两个指标共同变化所导致的。大数据试验区的建立是为了推进大数据与实体经济的深度融合,下文将通过试验区承担的七大任务来分析这两个指标的变化方向,进而确定大数据试验区建立影响全要素生产率的主导因素。

(一)纯技术进步的变化

纯技术进步是指不同时期生产前沿面的移动,即相同投入在不同时期的最优产出的变化(Farrell, 1957)。由于纯技术进步侧重于度量最优产出,新技术的出现和新业态的产生都有利于推动社会的纯技术进步。

大数据综合试验区的重要任务是实现数据的共享开放,促进数据资源的运用。试验区的经济个体获取信息的能力增强,在信息的传递、增值和倍增过程中,可以通过大数据技术得到对自己有价值的信息和知识,并分享自己的信息及对信息的理解。信息流动使区域人力资本得到积累,为地区创业创新奠定了人才基础。同时,数据可得性的提升为大众创新营造了良好的条件,创新的氛围将带动大数据的增值性公益性开发和创新应用。为了获取潜在的巨大利润并在激烈的竞争中崭露头角,初创企业往往会冒险尝试多元化的技术路线和商业模式(谢伏瞻, 2019),从而加快智慧城市、智慧安防、远程医疗、远程教育、无人驾驶、车联网等新业态的产生,数据红利在激烈的竞争中得到充分释放。

大数据时代深度学习技术的广泛使用还会加快企业自动化改革的步伐。随着人工智能的发展,自动化水平上升且成本下降。对于同一任务,培训劳动力的成本相对于机器较高。企业更倾向于将这项任务自动化,此前无法被机器替代的劳动任务将越来越多地实现自动化处理,单位劳动生产率大幅提升。企业还可以利用大数据技术降低管理决策成本,发现营销机会,定位了解客户,提供个性服务,使生产转向需求拉动,防止出现生产过剩,做出更有针对性的生产和销售决策。

最终,试验区内的大众创新和企业信息化改革将加速新业态的产生和发展,原先的金融、物流、医疗、教育等服务以一种技术水平更高、管理模式更先进、产品质量更优秀、产业链定位更精

准、社会效益更高的形态出现,促进地区的产业结构升级。例如,之前“一铝独大”的贵州白云区在大数据试验区的引领下,高端装备制造业、食品产业和医药产业迅速壮大。截至2017年底,白云区规模以上工业企业达123家,铝企业GDP占全区的比重从2011年的44%下降到11.5%,电子信息制造业产值从2011年的14.06亿元上升到50.6亿元,高新技术企业实现增加值21.3亿元,占规模以上工业增加值比重的23.3%。^①同时,新技术的引进和新业态的产生也需要信息基建的配套升级。贵州省在建立大数据试验区后,加快推进全省的互联网建设,实现了省内全域互联网覆盖,全省信息基础设施发展水平从全国第29位上升到第15位。乌兰察布对1029个行政村新增光纤覆盖,并对278个行政村的光纤进行改造,改造完成后实现全村光纤通达。互联网基础设施的配套改进和升级再次通过提升纯技术进步,提高了全要素生产率(郭家堂和骆品亮,2016)。

作为一个国家级的标杆平台,大数据试验区有力地促进了当地的大数据产业集聚。贵州在被选定为试验区后,面向优强企业和高端前沿研发机构开展引资引智活动,积极推进大数据重点企业、科研院所和行业协会开展技术交流与合作。河南也基于试验区积极招商引资,截至2018年已有102家大数据企业落地河南,其中不乏华为、浪潮、超图等国内优秀高科技企业。^②在产业集聚的基础上,数据要素的流通进一步通过政产学研的直接对接,减少了金融机构投资过程中的信息不对称,缓解了研发投资不足的问题,促进了地区创新创业。仅在2016—2017年,所有大数据试验区获得融资的大数据创投项目近千个,新申请大数据类技术发明专利超过8000项。^③

产业集聚还会吸引高技术外资企业入驻,外资入驻带来了更前沿的技术,同时加剧了区域内的竞争,促进国内企业加快研发的步伐。贵州大数据企业从2013年的不足1000家增长至2018年的8900多家,苹果、高通、微软、戴尔、惠普、英特尔、甲骨文等世界知名企业纷纷扎根贵州,大数据产业规模总量超过1100亿元。2014—2017年,贵州省规模以上电子信息制造业增加值、软件业务收入和网络零售交易额年均分别增长78.9%、35.9%和38.2%。^④

根据上述分析,本文认为建立大数据试验区对区域的纯技术进步具有促进作用。

(二) 技术使用效率的变化

给定要素投入,一个地区或企业的实际产出往往受到生产组织形式、管理水平和激励机制等因素的影响而无法达到最大值,技术使用效率度量了最大产出和实际产出的差距(Farrell, 1957),两者差距越小,技术使用效率越高,反之越低。下文将分析试验区建立所带来的技术使用效率变化。

大数据试验区虽然鼓励企业充分利用数据资源进行生产和决策,但是在网络效应的作用下,不可避免地会加剧互联网时代的垄断,不利于发挥技术使用效率。在互联网时代,一些已形成垄断的企业凭借积累的海量数据,能够比其他企业做出更加有效和精准的决策,算法也会通过深度学习不断迭代和升级,为用户提供更加个性化和精准化的服务,形成消费者集聚,进而积累更多的数据资源,巩固其垄断地位,加剧行业垄断。同时,部分传统行业运用大数据的意识不足,如果要求传统企业在短期内掌握大数据技术,并实现这些技术与实际生产的融合,企业需要付出极高的时间成本和训练成本,以调整内部组织结构和培训工人(Brynjolfsson和Hitt, 2000)。因此,这些企业更愿意暂时保持原有的生产模式,导致新技术的使用效率难以提升。进一步地,

① 信息来源:《贵阳日报》。

② 信息来源:《河南日报》。

③ 信息来源:《国家大数据综合试验区发展报告(2018)》。

④ 信息来源:《经济日报》。

大数据互联网企业掌握更多的数据,也能熟练运用大数据技术了解消费者偏好和市场趋势,很容易对传统行业形成冲击,侵占未转型的传统行业企业的市场份额。这种新技术所带来的“破坏性创新”同样不利于发挥技术使用效率。

试验区对当地政府的要求有效地弥补了企业等市场主体无法通过市场竞争实现资源有效配置的问题,政府对数据中心的整合以及对大数据制度的完善能够为试验区提供良好的生产环境和管理环境,促进技术使用效率的提升。一方面,数据中心的整合有助于提高政府的公共服务质量和治理效率。河南和贵州的大数据云平台通过移动设备端口实现对个人数据的采集,有力促进了对新冠疫情的监控、防止、溯源以及对物资的集中调度,极大提升了政府应对重大事件的能力。另一方面,贵州、广东、内蒙古等多地在被确立为试验区后,立即对管理体制进行完善,从数据共享规则、数据安全标准、数据隐私保护和数据资源权益维护等方面出台条例,以保障大数据技术在试验区的安全、有序、有效应用。

根据上述分析,建立大数据试验区对区域的技术使用效率既有正向作用,也有负向作用,最终的方向无法确定,需要实证分析给出答案。

(三)全要素生产率的变化

谢伏瞻(2019)指出,我国当前正处于新工业革命由导入期转入拓展期的阶段。本文认为,建立大数据试验区对全要素生产率的影响很可能符合技术导入期的特征(纯技术进步促进生产率提升),也有部分技术拓展期的特征(技术使用效率促进生产率提升);但整体而言,纯技术进步发挥主导作用,提升了区域的全要素生产率。

本文的分析框架和作用机制见表1。

表1 分析框架与作用机制

试验区七大任务	传导途径		作用机制
数据资源管理与共享开放 数据资源应用	个人:信息流通中促进人力资本积累、数据可得性激发创业激情→促进地区创新创业	基础设施的不断完善、新业态的产生与发展→产业结构升级	提升纯技术进步率
	企业:自动化程度大幅提升,自动化信息化改革加速		
	企业:网络效应促进大数据互联网企业形成垄断、部分传统行业运用大数据意识不足→对传统行业形成冲击		抑制技术使用效率
数据要素流通 大数据产业集聚	企业:向大数据试验区汇聚→研发机构、投资集团、制造企业实现对接→缓解研发投入不足的问题		提升纯技术进步率
大数据国际合作	外资:向大数据试验区汇聚→引入新技术、加剧区域内竞争		提升纯技术进步率
数据中心整合	政府:提高公共服务质量和治理效率		提升技术使用效率
大数据制度创新	政府:完善管理体制、填补制度空白		提升技术使用效率

三、模型与变量

本文的目的是探究以大数据产业为核心的数字经济发展能否促进区域全要素生产率提升,试验区则为本文的研究提供了一个良好的准自然实验环境。本文根据地级市所在省是否被选定为试验区,将样本划分为实验组和控制组,检验大数据人工智能对生产率提升的因果效应。

(一)模型设定

关于建立试验区对全要素生产率的影响,评估的难点在于识别因果效应:国家对于大数据试验区的选址很可能不是外生的,会考虑影响当地全要素生产率的禀赋等因素,而这些与区域特征相关的因素不可能被完全观测或控制,因而直接使用OLS估计方法必然会影响政策评价结果的一致性。本文使用广义双重差分法(Generalized DID)进行识别,这种双重差分法将一般DID

模型中处理组—时间的交互项更换为随个体—时间可变的政策虚拟变量,可以用于识别渐进推进政策的效应(Beck等,2010)。实证模型设定如下:

$$Y_{ipt} = \beta_0 + \beta_1 policy_{ipt} + \beta_2 X_{ipt} + \beta_3 W_{pt} + \alpha_i + \delta_t + \varepsilon_{ipt}$$

其中, i 表示地级市, p 表示地级市所在省份, t 表示年份。被解释变量 Y 表示地级市的全要素生产率 $tfpch$ 。 $policy$ 表示建立大数据试验区的政策变量,如果地级市 i 所在省份 p 在第 t 年为国家级大数据试验区,则 $policy$ 取值为1,否则为0。 $policy$ 的系数 β_1 是本文关注的参数。 X 为市级层面的控制变量, W 为省级层面的控制变量。时间固定效应 δ_t 的引入剔除了时间趋势的影响, α_i 为城市固定效应, ε_{ipt} 为随机扰动项。本文在估计时采用地级市层面的聚类标准误。

(二)变量定义

本文选取2013—2017年207个地级市的平衡面板数据,数据主要来自《中国城市统计年鉴》《中国人口和就业统计年鉴》《中国信息年鉴》《中国信息产业年鉴》《中国互联网络发展状况统计报告》、各省统计年鉴、各地级市统计年鉴、WIND数据库以及EPS数据库。

1. 被解释变量

鲁晓东和连玉君(2012)指出, $Malmquist$ 指数法是一种比较合适的计算宏观或中观单位(国家、省份或地级市)全要素生产率,并将其分解为纯技术进步和技术使用效率的方法。本文使用此方法计算全要素生产率 $tfpch$,其中投入和产出变量的处理如下:

产出变量是地级市实际GDP。本文将基期设定为2006年,并参考余泳泽等(2019)的方式,以各省的GDP指数对地级市名义GDP进行平减,得到实际GDP。

投入变量包括地级市的固定资本存量和城镇就业总人口。地级市的固定资本存量通过永续盘存法进行估计: $K_{ipt} = (1 - \delta)K_{ipt-1} + I_{ipt}/PI_{pt}$,其中 K 为地级市的固定资本存量; δ 为折旧率,参考张军等(2004)的研究,取值为9.6%; I 为固定资产投资, PI 为以2006年为基期平减后的省级固定资产投资指数。本文以2006年为基期,地级市基期资本存量可以根据现有文献的数据进行估计,具体方法如下: $K_{ip2006} = K_{p2006}I_{ip2006}/I_{p2006}$,其中2006年省级固定资本存量参考单豪杰(2008)的计算结果。

得到投入变量和产出变量后,本文利用DEAP2.1^①程序得到2013—2017年各地级市的全要素生产率。

2. 政策变量

政策变量 $policy$ 是本文关注的, $policy$ 表示是否建立大数据试验区,如果地级市所在省份为国家级大数据试验区,则 $policy$ 取值为1,否则为0。

3. 控制变量

大数据对全要素生产率的影响不一定只通过大数据综合试验区发挥作用,非试验区也可以运用大数据以其他途径提高本地的自动化和智能化水平,进而影响全要素生产率。因此,本文引入省级信息化程度 $info$,控制地区的自动化和智能化发展水平。同时,为了区分大数据和互联网对全要素生产率的不同影响,本文还控制了市级互联网普及率 $inter$ 。

参考现有文献,对外开放(Coe和Helpman,1995;简泽等,2014;陈蓉和许培源,2017)、研发实力和企业规模(周亚虹等,2012;李经路,2015)、人力资本质量(彭国华,2007;樊纲等,2011;程名望等,2019)以及工业化程度(戴维·N·韦尔,2007)都会影响企业或地区的全要素生产率,因此本文还控制了这些变量。变量定义与描述性统计见表2。

^① DEAP2.1是一款运用数据包络法计算的软件,可以用于计算Malmquist指数,得到全要素生产率。

表2 变量定义与描述性统计

变量符号	变量名称	变量定义	样本量	均值	方差
<i>tfpch</i>	全要素生产率	Malmquist 指数法计算得到	1035	0.9890	0.1058
<i>indust</i>	工业化程度	地级市第二产业 GDP 占比	1035	0.4702	0.1141
<i>foreign</i>	对外开放	地级市规模以上外资企业数量占比	1035	0.0638	0.0723
<i>firm</i>	企业数量	地级市规模以上企业数量的对数	1035	5.5635	1.2264
<i>inter</i>	互联网普及	地级市接入宽带互联网的用户占本市平均人口的比重	1035	0.2060	0.1522
<i>edu</i>	受教育程度	省级平均受教育程度,根据樊纲等(2011)的加权算法 ^① 计算得到	1035	9.0535	0.5641
<i>info</i>	信息化	地级市所在省年度新增电子信息产业专利数占本省当年新增总专利数的比重	1035	0.5201	0.1345

四、实证结果分析

(一)多重共线性检验

在回归分析之前,本文首先进行多重共线性检验。解释变量的方差膨胀因子检验显示,所有变量中 *VIF* 最大的值为 1.63,远小于经验法则所要求的临界值 10,因而不必担心变量间的多重共线性问题(见表 3)。

表3 解释变量的方差膨胀因子

变量	<i>indust</i>	<i>foreign</i>	<i>firm</i>	<i>inter</i>	<i>edu</i>	<i>info</i>	<i>policy</i>
<i>VIF</i>	1.16	1.16	1.63	1.58	1.21	1.02	1.07
1/ <i>VIF</i>	0.87	0.87	0.61	0.63	0.83	0.98	0.93

(二)回归结果分析

模型回归结果见表 4,其中列(1)只控制了地区固定效应和时间固定效应,列(2)进一步控制了其他的地区重要变量。结果显示,建立大数据试验区对区域的全要素生产率具有显著的促进作用,试验区的全要素生产率要比非试验区平均高出 0.0493 个单位。互联网普及率和信息化程度对区域的全要素生产率也起到显著的促进作用,说明在样本期内,网络化和自动化的推进有助于提升生产率。此外,规模企业数量的系数显著为正,说明区域内规模企业的数量越多,内部竞争压力越大,研发实力越强,进而区域的生产率越高。工业化程度的系数不显著,说明随着我国工业化程度的提高,传统生产模式对经济增长的贡献在逐步减小。在样本期内,平均受教育程度对全要素生产率也没有显著影响,可能的原因是:我国近年来越来越强调自主研发创新,教育方面也向自主创新倾斜,而这种倾斜在短期内未必能够提升生产率,反而会挤压模仿创新(郭家堂和骆品亮,2016),因此教育水平总体上对全要素生产率没有显著作用。

表4 模型回归分析

因变量: <i>tfpch</i>	(1)全样本	(2)全样本
<i>policy</i>	0.0433*** (2.73)	0.0493*** (3.05)
<i>indust</i>		0.170 (1.16)
<i>foreign</i>		-0.0284 (-0.17)
<i>firm</i>		0.0393* (1.89)
<i>inter</i>		0.156*** (2.63)
<i>edu</i>		-0.0282 (-1.05)
<i>info</i>		0.0849** (2.50)
<i>cons</i>	0.936*** (149.37)	0.818*** (2.82)
固定效应	城市、时间	城市、时间
<i>N</i>	1 035	1 035

注:括号内为使用稳健标准误计算的 *t* 值,***、**和*分别表示 1%、5% 和 10% 的显著性水平。下表同。

① 省级平均受教育程度=(文盲人数×0+小学学历人数×6+初中学历人数×9+高中学历人数×12+大专及以上学历人数×16)/6 岁以上总人口。

(三) 稳健性检验

为了保证政策评估结果的可靠性, 本文从广义 DID 识别条件、政策外生性、遗漏变量、模型设定、对照组和样本期选择六个方面进行了稳健性检验。

1. 广义 DID 识别条件

广义 DID 能够在一定程度上解决政策评估中的内生性问题, 但这是建立在一定假设基础上的, 要求在政策实施前, 试验区与非试验区的全要素生产率保持平行的发展趋势。参考 Beck 等 (2010) 的方法, 这里引入政策虚拟变量 $policyb2$ 、 $policyb1$ 和 $policya1$ (剔除末年政策虚拟变量避免多重共线), 取值为 1 的条件分别为“两年后为试验区”“一年后为试验区”和“一年前为试验区”。如果 $policyb2$ 和 $policyb1$ 的系数同时不显著, 则表明平行趋势假设成立。表 5 中列 (1) 结果显示, $policyb2$ 和 $policyb1$ 的系数均不显著, 平行趋势假设成立。

表 5 稳健性检验

因变量: $tfpch$	(1) 平行趋势假设检验	(2) 预期效应检验	(3) 一般 DID 估计	(4) PSM-DID
did			0.0464*** (2.71)	0.0501** (2.48)
$policyb2$	-0.0132 (-0.59)			
$policyb1$	-0.0128 (-0.57)	-0.0190 (-0.95)		
$policy$	0.0711*** (3.38)	0.0613*** (3.48)		
$policya1$	-0.0200 (-0.69)			
控制变量	控制	控制	控制	控制
固定效应	城市、时间	城市、时间	城市、时间	城市、时间
N	1 035	1 035	1 035	610

2. 政策外生性

广义 DID 要求政策实施前试验区和非试验区不能形成有效预期, 即保证政策的外生性。在控制个体效应和时间效应后, 作为国家战略的大数据综合试验区在理论上是外生政策, 但如果某些地区对政策有预期而提前采取措施, 则我们对政策效果的评价将是不准确的。因此, 在估计中加入 $policyb1$ 控制预期效应。表 5 中列 (2) 结果显示, $policyb1$ 的系数不显著, 且 $policy$ 系数的显著性和符号没有发生明显变化, 说明预期效应并不存在。

3. 遗漏变量

如果存在不可观测的遗漏变量 Z 进入 ε_{ipt} , 则可能会对估计结果产生影响。根据系数估计的表达式, 我们有 $\hat{\beta}_1 = \beta_1 + \theta COV(policy_{ipt}, Z_{ipt} | X_{ipt}, W_{pt}) / VAR(policy_{ipt} | X_{ipt}, W_{pt})$, 其中 θ 为不可观测的遗漏变量对全要素生产率的影响。这里采用周茂等 (2018) 的安慰剂检验方法对原假设 $\theta = 0$ 进行检验。如果存在一个安慰剂 $placebo_{ipt}$ 对全要素生产率的实际影响为 0, 则使用该安慰剂代替 $policy_{ipt}$ 所得到的回归系数是 $\hat{\beta}_{1, placebo} = \theta COV(placebo_{ipt}, Z_{ipt} | X_{ipt}, W_{pt}) / VAR(placebo_{ipt} | X_{ipt}, W_{pt})$ 。如果 $\hat{\beta}_{1, placebo}$ 为 0, 则可以反推 $\theta = 0$ 。我们随机安排政策得到安慰剂, 并重新估计广义 DID 模型, 循环 1 000 次, 得到 $\hat{\beta}_{1, placebo}$ 的密度函数, 其分布以 0 为中心, 均值为 0.00015, 置信度为 95% 的 $\beta_{1, placebo}$ 区间估计为 (-0.0432, 0.0488), 因此可以认为 $\theta = 0$, 即不可观测的遗漏变量不影响政策评估的结果。表 4 中列 (2) $policy$ 的回归系数 0.0493 处在该区间右侧, 可以认为政策不同于模拟生成的安慰剂, 对全要素生产率具有显著的促进作用。

4. 模型设定

虽然大数据试验区在全国是渐进式推广的,但是在2015年试验区只有贵州。试验区推广的峰值是在2016年,这里选取2016年作为政策的起始年份,重新估计后得到相似的结果(见表5中列(3))。

5. 对照组选择

为了避免潜在的选择性偏差问题,这里对试验区和非试验区的地区进行了倾向得分匹配,以保证实验组和对照组能够进行更加合理的比较。在样本匹配后,本文重新估计了政策效应,结果没有明显变化,表明上文结果是稳健的(见表5中列(4))。

6. 样本期选择

这里增加2012年数据重新进行了估计,并做了稳健性检验,以保证平行趋势假设在更长时间内依然成立,且结果不受样本期选择的影响。表6中核心变量的估计结果与表4和表5没有明显差异,说明结果对样本期选择不敏感。

表6 稳健性检验:增加2012年数据

因变量: <i>tfpch</i>	(1)基准回归	(2)平行趋势假设检验	(3)预期效应检验	(4)一般 DID 估计	(5)PSM-DID
<i>did</i>				0.0455*** (2.81)	0.0435** (2.28)
<i>policyb3</i>		0.0203 (1.15)			
<i>policyb2</i>		-0.0145 (-0.68)			
<i>policyb1</i>		-0.0124 (-0.55)	-0.0156 (-0.87)		
<i>policy</i>	0.0486*** (3.19)	0.0701*** (3.36)	0.0599*** (3.43)		
<i>policya1</i>		-0.0208 (-0.72)			
控制变量	控制	控制	控制	控制	控制
固定效应	城市、时间	城市、时间	城市、时间	城市、时间	城市、时间
<i>N</i>	1 242	1 242	1 242	1 242	732

注:如果三年后为试验区,则政策虚拟变量 *policyb3* 取值为 1,否则为 0。

(四)作用机制检验

大数据技术的广泛使用有望成为我国经济发展动能转换的基础,那么这种对全要素生产率的促进作用是由纯技术进步还是由技术使用效率主导?本文同样使用 *DEAP2.1* 程序,利用 *Malmquist* 指数法,得到样本期内各地级市的纯技术进步率 *techch* 和技术使用效率 *effch*。我们将这两个变量作为被解释变量,估计广义 *DID* 模型。表7中列(1)和列(2)结果表明,建立大数据试验区对纯技术进步具有显著的促进作用,对技术使用效率虽有正向影响,但系数并不显著。因此,大数据试验区对全要素生产率的促进作用主要是通过提升纯技术进步率实现的。

为了说明研究的稳健性,我们对试验区创新水平的变化和制造业企业的反应进行了检验。首先,我们选取省级年度新增电子专利数量的对数 *lnelepat* 作为被解释变量,表7中列(3)显示 *policy* 的系数显著为正,表明试验区的建立显著增加了与电子产业相关的创新行为,有助于推动新业态的产生和发展。其次,我们选取2013—2017年国A股上市公司作为样本,^①检验试验区的

① 样本的处理方式如下:仅保留2013年之前上市的公司,剔除样本期内退市和重组的公司以及有缺失值的样本。

建立是否有助于提升当地企业的智能化转型意识和研发投入。由于年报中往往会披露企业的近况和展望,我们通过对企业年报进行文本挖掘,构造企业智能化意识变量 *airate*。^①我们分别以企业智能化 *airate* 和研发投入占营业收入比重 *rds* 作为被解释变量,并参考现有文献控制了企业层面的变量,表 7 中列(4)和列(5)显示大数据试验区的建立显著提升了区域内企业的智能化意识和研发投入力度。此外,表 7 中列(6)显示, *policy* 和 *airate* 的系数均显著为正,说明试验区的建立通过增强智能化意识,增大了企业的研发投入力度。

表 7 作用机制检验

因变量	(1) <i>techch</i>	(2) <i>effch</i>	(3) <i>lnlepat</i>	(4) <i>airate</i>	(5) <i>rds</i>	(6) <i>rds</i>
<i>policy</i>	0.0115** (2.01)	0.0190 (0.89)	0.2328** (2.43)	0.0954** (2.21)	0.359*** (2.60)	0.315** (2.32)
<i>airate</i>						0.467*** (6.49)
控制变量	控制	控制	控制	控制	控制	控制
固定效应	城市、时间	城市、时间	省份、时间	行业、时间	行业、时间	行业、时间
<i>N</i>	1 035	1 035	150	4 808	4808	4 808

注:列(1)和列(2)中的控制变量与表 4 相同,列(3)中的控制变量均是省级层面的变量。参考杨德明和刘泳文(2018)的研究,列(4)、列(5)和列(6)中的控制变量包括经营费用率、财务杠杆率、薪酬前三位的高管工资总和的对数、公司规模(总资产的对数)以及是否国有企业。

(五)异质性分析

一个值得关注的问题是,现阶段我国区域发展不平衡,大数据综合试验区作为一项国家战略,对不同地区生产率的促进作用是否存在差异?为了回答这个问题,下文将展开更加详细的讨论。

1. 跨越式发展

不同于前几次工业革命,以智能化、网络化和数字化为主导的新工业革命对地区原有的发展水平并没有很强的依赖性,具有一定互联网基础的落后地区也可以抓住新工业革命的机遇,促进大数据、人工智能与本地传统优势行业的融合,在追赶发达地区的同时注重质量和效率,实现跨越式发展。本文根据政策发生前一年各地的人均 *GDP* 是否大于当年全国人均 *GDP*,将所有样本划分为经济发达地区和欠发达地区,分样本进行了回归分析。表 8 中列(1)和列(2)显示,大数据试验区对经济较发达地区全要素生产率的促进作用不显著,而对欠发达地区的促进作用更大且显著,组间差异检验显示两组的政策效果存在显著差别。类似地,本文根据政策实施前各省专利数的中位数,将所有样本划分高创新地区和低创新地区。表 8 中列(3)和列(4)显示,试验区对全要素生产率的促进作用在创新基础薄弱的地区更加显著。上述结果说明增长动能相对不足的地区可以借助大数据产业实现跨越式发展。

2. 产业结构高度

大数据的应用需要一定的产业基础,如果一个地区的主导产业是第二三产业,则该地区在生产生活中所积累的数据量较大,应用数据的意识较强。同时,该地区对大数据的需求也较大,

^① 文本挖掘步骤如下:首先,使用中文 *wiki* 训练 *word2vec* 模型,获得与“大数据”“人工智能”最相关的 50 个词,将两组词去重后得到人工智能词典。这样做减少了手工收集词汇的工作量,也大幅度提高了词典构造的客观性(沈艳等, 2019);然后,参考王克敏等(2018)的方法对公司年报进行清洗;最后,获取各公司每年年报中披露的人工智能相关词汇词频和总词汇数,构造公司人工智能意识变量 *airate*=人工智能相关词汇词频/总词汇数。

需要大数据产业推动地区发展。本文认为,产业结构高度高的地区会更加受益于大数据的发展。参考现有文献,本文计算了各地的产业结构高度指数,^①并根据政策发生前各地该指数的中位数,划分样本进行了回归分析。表8中列(5)和列(6)显示,试验区的建立对高产业结构地区的全要素生产率具有显著的促进作用,而对低产业结构地区没有显著影响,且组间差异检验显示两组的政策效果存在显著差别。

表8 异质性分析

因变量: <i>tfpch</i>	(1)经济较发达地区	(2)经济欠发达地区	(3)高创新地区	(4)低创新地区	(5)高产业结构地区	(6)低产业结构地区	(7)劳动力净增加地区	(8)劳动力净流失地区
<i>policy</i>	0.0183 (0.78)	0.0736*** (3.24)	0.0191 (1.08)	0.0842*** (3.00)	0.0761*** (4.01)	-0.00631 (-0.20)	0.0164 (1.00)	0.0794*** (2.69)
控制变量	控制	控制	控制	控制	控制	控制	控制	控制
固定效应	城市、时间	城市、时间	城市、时间	城市、时间	城市、时间	城市、时间	城市、时间	城市、时间
<i>N</i>	620	415	575	460	555	480	545	490
组间差异检验	2.94 <i>p</i> =0.0864		3.96 <i>p</i> =0.0467		5.17 <i>p</i> =0.0229		3.85 <i>p</i> =0.0585	

3. 劳动力供给状况

不同于延迟退休年龄、放开二胎和户籍改革等从供给端增加劳动供给的现有政策,大数据和人工智能的应用可以减少对劳动力的需求(Acemoglu 和 Restrepo, 2018)。本文根据政策发生前三年各地的劳动力净增加是否大于0,划分样本进行了回归分析。表8中列(7)和列(8)显示,试验区的建立对劳动力净流失地区生产率的促进作用更大且显著。这说明劳动力供给不足的地区可以通过发展大数据和人工智能产业,缓解人口流出所造成的发展动能不足的问题。

4. 边际分析

广义 DID 回归中 *policy* 的系数表示政策实施对全要素生产率的平均影响,但有时我们会对某些控制变量处在某一特定水平下的政策效果感兴趣。现有文献普遍认为,信息通讯技术设备和互联网是大数据和人工智能产生和发展的基础(许宪春和张美慧, 2020),因此互联网和电子设备可以视为大数据时代的“基础设施”。对于“基础设施”完善程度不同的地区,试验区提升全要素生产率的作用可能是不同的。我们基于互联网普及和信息化两个变量进行了边际分析。图1左图是在不同互联网普及率下的边际分析结果,可以发现大数据试验区对互联网普及程度较低地区的全要素生产率具有更加显著的促进作用。一种可能的解释是引进新技术需要基建的配套升级,作用机制分析中已做论述,在此不再赘述。建立试验区对信息化程度较高地区的生产率具有更加显著的促进作用(见图1右图),这是因为:自动化设备和信息化配件往往是采集大数据的重要端口,当数据成为一种生产资源时,自动化和信息化程度高的地区具有更强的“资源采掘和分解”能力,能够进行更大规模的数据决策和自动化生产,进而提升生产率。

五、结论与政策建议

本文利用2013—2017年我国地级市面板数据,基于大数据综合试验区,分析了大数据产业发展对全要素生产率的作用,主要得到以下结论:第一,在经济发展新时期,大数据综合试验区的建立对全要素生产率具有显著的促进作用,这种促进作用主要是由纯技术进步主导的。试验区的建立极大促进了与电子产业相关的区域创新行为,并增强了企业的自动化意识,促使企业

^① 参考徐德云(2008)的研究,计算方法如下:第一产业 GDP 占比×1+第二产业 GDP 占比×2+第三产业 GDP 占比×3。

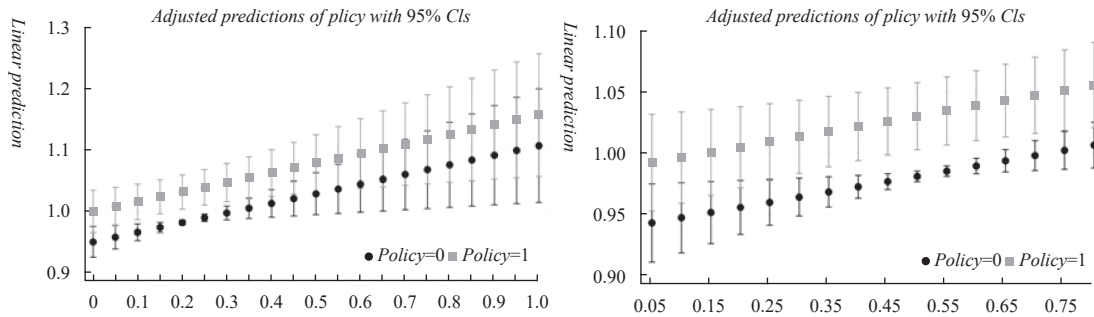


图 1 边际分析

加大研发力度。第二,试验区的建立对欠发达地区和创新不足地区的促进作用更加显著,因此相对落后的地区完全可以借助数字产业实现跨越式发展。第三,大数据试验区的建立对劳动力流失严重地区的全要素生产率具有更强的促进作用,可以缓解当地人口流出所造成的劳动力有效供给不足的问题。

基于上述结论,本文认为国家应以大数据试验区的实践作为基础,继续大力发展大数据和人工智能技术,抓住新工业革命的机遇,缩小与发达国家的差距。首先,政府应向公众适度开放数据资源,鼓励大众通过数字产业进行创业创新,支持制造业企业向自动化和智能化转型升级,并给予相关企业适当的研发补贴。其次,在各地区发展尚不平衡的背景下,政府要重点关注各地区大数据产业发展所带来的异质性作用。对于老龄化或劳动力流失而导致经济增长动能不足的地区,可以通过发展大数据和人工智能产业来缓解发展压力,实现跨越式发展。最后,政府要继续推进互联网普及,完善信息技术基础设施,发展新基建,为地方企业实现数字转型、智能升级和融合创新提供完善的基础设施体系。

参考文献:

[1]陈蓉,许培源. 进出口产品多样化的生产率效应:来自中国制造业的经验研究[J]. 经济与管理评论,2017,(2): 139-145.

[2]陈彦斌,林晨,陈小亮. 人工智能、老龄化与经济增长[J]. 经济研究,2019,(7): 47-63.

[3]程名望,贾晓佳,仇焕广. 中国经济增长(1978-2015):灵感还是汗水?[J]. 经济研究,2019,(7): 30-46.

[4]戴维·N·韦尔. 经济增长[M]. 金志农,古和今,译. 北京:中国人民大学出版社,2007.

[5]樊纲,王小鲁,马光荣. 中国市场化进程对经济增长的贡献[J]. 经济研究,2011,(9): 4-16.

[6]郭家堂,骆品亮. 互联网对中国全要素生产率有促进作用吗?[J]. 管理世界,2016,(10): 34-49.

[7]韩先锋,宋文飞,李勃昕. 互联网能成为中国区域创新效率提升的新动能吗[J]. 中国工业经济,2019,(7): 119-136.

[8]简泽,张涛,伏玉林. 进口自由化、竞争与本土企业的全要素生产率——基于中国加入WTO的一个自然实验[J]. 经济研究,2014,(8): 120-132.

[9]李经路. 公司规模对研发投入影响的研究——2010-2014年创业板的证据[J]. 兰州财经大学学报,2015,(6): 89-97.

[10]鲁晓东,连玉君. 中国工业企业全要素生产率估计:1999-2007[J]. 经济学(季刊),2012,(2): 541-558.

[11]马克·珀迪,邱静,陈笑冰. 埃森哲:人工智能助力中国经济增长[J]. 机器人产业,2017,(4): 80-91.

[12]彭国华. 我国地区全要素生产率与人力资本构成[J]. 中国工业经济,2007,(2): 52-59.

[13]邱泽奇,张树沁,刘世定,等. 从数字鸿沟到红利差异——互联网资本的视角[J]. 中国社会科学,2016,(10): 93-115.

- [14]单豪杰. 中国资本存量 K 的再估算: 1952~2006 年[J]. 数量经济技术经济研究, 2008, (10): 17-31.
- [15]沈艳, 陈赟, 黄卓. 文本大数据分析在经济学和金融学中的应用: 一个文献综述[J]. 经济学(季刊), 2019, (4): 1153-1186.
- [16]王克敏, 王华杰, 李栋栋, 等. 年报文本信息复杂性与管理者自利——来自中国上市公司的证据[J]. 管理世界, 2018, (12): 120-132.
- [17]王修华, 赵亚雄. 数字金融发展是否存在马太效应? ——贫困户与非贫困户的经验比较[J]. 金融研究, 2020, (7): 114-133.
- [18]谢伏瞻. 论新工业革命加速拓展与全球治理变革方向[J]. 经济研究, 2019, (7): 4-13.
- [19]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, (4): 1557-1580.
- [20]徐德云. 产业结构升级形态决定、测度的一个理论解释及验证[J]. 财政研究, 2008, (1): 46-49.
- [21]许宪春, 张美慧. 中国数字经济规模测算研究——基于国际比较的视角[J]. 中国工业经济, 2020, (5): 23-41.
- [22]许竹青, 郑风田, 陈洁. “数字鸿沟”还是“信息红利”? 信息的有效供给与农民的销售价格——一个微观角度的实证研究[J]. 经济学(季刊), 2013, (4): 1513-1536.
- [23]杨德明, 刘泳文. “互联网+”为什么加出了业绩[J]. 中国工业经济, 2018, (5): 80-98.
- [24]余泳泽, 刘大勇, 龚宇. 过犹不及事缓则圆: 地方经济增长目标约束与全要素生产率[J]. 管理世界, 2019, (7): 26-42.
- [25]张军, 吴桂英, 张吉鹏. 中国省际物质资本存量估算: 1952-2000[J]. 经济研究, 2004, (10): 35-44.
- [26]张勋, 万广华, 张佳佳, 等. 数字经济、普惠金融与包容性增长[J]. 经济研究, 2019, (8): 71-86.
- [27]张勇进, 王璟璇. 主要发达国家大数据政策比较研究[J]. 中国行政管理, 2014, (12): 113-117.
- [28]郑世林, 周黎安, 何维达. 电信基础设施与中国经济增长[J]. 经济研究, 2014, (5): 77-90.
- [29]周茂, 陆毅, 杜艳, 等. 开发区设立与地区制造业升级[J]. 中国工业经济, 2018, (3): 62-79.
- [30]周亚虹, 贺小丹, 沈瑶. 中国工业企业自主创新的影响因素和产出绩效研究[J]. 经济研究, 2012, (5): 107-119.
- [31]Acemoglu D, Restrepo P. The race between machine and man: Implications of technology for growth, factor shares and employment[R]. NBER Working Paper No.22252, 2016.
- [32]Acemoglu D, Restrepo P. Artificial intelligence, automation and work[R]. NBER Working Paper No.24196, 2018.
- [33]Beck T, Levine R, Levkov A. Big bad banks? The winners and losers from bank deregulation in the united states[J]. *Journal of Finance*, 2010, 65(5): 1637-1667.
- [34]Bonfadelli H. The internet and knowledge gaps: A theoretical and empirical investigation[J]. *European Journal of Communication*, 2002, 17(1): 65-84.
- [35]Brynjolfsson E, Hitt L M. Beyond computation: Information technology, organizational transformation and business performance[J]. *Journal of Economic Perspectives*, 2000, 14(4): 23-48.
- [36]Coe D T, Helpman E. International R&D spillovers[J]. *European Economic Review*, 1995, 39(5): 859-887.
- [37]Farrell M J. The measurement of productive efficiency[J]. *Journal of the Royal Statistical Society: Series A (General)*, 1957, 120(3): 253-281.
- [38]Hanson R. Economic growth given machine intelligence[EB/OL]. Berkeley: University of California, <http://mason.gmu.edu/~rhanson/aigrow.pdf>, 2001.
- [39]Jensen R. The digital divide: Information (technology), market performance, and welfare in the South Indian fisheries sector[J]. *The Quarterly Journal of Economics*, 2007, 122(3): 879-924.
- [40]Röller L H, Waverman L. Telecommunications infrastructure and economic development: A simultaneous approach[J]. *American Economic Review*, 2001, 91(4): 909-923.
- [41]Solow R M. Technical change and the aggregate production function[J]. *The Review of Economics and Statistics*, 1957, 39(3): 312-320.

Development of Digital Economy and Regional Total Factor Productivity: An Analysis Based on National Big Data Comprehensive Pilot Zone

Qiu Zixun, Zhou Yahong

(School of Economics, Shanghai University of Finance and Economics, Shanghai 200433, China)

Summary: Major industrialized countries, including China, are gradually aging, and resource constraints are increasingly tense. The traditional production system can no longer meet the requirements of sustainable economic development. However, with the rapid development of digital economy, it is becoming the main trend to promote the high-quality development of economy through the development of digital economy. At the same time, people are generally concerned about not only whether big data and AI can increase productivity but also whether Digital Gap will be deepened.

The Chinese government also attaches importance to the development of digital economy and started the construction of National Big Data Comprehensive Pilot Zone in 2015. Through undertaking seven tasks, the Pilot Zone has promoted the development of regional digital economy. Based on this rare quasi-natural experiment, this paper evaluates the effect of digital economy development on regional total factor productivity. We use the balanced panel data of China's prefecture-level cities from 2013 to 2017 to estimate the Generalized DID model, and find that the Pilot Zone significantly improves the regional total factor productivity, which is dominated by pure technological progress. At the same time, the Pilot Zone helps to increase the innovation related to regional and digital industries, and enhance the intellectualization and R&D level of manufacturing enterprises in the zone.

In addition to the average effect of the Pilot Zone on productivity, we also test whether the development of digital economy can help relatively backward areas achieve leapfrog development. We find that the Pilot Zone has a stronger promotion effect on total factor productivity in less developed areas, less innovative areas and areas with higher industrial structure. In addition, the Pilot Zone has a greater role in promoting the total factor productivity of areas with net labor loss. Therefore, we suggest that regions with insufficient economic growth momentum, especially those with aging and serious labor outflow, can develop big data and artificial intelligence industry to alleviate the pressure of development.

The possible marginal contribution of this paper lies in the following aspects: First, in the environment of benign interaction between government and effective market, the government has issued regional industrial policies to promote the development of big data industry, but no literature has investigated whether these policies will promote regional productivity. Second, there is a lack of empirical research on whether big data industry can promote productivity at the meso level which is directly related to the local economy. From the perspective of National Big Data Comprehensive Pilot Zone, this paper analyzes it as a policy variable to fill the gap in this aspect. Finally, although the existing literature generally agrees that the Internet and intellectualization play a positive role in economic growth, it does not reach a consensus on whether the digital divide will continue to deepen. Through the heterogeneity analysis, this paper examines whether the relatively backward regions can achieve leapfrog development through the development of digital industry under the background of unbalanced regional development in China.

Key words: big data; total factor productivity; National Big Data Comprehensive Pilot Zone

(责任编辑 康健)