

引用格式:石虹,余少龙.数据要素集聚如何驱动企业新质生产力发展?[J].技术经济,2024,43(12):35-46.

YU Shaolong, SHI Hong. How can data element clustering drive the development of new quality productivity in enterprises? [J]. Journal of Technology Economics, 2024, 43(12): 35-46.

数据要素集聚如何驱动企业新质生产力发展?

石虹,余少龙

(贵州大学经济学院,贵阳 550025)

摘要:基于2012—2022年沪深A股上市公司的数据测度企业新质生产力水平,以国家级大数据综合试验区作为准自然实验构造多期双重差分模型研究数据要素集聚对企业新质生产力的影响。研究表明,数据要素集聚推动了企业新质生产力发展,这一结论在PSM-DID、安慰剂检验等稳健性检验后仍成立。机制检验表明,数据要素集聚可以通过提高人力资本水平、推动绿色技术创新赋能企业新质生产力发展;随着行业竞争、媒体关注的增强,数据要素集聚对企业新质生产力的推动作用增强。异质性分析表明,数据要素集聚对企业新质生产力的推动作用在非国有企业、技术密集型企业、高科技行业以及数字基础设施建设较完善的地区中更加显著。研究结论为如何利用新型生产要素培育新质生产力提供了启示。

关键词:新质生产力;数据要素集聚;多期DID;准自然实验;高质量发展

中图分类号:F273 **文献标志码:**A **文章编号:**1002-980X(2024)12-0035-12

DOI:10.12404/j.issn.1002-980X.J24070516

一、引言

随着我国进入高质量发展阶段,传统生产力已经难以满足现阶段发展需求,如何培育与现阶段发展相匹配的新质生产力成为亟须回答的关键议题。自新质生产力这一概念首次提出以来,到2024年3月政府工作报告指出要“大力推进现代化产业体系建设,加快发展新质生产力”,新质生产力已然成为推进现代化产业体系建设,实现经济高质量发展的重要动力,然而,如何培育新质生产力仍处于探索阶段。数字经济背景下,数据要素推动了新旧动能转换、传统生产方式变革,逐渐成为高质量发展阶段的关键生产要素,而企业作为经济高质量发展的重要微观基础,以新型生产要素培育企业新质生产力应当是推动高质量发展的题中应有之义。为顺应数字经济发展潮流,推动大数据应用,2015年国务院印发《促进大数据发展行动纲要》,同年9月,在贵州省设立国家级大数据综合试验区,而后于2016年10月在京津冀地区、珠江三角洲、上海、河南、重庆、沈阳及内蒙古设立第二批大数据试验区,大数据试验区的设立为推广大数据应用,促进数据成为新型生产要素夯实了基础。2020年国务院印发《关于构建更加完善的要素市场化配置体制机制的意见》,正式将数据视作一种新型生产要素。根据中国信息通信研究院发布的《中国数字经济发展研究报告(2023年)》显示,2021年我国数据产量达到6.6泽字节,同比增加29.4%,占全球总量的9.9%,且现有实际运营中的数据交易所共30家,筹建中有6家。大数据试验区内数据要素集聚和应用的效果明显,以首个国家级大数据(贵州)综合试验区为例,截至2022年,贵州已建成的政府数据共享交换平台挂接数据资源1.38万个、信息项23.89万个,交换数据2.46万亿余条;搭建的贵州省数据流通交易平台入驻专业数据商近100家,上架数据产品和服务150余个。而在国家工业信息安全发展研究中心发布的《中国数据要素市场发展报告(2021—2022)》中,通过对726家上市公司共5000多个项目信息的分析发现,数据要素使得工业企业业务

收稿日期:2024-07-05

基金项目:贵州省理论创新课题专项项目“振兴贵州装备制造业对策研究”(GZLCZB-2022-3-10);环境保护部课题项目“促进部分‘双高’产品退出市场的经济政策研究”(2018A162);贵州省高校人文社会科学研究基地项目“马克思主义经济学与‘双碳’目标的地方实践研究”(23GZGXRWJD021)

作者简介:石虹,贵州大学经济学院副教授,硕士研究生导师,研究方向:可持续发展;余少龙,贵州大学经济学院硕士研究生,研究方向:数字经济。

平均增长41.18%,生产效率提高42.8%,产品研发周期缩短15.33%,能源利用率提高10.19%。综上,随着数据要素资源的增加,以及数据交易市场的完善,数据要素已在企业生产经营的各个方面发挥了显著作用,为培育企业新质生产力提供了可能。那么,数据要素集聚能否赋能企业新质生产力发展?其中有怎样的实现路径,又受到哪些因素的影响?研究以上问题有助于理清新型生产要素与新质生产力之间的关系,为发挥新质生产力区别于传统生产力的作用,助推高质量发展提供启示。

现有研究大多从理论方面阐述推动新质生产力发展的影响因素和新质生产力发展的经济后果,而关于如何使用新型生产要素培育新质生产力的研究较少。就数据要素如何发挥新型生产要素价值来看,数据要素是基于行动者网络、数字基础设施和制度构建三个方面进行即时价值转换和实现^[1],设立大数据试验区吸引众多大数据企业聚集,实现了数据的即时采集与实时共享,有助于完善行动者网络;大数据试验区内云计算、5G基站等数字基础设施建设的完善丰富了数据收集的范围和种类,加快了数据流通速度;而大数据试验区内相关数据政策则为数据要素产权界定、数据资源分配提供了制度保障。综上,设立大数据试验区有助于夯实数据要素价值实现的三大基础,实现数据要素集聚,为发挥数据要素这一新型生产要素的乘数效应,助力企业新质生产力发展提供了现实条件。因此,本文以大数据试验区作为准自然实验构造多期双重差分模型,研究数据要素集聚如何培育企业新质生产力。

与现有研究相比,本文边际贡献如下:第一,从新型生产要素的视角出发,研究数据要素集聚对企业新质生产力的影响,补充了数据要素、新质生产力的相关研究;第二,构造多期双重差分模型进行因果识别,一定程度上避免了变量测度误差等问题导致的估计结果偏误,较为准确地识别了数据要素集聚对企业新质生产力的影响;第三,基于企业内、外部视角检验数据要素集聚推动企业新质生产力发展的作用机制,以期打开数据要素影响企业新质生产力的机制“黑箱”。

二、文献回顾

作为一种新型生产要素,数据要素已成为推动经济高质量发展的核心动力,现有研究主要关注数据要素的赋能作用及其市场化的影响两个方面。就数据要素的赋能作用而言,从创新视角出发,数据要素可以通过提高传统生产要素的配置效率改善企业创新绩效^[2];且数据要素与传统生产要素融合形成的新型生产要素组合,可以推动数据要素价值实现,如数据要素与人力资本形成的要素组合促进了企业创新^[3-4];但借助数据要素进行创新活动也可能受制于数据要素“陷阱”,导致社会整体创新水平下降^[5]。从产业结构视角出发,数据要素通过优化生产要素配置推动产业结构调整^[6],同时数据要素与产业深度融合产生的数字技术红利,有助于促进产业结构升级^[7]。从区域发展来看,数据要素集聚可以缩小区域差距,推动区域协调发展^[8],助力共同富裕目标实现^[9]。就数据要素市场化的影响而言,在微观层面,数据要素市场化通过成本节约、融资约束以及技术创新效应推动了企业数字化转型,但这种推动作用呈现边际递减特征^[10];在中观层面,数据要素市场化通过推动产业融合、增强创新效能提高了产业数字化水平^[11];在宏观层面,数据要素市场化通过成本、要素配置及技术创新效应推动了数字经济与实体经济的融合^[12]。

现有关于新质生产力的研究主要从新质生产力的概念界定、如何培育新质生产力及新质生产力的经济后果三方面展开。就概念界定而言,学界普遍认为新质生产力是以科技创新为主导的生产力,是符合高质量发展阶段要求的生产力^[13],是包含新型劳动者、劳动资料和劳动对象的生产力。就如何培育新质生产力而言,已有研究表明高水平科技自立自强是新质生产力发展的关键因素^[14];而在金融方面,金融集聚推动了新质生产力发展^[15]。就新质生产力的经济后果而言,作为一种经济增长新动能,新质生产力可以推动经济高质量发展^[16]。

现有研究已从多方面探讨了数据要素的经济效应,且对新质生产力的概念进行了清晰的界定,以及从理论上为如何培育新质生产力提供了较为明确的方向。梳理上述数据要素、新质生产力的相关研究发现,创新既是数据要素赋能的主要对象,也是数字要素市场化产生影响的主要实现路径,那么数据要素集聚是否能赋能以创新为主导、以实现颠覆性创新突破为目标的新质生产力?鲜有研究关注到这一方面且尚未有研究通过实证检验数据要素集聚与新质生产力的关系。因此,本文构造多期双重差分模型研究数据要素集

聚对新质生产力的影响,能够为理清数据要素集聚与新质生产力二者关系提供实证参考,并从新型生产要素的角度补充企业新质生产力的影响因素研究。

三、理论分析与研究假设

(一)数据要素集聚与企业新质生产力

基于新型劳动者、新型劳动资料及新型劳动对象等新质生产力构成要素,对数据要素集聚如何培育企业新质生产力的展开分析。

就数据要素集聚与新型劳动者而言,数据要素集聚有助于数据要素与传统生产要素、传统生产工具的深度融合,由此催生出的新型劳动对象和新型劳动资料,要求与之适配的新型劳动者;因此,企业将员工的劳动者角色从简单劳动力提供者转变为创新和智能化工作的执行者^[17],即企业员工由传统劳动者向新型劳动者转型;另外,数据要素集聚配套发展的人工智能技术将很大程度替代传统劳动力,提高企业对数字化人才的需求,企业通过招聘高学历、高技能的劳动者或培养已有员工数字技能、数字素养将推动新型劳动者的出现,而新型劳动者使用新型劳动资料作用于新型劳动对象,实现了企业传统生产力向新质生产力的蜕变。

就数据要素集聚与新型劳动资料而言,与其他生产要素一样,数据要素需要与生产资料相结合,才能成为现实生产要素^[18]。一方面,数据要素集聚并应用于生产活动,拓展了传统劳动资料的范畴,加快了企业“机器换人”进程,有效实现了企业生产设备的智能化。并同时作用于企业生产经营各个环节,而生产环节的智能化有助于数据在不同部门间实时高效流动,实现了数据要素与其他生产要素的深度融合^[19],由此产生的规模收益递增有助于企业新质生产力发展。另一方面,数据要素集聚使得大数据、5G等配套技术发展更加完善,依托于数字技术,劳动资料逐渐数字化、智能化并向新型劳动资料转变,而新型劳动资料与资本、管理等生产要素的高效率结合,能够推动新质生产力发展。

就数据要素集聚与新型劳动对象而言,数据要素集聚既推动了数据要素作为生产要素投入企业生产经营活动,直接作用于企业新质生产力发展;又促进了数据要素与传统生产要素相结合,推动传统生产要素创新式变革演化,实现传统劳动对象向新型劳动对象的转变,促进企业新质生产力发展。一方面,作为新型生产要素,数据要素可视为一种新型劳动对象,区别于传统生产要素,数据要素具有低复制成本、非排他性、非竞争性等技术-经济特征^[20],不仅可以被重复利用、不断再生,而且在投入生产过程中产生新的数据可以实现数据要素的价值积累。另一方面,凭借高渗透性,数据要素融入生产活动并与传统生产要素相结合,有助于重构生产要素体系,拓展传统经济增长理论边界^[21];进而通过生产要素组合激发劳动、资本、管理等要素潜能,实现传统要素的价值增值,变革传统边际收益递减规律^[22],助推企业新质生产力发展。

基于以上分析,提出假设1:

数据要素集聚有助于企业培育新质生产力(H1)。

(二)数据要素集聚推动新质生产力的作用机制

从企业内部视角出发,基于企业内部的人力资本水平和绿色创新检验数据要素集聚推动新质生产力的实现路径。劳动者是生产力中最活跃的因素,提高人力资本水平有助于企业传统劳动者向新型劳动者转变,而数据要素集聚增强了数据要素对企业人力资本的替代效应和互补效应。从替代效应来看,数据要素与人工智能的融合对低技能劳动力产生挤出效应,减少了企业对低技能、低学历劳动力的需求,同时倒逼企业员工提高数字素养,增强劳动技能^[23],有助于改善企业人力资本结构,提高人力资本水平;另外,随着传统劳动要素的份额下降,数据要素推动形成的数据要素-传统要素组合,将强化数据要素的替代效应^[24],推动企业新质生产力发展。从互补效应来看,在数据要素赋能下,劳动要素属性得到拓展并进一步与数字技术结合,可以激发传统劳动力生产潜能,提高劳动生产率^[25],推动企业新质生产力发展。

区别于需要大量资源投入和高度消耗资源的生产力发展方式,新质生产力是科技创新发挥主导作用的生产力^[26]。因此,推动企业技术创新尤其是既能提高企业生产效率,又能减少资源消耗、降低环境污染进而兼顾经济、环境效益的绿色技术创新是企业发展新质生产力的关键。然而,由于资源要素流通阻塞导致的资源约束,企业绿色创新积极性不高,绿色创新能力和产出较低;数据要素的流通不受地理空间限制,在与

传统生产要素相结合后,可以有效破除生产要素流通壁垒,缓解创新资源约束,提高企业绿色创新积极性,推动企业绿色创新;数据要素集聚又进一步增强了数据要素的绿色属性,为企业绿色创新提供丰富的数据资源,提高了企业绿色创新产出质量。而绿色技术创新有助于促进传统高耗能、高污染的生产方式变革,实现企业生产绿色转型,为培育企业新质生产力提供技术支持。

基于上述分析,提出假设 2:

数据要素集聚通过提高人力资本水平、推动绿色创新促进企业新质生产力发展(H2)。

从企业外部视角出发,基于企业外部的行业竞争和媒体关注检验外部因素对数据要素集聚推动新质生产力的调节效应。就行业竞争来看,首先,大数据应用既有“同群效应”又存在“竞争效应”,达维多定律认为在市场竞争中,先进入市场可以带来较大的市场份额和高额利润。因此,优先使用数据要素进行生产经营活动并以此培育新质生产力的企业在行业竞争中会取得一定先发优势。其次,作为新型生产要素,数据要素有助于推动传统产业数字化转型,同时催生出人工智能、云计算等新兴产业,并通过数字经济与实体经济的深度融合重塑产业链结构,推动市场分工体系及产业链价值重新分配^[27],逐渐淘汰难以进行数字化转型或转型较慢的企业,企业面临来自行业内外部的双重竞争,行业竞争越激烈,企业利用数据要素发展新质生产力进而提高生产效率的意愿就越强。最后,随着数字技术的普及应用、数据要素价值的深入挖掘,数据已成为企业发展的关键动力,数据要素应用有助于企业成本控制和风险管理,使企业在日益激烈的行业竞争中保持竞争优势^[17]。因此,企业利用数据要素培育新质生产力既是企业发展的需要,也是企业保持竞争优势的必然要求。

就外部关注来看,作为信息传递媒介,媒体是公众、政府及资本市场接触企业信息的重要渠道,通过实时解读、报道企业相关信息对企业决策、生产经营产生外部监督作用^[28],有利于企业合理配置资源、进行科学决策,发展新质生产力。一方面,媒体关注会提高企业信息透明度,缓解企业与相关主体信息不对称问题,有助于加强各个主体对数据要素的协同应用,充分发挥数据要素作为新型生产要素的乘数效应,推动企业新质生产力发展。另一方面,通过媒体报道,企业发展新质生产力推动生产方式绿色化、数字化转型,有助于提高企业声誉价值;根据声誉理论,企业声誉提高会改善企业形象,赢得消费者、投资者的信任,有助于增强企业市场竞争力,这提高了企业发展新质生产力的积极性。

基于上述分析,提出假设 3:

行业竞争程度、媒体关注强度的提高会增强数据要素集聚对企业新质生产力的推动作用(H3)。

四、研究设计

(一) 样本选择与数据来源

以 2012—2022 年沪深 A 股上市企业作为研究对象,对企业数据处理如下:①剔除连续两个会计年度亏损(ST)、存在退市风险(*ST)、已退市(PT)及上市时间不足 3 年的企业样本;②剔除金融、保险行业及数据缺失过多的企业样本;③对连续变量进行 1% 和 99% 分位的缩尾处理,数据来源于中国研究数据服务平台(CNRDS)和国泰安(CSMAR)数据库。

(二) 模型设定

以大数据试验区为准自然实验,构建多期双重差分模型研究数据要素集聚对企业新质生产力的影响,模型设定如式(1)所示。

$$NQF_{it} = \alpha_0 + \alpha_1 DATA_{it} + \alpha_2 Control_{it} + FirmFE + YearFE + \varepsilon_{it} \quad (1)$$

其中: NQF_{it} 为 t 年 i 企业的新质生产力水平; $DATA_{it}$ 为数据要素集聚; $Control_{it}$ 为控制变量; $FirmFE$ 、 $YearFE$ 分别为企业、年份固定效应; α 为估计系数; ε_{it} 为随机扰动项。

(三) 变量选取

1. 被解释变量

企业新质生产力(NQF)。新质生产力发挥作用的过程就是新型劳动者通过新型劳动工具作用于新型劳动对象的过程^[29]。作为一种新型生产要素,数据要素本身可视作一种新型劳动对象。同时,数据要素在

与传统要素结合的过程中推动了传统劳动对象向新型劳动对象的转变。因此,结合现有研究对新质生产力水平的测度指标^[30-31]。本文从新型劳动者、新型劳动资料及新型劳动对象三方面构造新质生产力指标体系,使用熵值法测度企业新质生产力水平,具体指标如表1所示。

表1 指标体系

总指标	一级指标	二级指标	指标定义	权重(%)
新质 生产力	劳动者	研发人员占比	研发人员数量/总员工数量	8.43
		高学历人员占比	本科及以上学历人数/总员工数量	14.23
		独立董事占比	独立董事人数/董事会人数	0.97
	劳动资料	总资产周转率	营业收入/平均资产总额	1.70
		权益乘数倒数	所有者权益/总资产	0.61
		企业现金流	经营活动产生的现金流量/总资产	0.34
		创新产出	企业当年获得专利数量	12.08
		数字化无形资产占比	数字化无形资产/无形资产	27.19
		研发投入占比	(研发费用-直接投入)/营业收入	3.91
		研发投入折旧摊销占比	(研发费用-折旧摊销)/营业收入	6.49
		研发投入折旧租赁占比	(研发费用-租赁费)/营业收入	3.20
		劳动对象	固定资产占比	固定资产/总资产
	无形资产占比		无形资产/总资产	19.08
	制造费用占比		制造费用/费用支出	0.03

2. 核心解释变量

数据要素集聚(*DATA*)。以大数据试验区的试点地区企业虚拟变量(*Treat*)与试点时间虚拟变量(*Post*)的交互项衡量大数据试验区设立的政策效果,作为数据要素集聚的代理变量。具体而言,将大数据试验区试点地区企业作为实验组(*Treat* = 1),其他企业作为对照组(*Treat* = 0);根据试验区设立前后设置时间虚拟变量(*Post*),设立前为0,设立后为1。

3. 机制变量

人力资本水平(*Human*),以企业本科及以上学历员工数量占比衡量;绿色技术创新(*Patent*),参考已有研究,以企业当年绿色专利申请总量加1后取自然对数衡量^[32];行业竞争程度(*HHI*),参考已有研究,使用赫芬达尔指数衡量行业竞争程度,赫芬达尔指数值越大表示行业竞争程度越低,反之则行业竞争程度越高^[33];媒体关注(*Media*),参考已有研究,使用媒体报道标题中出现企业名称总数加1后取自然对数衡量^[34]。

4. 控制变量

为控制企业特征,选取控制变量如下:企业规模(*Size*),以总资产的自然对数衡量;资产负债率(*Lev*),以总负债与总资产的比值衡量;资产净利润率(*Roa*),净利润与总资产的比值衡量;营业收入增长率(*Growth*),以本年营业收入与上一年营业收入的比值减1衡量;大股东持股(*Top1*),以第一大股东持股比例衡量;董事会规模(*Board*),以董事会人数的自然对数衡量;企业年龄(*Firmage*),以当年年份减去公司成立年份后取自然对数衡量;两职合一(*Dual*),若董事长兼任总经理取值为1,反之为0。

表2 描述性统计

变量	观测值	平均值	标准差	最大值	最小值
<i>NQF</i>	26378	3.5929	2.8091	40.6093	0.1140
<i>DATA</i>	26378	0.2782	0.4481	1	0
<i>Size</i>	26378	22.3433	1.3039	26.4523	19.5850
<i>Lev</i>	26378	0.4284	0.2006	0.9079	0.0349
<i>Roa</i>	26378	0.0397	0.0645	0.2473	-0.3730
<i>Growth</i>	26378	0.1604	0.4061	4.0242	-0.6576
<i>Top1</i>	26378	0.3430	0.1472	0.7578	0.0802
<i>Board</i>	26378	2.1216	0.1959	2.7081	1.6094
<i>Firmage</i>	26378	2.9433	0.3200	3.6109	1.6094
<i>Dual</i>	26378	0.2734	0.4457	1	0

续表

变量	观测值	平均值	标准差	最大值	最小值
<i>Patent</i>	26378	0.0280	0.1035	1	0
<i>Human</i>	26378	0.1088	0.1359	0.9971	0
<i>HHI</i>	26378	0.1384	0.1497	1	0
<i>Media</i>	26378	4.3972	1.1054	8.9574	0

五、实证结果分析

(一) 基准回归分析

如表 3 所示, (1) 列中未考虑企业特征时, 数据要素集聚对企业新质生产力的估计系数为 0.1897 且在 1% 的水平上显著; (2) 列中加入控制变量控制企业特征后, 数据要素集聚的估计系数为 0.1616 仍在 1% 的水平上显著, 这表明数据要素集聚可以显著推动企业新质生产力发展。政府部门应当重视大数据试验区设立引发的数据要素集聚, 充分发挥数据要素的“乘数效应”, 以新型生产要素培育新质生产力; 而企业则应积极响应政府相关政策, 调整企业发展战略, 借助数据要素集聚这一有利外部条件来发展新质生产力, 推动企业高质量发展。假设 1 得到证实。

表 3 基准回归结果

变量	(1)	(2)
	<i>NQF</i>	<i>NQF</i>
<i>DATA</i>	0.1897 ^{***} (0.0368)	0.1616 ^{***} (0.0355)
<i>Size</i>		0.0703 ^{***} (0.0273)
<i>Lev</i>		-0.7032 ^{***} (0.1158)
<i>Roa</i>		-0.8847 ^{***} (0.3033)
<i>Growth</i>		0.0641 ^{***} (0.0244)
<i>Top1</i>		-0.2337 (0.1538)
<i>Board</i>		-2.4985 ^{***} (0.1094)
<i>Dual</i>		-0.0231 (0.0321)
<i>Firmage</i>		0.9922 ^{***} (0.1597)
<i>Constant</i>	3.5402 ^{***} (0.0128)	4.7692 ^{***} (0.7387)
企业固定	是	是
年份固定	是	是
<i>Obs</i>	26378	26378
<i>R</i> ²	0.784	0.792

注: *、**、*** 分别表示 10%、5%、1% 的显著性水平, 括号内为稳健标准误。

(二) 稳健性检验

1. 平行趋势检验

使用多期双重差分模型进行因果识别需要满足平行趋势假设, 即在政策冲击前实验组和对照组的研究对象保持一致的变化趋势, 但大数据试验区并非同一年设立, 需要设定不同政策冲击年份的时间虚拟变量。使用事件研究法进行平行趋势检验, 模型设定如式(2)所示。

$$NQF_{it} = \beta_0 + \beta_1 Before4_{it} + \beta_2 Before3_{it} + \beta_3 Before2_{it} + \beta_4 Before1_{it} + \beta_5 Current_{it} + \beta_6 After1_{it} + \beta_7 After2_{it} + \beta_8 After3_{it} + \beta_9 After4_{it} + \beta_{10} Control_{it} + FirmFE + YearFE + \varepsilon_{it} \quad (2)$$

其中: $Before_{it}$ 为大数据试验区设立前的虚拟变量; $Current_{it}$ 为设立当期的虚拟变量; $After_{it}$ 为设立后的虚拟变量; β 为估计系数; $Before1_{it} \sim Before4_{it}$ 为大数据试验区设立前 1~4 年的虚拟变量; $After1_{it} \sim After4_{it}$ 为大数据试验区设立后 1~4 年的虚拟变量。如图 1 所示, 试验区设立前的估计系数均不显著, 这表明大数据试验区设立前, 实验组和对照组企业在新质生产力水平上无显著差异, 通过平行趋势检验; 试验区设立当年的估计系数不显著, 而试验区设立后几年的估计系数均显著为正, 这表明数据要素集聚可以显著推动企业新质生产力发展, 但该推动作用具有一定的时滞性。上述结果表明, 企业新质生产力水平提高不是由于实验组与对照组的系统性差异, 而是由大数据试验区设立引发的数据要素集聚所致。试验区设立后系数值逐渐变小且第四期的系数明显变小, 这可能是由于在新质生产力发展初期, 企业尚未形成系统性的新质生产力培育模式, 企业使用新型生产要素来组织生产时, 数据要素集聚对新质生产力的推动作用较为明显; 而当新质生产力发展到一定水平时, 数据要素集聚虽仍有显著推动作用, 但企业新质生产力发展可能更多地依靠相关政策或其他因素的助力。

2. 安慰剂检验

考虑可能存在潜在因素或其他政策对估计结果造成影响。随机抽取 1000 家企业作为实验组, 其他企业作为对照组, 设定虚拟政策冲击时间来构造虚拟的政策交互项以基准回归模型进行估计, 并重复上述过程

500次进行安慰剂检验。如图2所示,安慰剂检验的系数值聚集于0附近且 $P>0.1$ 即在10%的水平上不显著,同时基准回归中系数为0.1616与安慰剂检验结果有较大差异,这表明前文估计结果未受到其他潜在因素或政策的影响,数据要素集聚可以推动企业新质生产力发展这一结论具有稳健性。

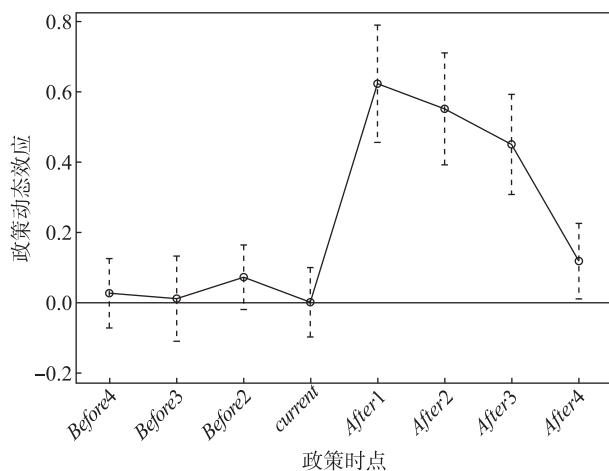


图1 平行趋势检验

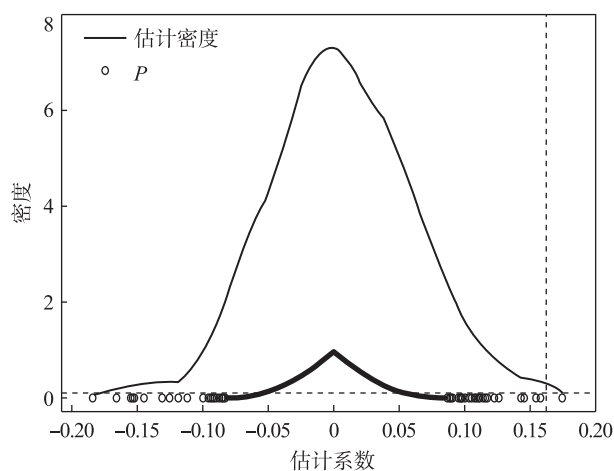


图2 安慰剂检验

3. 倾向得分匹配法

由于大数据试验区并非真正的自然实验,仍可能存在选择性偏差导致估计结果偏误。使用倾向得分匹配法(PSM)缓解样本自选择偏误导致的估计结果偏差,首先,以控制变量为协变量构造Logit模型,采用1:1最近邻匹配方法进行样本匹配,再以基准回归模型进行估计。如表4所示,(1)列中数据要素集聚的估计系数显著为正,这表明PSM-DID的估计结果与前文结果基本一致,数据要素集聚可以显著推动企业新质生产力发展。

4. 排除同期政策

考虑到“宽带中国”试点政策与大数据试验区均有推进数字基础设施建设的效果,本文在基准回归模型中加入“宽带中国”政策项进行稳健性检验。如表4所示,(2)列中数据要素集聚的估计系数显著为正,这表明排除同期政策影响后,数据要素集聚仍对企业新质生产力有促进作用。

5. 政策外生性

多期双重差分模型需要满足政策外生性假设,但大数据试验区设立前,企业等相关主体可能会进行预期准备工作。因此,在基准回归中加入试验区设立前一年的政策项(Before1)进行稳健性检验。如表4所示,(3)列中前一年政策项系数不显著且数据要素集聚的估计系数显著为正,这表明控制企业等主体的预期效应后,数据要素集聚对企业新质生产力发展仍有积极影响。

6. 剔除直辖市企业样本

相较于地级市,直辖市拥有较好的基础设施水平、营商环境,为企业发展新质生产力提供了良好的外部环境,因此,本文剔除直辖市企业样本进行稳健性检验。如表4所示,(4)列中数据要素集聚的估计系数显著为正,这表明剔除直辖市企业样本后,数据要素集聚仍可显著推动企业新质生产力发展。上述稳健性检验结果加强了前文结论的稳健性。

表4 稳健性检验

变量	(1)	(2)	(3)	(4)
	<i>NQF</i>	<i>NQF</i>	<i>NQF</i>	<i>NQF</i>
DATA	0.1418 ^{***} (0.0544)	0.1616 ^{***} (0.0355)	0.1622 ^{***} (0.0356)	0.2271 ^{***} (0.0465)
“宽带中国”		-0.0045(0.0317)		
Before1			0.0253(0.0429)	
Constant	5.2300 ^{***} (1.1000)	4.7738 ^{***} (0.7399)	4.7691 ^{***} (0.7387)	6.3890 ^{***} (0.7783)
控制变量	是	是	是	是

续表

变量	(1)	(2)	(3)	(4)
	<i>NQF</i>	<i>NQF</i>	<i>NQF</i>	<i>NQF</i>
企业固定	是	是	是	是
年份固定	是	是	是	是
<i>Obs</i>	14352	26378	26378	20997
<i>R</i> ²	0.802	0.792	0.792	0.795

注：*、**、*** 分别表示 10%、5%、1% 的显著性水平；括号内为稳健标准误。

7. 培根分解诊断

使用分批设立的大数据试验区构建多期双重差分模型进行估计时,可能会出现将早期实验组作为晚期实验组的对照组,进而导致“负权重”问题。因此,借鉴已有研究的方法进行培根分解诊断^[35]。如表 5 所示,以较早和尚未设立试验区的地区企业作为对照组的权重分别为 0.009、0.100;而以从未设立试验区的地区企业作为对照组的权重为 0.891,这表明使用大数据试验区设立地区企业作为实验组、其他地区企业作为对照组可以较为准确的识别数据要素集聚对企业新质生产力的影响。

8. 两阶段双重差分模型

双重差分模型在使用控制个体、时间双向固定效应进行估计时,需要满足同质性处理假设,而基于分批设立的大数据试验区来构造多期双重差分模型,得到的估计值是不同实验组在不同时期处理效应的加权平均值,由于难以避免的“负权重”问题,导致估计结果存在偏误^[36]。借鉴现有研究,使用两阶段双重差分模型缓解异质性处理效应^[37]。首先,估计移除个体、时间固定效应的残差值;再将所估计的残差值与政策交互项进行回归得到平均处理效应。如表 5 所示,数据要素集聚的估计系数显著为正,这表明以两阶段双重差分模型缓解异质性处理问题后,数据要素集聚可以促进企业新质生产力发展这一结论仍成立。

表 5 模型有效性检验

(1) 培根分解诊断			(2) 两阶段双重差分模型	
作为控制组类型	权重	处理效应	变量	<i>NQF</i>
“较早接受的处理组”	0.009	-0.506	<i>DATA</i>	0.5839 ***
“尚未接受的处理组”	0.100	0.690		(0.0645)
“从未接受的处理组”	0.891	0.211	控制变量/企业/年份	是

注：*、**、*** 分别表示 10%、5%、1% 的显著性水平；括号内为稳健标准误。

(三) 机制检验

1. 实现路径

从企业内部视角出发,检验数据要素集聚促进企业新质生产力发展的实现路径。借鉴已有研究^[38],构造实证模型检验数据要素集聚对企业人力资本水平、绿色技术创新的影响,再结合前文理论部分的论述理清实现路径。模型设定如式(3)所示。

$$M_{it} = \mu_0 + \mu_1 DATA_{it} + \mu_2 Control_{it} + FirmFE + YearFE + \varepsilon_{it} \quad (3)$$

其中: M_{it} 为人力资本水平、绿色技术创新; μ 为估计系数。

如表 6 所示,(1)列中数据要素集聚对人力资本水平的估计系数为 0.0115 且在 1% 的水平上显著,这表明数据要素集聚有助于改善企业人力资本结构,提高人力资本水平。(2)列中数据要素集聚对绿色技术创新的估计系数为 0.0046 且在 5% 的水平上显著,这表明数据要素集聚可以推动企业绿色技术创新。设立大数据试验区不仅有助于数据要素的集聚,还可以汇集

表 6 实现路径检验

变量	(1)	(2)
	<i>Human</i>	<i>Patent</i>
<i>DATA</i>	0.0115 *** (0.0020)	0.0046 ** (0.0023)
<i>Constant</i>	-0.5121 *** (0.0446)	0.0618 (0.0466)
控制变量	是	是
企业固定	是	是
年份固定	是	是
<i>Obs</i>	26378	26378
<i>R</i> ²	0.791	0.420

注：*、**、*** 分别表示 10%、5%、1% 的显著性水平；括号内为稳健标准误。

高新技术人才,为企业招聘可以推动新质生产力发展的技术研发人员提供了良好外部环境。而新型劳动者参与企业技术研发,可以推进包括绿色技术创新在内的创新活动,为企业新质生产力发展提供人才支撑;同时,企业应用绿色技术有助于实现生产方式绿色转型,变革高耗能高污染的生产模式,为新质生产力发展提供技术支持。实证结果结合上述分析表明,数据要素集聚可以通过提高企业人力资本水平、推动绿色技术创新,实现企业传统劳动者向新型劳动者的转变以及生产技术的变革,进而促进企业新质生产力发展。假设2得到证实。

2. 调节效应

从企业外部视角出发,检验行业竞争程度和媒体关注强度在数据要素集聚推动企业新质生产力发展过程中的调节效应。模型设定如式(4)所示。

$$NQF_{it} = \lambda_0 + \lambda_1 DATA_{it} + \lambda_2 Med_{it} + \lambda_3 DATA_{it} \times Med_{it} + \lambda_4 Control_{it} + FirmFE + YearFE + \varepsilon_{it} \quad (4)$$

其中: Med_{it} 为行业竞争程度、媒体关注强度; λ 为估计系数。

表7 调节效应检验

变量	(1)	(2)
	NQF	NQF
$DATA$	0.1666 ^{***} (0.0356)	0.1823 ^{***} (0.0366)
HHI	-0.4958 ^{***} (0.1035)	
$DATA \times HHI$	-1.1032 ^{***} (0.1694)	
$Media$		0.0194 (0.0130)
$DATA \times Media$		0.0855 ^{***} (0.0266)
$Constant$	4.7421 ^{***} (0.7357)	4.7898 ^{***} (0.7381)
控制变量	是	是
企业固定	是	是
年份固定	是	是
Obs	26378	26378
R^2	0.793	0.792

注: *、**、*** 分别表示 10%、5%、1% 的显著性水平;括号内为稳健标准误。

如表7所示,(1)列中数据要素集聚与行业竞争程度的交互项估计系数为-1.1032且在1%的水平上显著,这表明行业竞争程度在数据要素集聚推动企业新质生产力发展过程中发挥了显著的调节效应,由于行业竞争程度为负向指标,指标越小则行业竞争程度越高。因此,行业竞争程度越高,则数据要素集聚对企业新质生产力的推动作用越强。(2)列中数据要素集聚与媒体关注强度的交互项估计系数为0.0855且在1%的水平上显著,这表明媒体关注在数据要素集聚推动企业新质生产力发展过程中发挥了正向调节效应,即随着媒体关注不断加强,数据要素集聚对企业新质生产力的推动作用增强。假设3得到证实。

(四) 异质性分析

1. 所有权性质

就企业特征而言,不同所有权性质的企业在新质生产力发展上存在较大差异。因此,依据企业所有权性质将样本划分为国有、非国有企业进行分组回归。如表8所示,(1)列、(2)列中数据要素集聚对国有企业新质生产力发展的估计系数不显著,但对非国有企业的估计系数显著为正,这表明相较于国有企业,数据要素集聚对非国有企业的新质生产力发展有显著推动作用。这可能是由于新质生产力的发展不仅需要新型生产要素,也需要与之相匹配的新型劳动者,而国有企业的人力结构相对固化,传统劳动者向新型劳动者的转变较为缓慢,数据要素集聚对国有企业新质生产力的影响尚未显著发挥出来。

表8 企业层面异质性

变量	(1)	(2)	(3)	(4)	(5)
	国有	非国有	劳动密集型	技术密集型	资本密集型
	NQF	NQF	NQF	NQF	NQF
$DATA$	-0.0857(0.0833)	0.3629 ^{***} (0.0870)	-0.1617 [*] (0.0872)	0.4698 ^{***} (0.1004)	0.0029(0.0904)
$Constant$	4.3682 ^{**} (1.7373)	4.0051 ^{***} (1.4393)	1.1406(1.7230)	6.8915 ^{***} (1.8346)	9.5972 ^{***} (1.7481)
控制变量	是	是	是	是	是
企业固定	是	是	是	是	是
年份固定	是	是	是	是	是
Obs	9866	16512	9172	12013	5193
R^2	0.735	0.691	0.692	0.703	0.796

注: *、**、*** 分别表示 10%、5%、1% 的显著性水平;括号内为稳健标准误。

2. 企业类型

就企业依靠的主要生产要素类型而言,技术密集型企业相较于劳动、资本密集型企业对数据要素的依赖度更强。因此,依据企业生产要素密集度将样本划分为劳动、技术及资本密集型企业进行分组回归。如表 8 所示,(3)列中数据要素集聚的估计系数仅在 10%的水平上显著,(5)列中数据要素集聚的估计系数不显著;而(4)列中数据要素集聚的估计系数在 1%的水平上显著为正,这表明相较于劳动、资本密集型企业,数据要素集聚可以显著推动技术密集型企业新质生产力发展。

3. 行业属性

就行业特征而言,高科技行业拥有较为先进的技术设备及高效的研发团队,在新型劳动者与新型生产资料上存在较大优势。参考已有研究,根据所属行业^①将样本划分为高科技、非高科技行业进行分组回归^[39]。如表 9 所示,(1)列中数据要素集聚的估计系数显著为正,(2)列中数据要素集聚的估计系数不显著,这表明数据要素集聚对新质生产力发展的推动作用在高科技行业中更加显著。高科技行业在研发人员等新型劳动者和人工智能设备等新型劳动资料上存在比较优势,对数据要素这一新型生产要素的利用更加合理高效,新型劳动者、劳动资料以及劳动对象的结合在高科技行业中更加密切,因此,数据要素集聚显著推动了高科技行业新质生产力发展。

4. 数字基础设施建设

就地区特征而言,地区数字基础设施建设越完善,越有利于发挥数据要素的“乘数效应”。因此,以互联网普及率作为数字基础设施建设的代理变量,再以数字基础设施建设的中位数划分样本进行分组回归。如表 9 所示,(3)列中数据要素集聚的估计系数显著为正,(4)列中数据要素集聚的估计系数不显著,这表明在数字基础设施建设较为完善的地区,数据要素集聚对企业新质生产力的推动作用更明显。

表 9 行业、地区层面异质性

变量	(1)	(2)	(3)	(4)
	高科技	非高科技	较完善	较薄弱
	<i>NQF</i>	<i>NQF</i>	<i>NQF</i>	<i>NQF</i>
<i>DATA</i>	0.4333*** (0.0851)	-0.1013 (0.0692)	0.2074*** (0.0789)	0.1617 (0.1349)
<i>Constant</i>	6.6876*** (1.4307)	3.1159** (1.3189)	3.6091** (1.5003)	8.9129*** (1.8227)
控制变量	是	是	是	是
企业固定	是	是	是	是
年份固定	是	是	是	是
<i>Obs</i>	15472	10906	13367	13011
<i>R</i> ²	0.705	0.741	0.701	0.688

注: *、**、*** 分别表示 10%、5%、1% 的显著性水平;括号内为稳健标准误。

七、结论与启示

现有研究关于数据要素与新质生产力关系的实证研究较少,且对企业新质生产力水平的测度尚未统一。本文基于 2012—2022 年沪深 A 股上市公司的数据,从新型劳动者、劳动资料及劳动对象出发构建指标衡量体系对企业新质生产力水平进行测度;区别于现有研究使用企业年报词频或构建指标衡量体系测度数据要素这一指标,本文基于大数据试验区这一准自然实验构造多期 DID 研究数据要素集聚对企业新质生产力的影响和作用机制,以期识别数据要素集聚提供新的思路,同时深化对数据要素集聚与企业新质生产力发展间关系的认识。研究发现:第一,数据要素集聚可以显著推动企业新质生产力发展,该结论在安慰剂检验、PSM-DID、培根分解诊断等一系列稳健性检验后仍成立。第二,人力资本水平、绿色技术创新是数据要素集聚推动企业新质生产力发展的实现路径;同时,数据要素集聚对企业新质生产力的推动作用受到行业竞争程度、媒体关注强度的调节作用。第三,在非国有企业、技术密集型企业、高科技行业以及数字基础设施

① 经济合作与发展组织(OECD)规定的高科技行业包括计算机、电子、信息技术、生物制药及通信行业,结合《上市公司行业分类指引》,高科技行业包括:C25、C26、C27、C28、C29、C31、C32、C34、C35、C36、C37、C38、C39、C40、C41、I63、I64、I65 和 M73。

施建设较为完善地区中,数据要素集聚对企业新质生产力的推动作用更加显著。

基于研究结论,有以下政策启示:

第一,强化数据要素供给,以新型生产要素培育新质生产力。首先,政府部门应当拓展数据要素规模,完善数据要素交易市场,深挖数据要素潜力,为发挥数据要素作用夯实基础。具体而言,拓宽大数据试验区的辐射范围,推广大数据试验区建设经验,加快5G、工业互联网等“新基建”的建设,为数据要素资源采集、数据要素价值挖掘提供底层支撑;深化数据要素市场化改革,统筹数据交易平台建设,推动形成数据要素全国统一大市场,实现数据要素跨行业、跨区域流通,合理配置数据要素资源,释放数据要素价值,助力企业培育新质生产力。其次,企业方面应当响应政府号召,积极引入数字技术、智能设备并完善员工培训和人才引进机制,推动生产工具的数字化、智能化,强化数据挖掘能力,提高数据要素利用效率,推动数据要素与传统生产要素的组合应用,培育企业新质生产力。最后,加强政府、企业的协同合作,并逐步扩展到各个领域的协作共赢,推动数据技术产学研合作平台、数据共享交换平台建设,打破不同企业、不同行业间数据要素流通壁垒,实现数据潜力挖掘、数据要素应用在各领域协作共赢。

第二,发挥数据要素作用,畅通实现路径堵点;重视企业外部因素,强化数据赋能效果。加强数字人才引进以及员工技能培养,提高企业人力资本水平,为利用数据要素提供人才支持;提高企业绿色创新能力,增加绿色创新产出,为使用数据要素培育新质生产力提供技术支撑。发挥行业竞争、媒体关注在数据要素赋能企业新质生产力发展中的作用,应当充分认识到变革传统生产方式、发展新质生产力才能保持竞争优势,树立良好的企业形象,提高企业声誉价值。

第三,避免盲目照搬现有新质生产力发展模式,做到因地制宜、因企施策。充分考虑不同企业、不同行业以及不同地区存在的差异性,有针对性的推动数据要素发挥对新质生产力的赋能作用。同时,企业应当加强对新质生产力经济后果的认识,调整企业战略方向,借助数字经济发展浪潮,以新型生产要素培育新质生产力,助力企业高质量发展。

参考文献

- [1] 谢康,张祎,吴瑶.数据要素如何产生即时价值:企业与用户互动视角[J].中国工业经济,2023(11):137-154.
- [2] 宋炜,曹文静,周勇.数据要素赋能、研发决策与创新绩效——来自中国工业的经验证据[J].管理评论,2023,35(7):112-121.
- [3] 刘启雷,张媛,雷雨嫣,等.数字化赋能企业创新的过程、逻辑及机制研究[J].科学学研究,2022,40(1):150-159.
- [4] 陶长琪,丁煜.数据要素何以成为创新红利?——源于人力资本匹配的证据[J].中国软科学,2022(5):45-56.
- [5] 徐翔,赵墨非,李涛,等.数据要素与企业创新:基于研发竞争的视角[J].经济研究,2023,58(2):39-56.
- [6] 朱晓武,魏文石,王靖雯.数据要素、新型基础设施与产业结构调整路径[J].南方经济,2024(1):107-123.
- [7] 陈晓佳,徐玮.数据要素、交通基础设施与产业结构升级——基于量化空间一般均衡模型分析[J].管理世界,2024,40(4):78-98.
- [8] 苏婧,赵城,王必达.数据要素集聚能实现区域协调发展吗——论大国区域发展的“集聚”与“协调”[J].财经科学,2022(6):62-77.
- [9] 彭国柱,周湘莲.数据要素配置与共同富裕:基于国家级大数据综合试验区的准自然实验[J].经济问题探索,2024(4):87-102.
- [10] 郑国强,张馨元,赵新宇.数据要素市场化如何驱动企业数字化转型?[J].产业经济研究,2023(2):56-68.
- [11] 陈寒钰,赵紫凤,刘超.数据要素市场建设促进了产业数字化发展吗——基于数据交易平台的准自然实验[J].宏观经济研究,2024(1):61-74.
- [12] 王晓丹,石玉堂,刘达.数据要素市场化配置对数实融合的影响研究——基于数据交易平台设立的准自然实验[J].广东财经大学学报,2024,39(2):44-58.
- [13] 胡洪彬.习近平总书记关于新质生产力重要论述的理论逻辑与实践进路[J].经济学家,2023(12):16-25.
- [14] 周文康.高水平科技自立自强助推新质生产力形成发展:理据、优势与进路[J].技术经济,2024,43(4):15-25.
- [15] 任宇新,吴艳,伍喆.金融集聚、产学研合作与新质生产力[J].财经理论与实践,2024,45(3):27-34.
- [16] 徐政,郑霖豪,程梦瑶.新质生产力赋能高质量发展的内在逻辑与实践构想[J].当代经济研究,2023(11):51-58.
- [17] 刘明慧,李秋.财税政策何以驱动新质生产力发展?[J].上海经济研究,2024(3):31-41.
- [18] 谢康,夏正豪,肖静华.大数据成为现实生产要素的企业实现机制:产品创新视角[J].中国工业经济,2020(5):42-60.
- [19] 钞小静,王宸威.数据要素对制造业高质量发展的影响——来自制造业上市公司微观视角的经验证据[J].浙江工商大学学报,2022(4):109-122.
- [20] 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.
- [21] 陈晓红,李杨扬,宋丽洁,等.数字经济理论体系与研究展望[J].管理世界,2022,38(2):208-224,13-16.
- [22] 王谦,付晓东.数据要素赋能经济增长机制探究[J].上海经济研究,2021(4):55-66.
- [23] 于柳管,高煜.数据要素如何驱动制造业生产率提升[J].财经科学,2024(1):76-90.

- [24] ACEMOGLU D, RESTREPO P. The race between man and machine: Implications of technology for growth, factor shares, and employment[J]. *American Economic Review*, 2018, 108(6): 1488-1542.
- [25] 杜传忠, 疏爽, 李泽浩. 新质生产力促进经济高质量发展的机制分析与实现路径[J]. *经济纵横*, 2023(12): 20-28.
- [26] 周文, 许凌云. 论新质生产力: 内涵特征与重要着力点[J]. *改革*, 2023(10): 1-13.
- [27] 刘培. 数据要素影响绿色生态效率的机制与效应研究[J]. *经济经纬*, 2023, 40(6): 136-148.
- [28] 邹颖, 祁亚, 石福安. 共同机构投资者改善了投融资期限错配吗? ——基于供应链关系治理机制[J]. *上海财经大学学报*, 2024, 26(2): 18-34.
- [29] 黄群慧, 盛方富. 新质生产力系统: 要素特质、结构承载与功能取向[J]. *改革*, 2024(2): 15-24.
- [30] 韩文龙, 张瑞生, 赵峰. 新质生产力水平测算与中国经济增长新动能[J]. *数量经济技术经济研究*, 2024, 41(6): 5-25.
- [31] 宋佳, 张金昌, 潘艺. ESG 发展对企业新质生产力影响的研究——来自中国 A 股上市企业的经验证据[J]. *当代经济管理*, 2024, 46(6): 1-11.
- [32] 罗双成, 刘建江, 熊智桥. 人才政策支持与重污染企业绿色创新绩效——来自高层次人才补助的经验证据[J]. *产业经济研究*, 2024(1): 56-70.
- [33] 钟廷勇, 马富祺, 唐嘉尉, 等. 国家审计能够降低企业碳风险吗? [J]. *审计研究*, 2023(4): 41-54.
- [34] 魏延鹏, 毛志宏, 王浩宇. 国有资本参股对民营企业 ESG 表现的影响研究[J]. *管理学报*, 2023, 20(7): 984-993.
- [35] GOODMAN-BACON A. Difference-in-differences with variation in treatment timing[J]. *Journal of Econometrics*, 2021, 225(2): 254-277.
- [36] DE CHAISEMARTIN C, D'HAULTFOEUILLE X. Two-way fixed effects estimators with heterogeneous treatment effects[J]. *American Economic Review*, 2019, 110(9): 2964-2996.
- [37] GARDNER J. Tow-stage differences in differences[R]. Cambridge: NBER, Working Paper, 2021.
- [38] 江艇. 因果推断经验研究中的中介效应与调节效应[J]. *中国工业经济*, 2022(5): 100-120.
- [39] 彭红星, 毛新述. 政府创新补贴、公司高管背景与研发投入——来自我国高科技行业的经验证据[J]. *财贸经济*, 2017, 38(3): 147-161.

How Can Data Element Clustering Drive the Development of New Quality Productive Forces in Enterprises?

Shi Hong, Yu Shaolong

(School of Economics, Guizhou University, Guiyang 550025, China)

Abstract: Based on the data of A-share listed companies in Shanghai and Shenzhen from 2012 to 2022 to measure the level of new quality productive forces of enterprises (NQP), a multi-period difference-in-differences model was constructed to study the impact of data factor agglomeration on the new quality productive forces of enterprises with the national-level big data comprehensive experimental zone as a quasi-natural experiment. It shows that data factor agglomeration promotes the development of new quality productive forces of enterprises, and this conclusion still holds after PSM-DID, placebo test and other robustness tests. Mechanism tests show that data factor agglomeration can empower the development of firms' new quality productive forces by improving human capital level and promoting green technology innovation; with the increase of industry competition and media attention, the role of data factor agglomeration in promoting firms' new quality productive forces increases. Heterogeneity analysis shows that the effect of data factor agglomeration on new productivity of enterprises is more significant in non-state-owned enterprises, technology-intensive enterprises, high-tech industries and regions with better digital infrastructure. The findings provide insights into how to utilize new factors of production to cultivate new productivity.

Keywords: new quality productive forces; data element agglomeration; multi-period DID; quasi-natural experiments; high-quality development