

DOI: 10.13253/j.cnki.ddjgl.2024.05.004

大数据发展与企业创新能力提升

陈蕊¹, 王宏伟^{1,2,3}

(1. 中国社会科学院数量经济与技术经济研究所, 北京 100732; 2. 中国社会科学院项目评估与战略规划研究咨询中心, 北京 100732; 3. 中国社会科学院大学, 北京 102488)

[摘要] 大数据发展提高了信息交流与知识共享的速度及效率, 成为推动中国建设世界科技强国, 实现高水平科技自立自强的核心动力。文章基于2012—2021年A股上市制造业企业数据, 探究大数据发展对企业创新能力的影响及中介机制。研究表明: 大数据发展显著促进了企业创新能力提升, 该结论在控制内生性问题及稳健性检验后依然成立。机制分析表明, 大数据发展通过缓解企业融资约束、优化人力资本结构和提高管理效率三种途径提高企业创新能力。进一步研究发现: 劳动力市场、资本市场与技术市场等要素细分市场发展增强了大数据发展对企业创新的作用效果, 且大数据发展对企业创新能力的促进作用在东部地区、地方政府数据开放水平高的地区、国有企业和大型企业中的效果更明显。通过验证大数据发展在企业创新能力提升中的重要作用, 为大数据发展与企业核心技术突破提供了有益的参考依据。

[关键词] 大数据发展; 企业创新能力; 要素市场; 制造业

[中图分类号] F273.1

[文献标识码] A

[文章编号] 1673-0461(2024)05-0030-13

一、引言

企业作为我国科技创新体系的重要主体, 其创新能力是国家竞争力、创造力和综合国力提升的动力源泉。为增强企业技术创新水平, 培育发展动能, 厚植发展优势, “十四五”规划纲要指出要“提升企业技术创新能力”, 并提出要“强化企业创新主体地位, 促进各类创新要素向企业集聚, 形成以企业为主体、市场为导向、产学研深度融合的技术创新体系”。但目前我国企业普遍存在创新能力不足、创新成果质量低、核心技术对外依存度高等问题^[1], 所以, 如何提高企业技术创新水平对增强企业市场竞争力和国民经济高质量发展至关重要。现有研究主要分析了市场化水平、市场竞争、数字经济、数字化转型等多种因素在企业技术创新中的关键作用^[2-5]。然而, 在我国大数据相关政策和措施加速推进的背

景下, 正处于快速发展阶段的大数据发展对企业创新能力可能产生的影响尚未引起学术界足够关注。

大数据技术主要通过采集、储存、加工和分析实现数据要素的核心价值, 并提取高价值的知识信息^[6], 而知识是创新的重要要素, 所以, 大数据发展在提高企业技术创新水平中发挥着非常关键的作用。理论层面上, 大数据发展能够引起生产研发模式转型^[7], 为企业拓展已有知识的搜索空间, 并利用数据分析组合现有技术来为创造新技术提供相关支持^[8], 不仅能够缓解企业融资约束、加强金融集聚、优化人力资本结构, 还能降低信息不对称和改善公司治理水平^[9-12], 进而可能对企业创新活动产生影响。实践层面上, 我国大数据产业规模从2016年的2480.6亿元增长到2021年的1.3万亿元, 平均增长率为46.26%^①, 大数据产业价值不断提升, 增长潜力巨大。因此,

收稿日期: 2023-11-10

基金项目: 国家自然科学基金专项项目“面向2040消费需求的重点领域工程科技发展方向研究”(L222400054); 国家自然科学基金重大项目“建设人才强国背景下激发科技人才创新活力研究”(21ZDA014)。

作者简介: 陈蕊(1993—), 女, 安徽六安人, 中国社会科学院数量经济与技术经济研究所博士后, 研究方向为科技创新与企业创新; 王宏伟(1970—), 女, 黑龙江萝北人, 中国社会科学院数量经济与技术经济研究所研究员、博士生导师, 研究方向为科技创新与经济增长、科技创新政策分析和评估。

大数据发展可能是提高企业技术创新水平的一个关键途径, 挖掘大数据发展影响企业创新能力的基本规律是值得研究的重要课题。基于此, 本文从理论和实证两个方面分析大数据发展对企业创新能力的影响效应, 以期从大数据发展视角为提高中国企业技术创新水平提供理论解释和经验证据。

和已有文献相比, 本文的贡献主要包括, 第一, 构建综合指标衡量大数据发展水平, 探究大数据发展如何作用于企业创新能力提升, 为大数据发展影响企业创新能力的研究提供了新的经验证据。第二, 挖掘要素细分市场在大数据发展作用于企业创新能力提升中的关键作用。研究发现, 为了更好地提高企业创新能力, 大数据发展需要和劳动力市场、资本市场及技术市场深度融合。这不仅深化了对大数据发展影响企业技术创新内在规律的理解和认识, 还为更好发挥大数据在提升企业创新能力中的作用提供了政策启示。第三, 考察在不同地区和不同地方政府数据开放水平下大数据发展对企业创新能力的影响差异, 以及大数据发展对不同规模和不同所有制企业创新能力的影响差异, 进一步丰富了大数据发展和企业创新能力的相关研究。

二、理论分析与研究假设

基于大数据在中国飞速发展的现实情况, 以及我国经济进入转型期亟须提高企业创新能力的现实需求, 通过对已有研究成果的系统梳理, 构建如下理论框架, 探讨分析大数据发展对企业创新能力的影响效应。

(一) 大数据发展对企业创新能力影响的直接效应

大数据发展可能对企业创新能力产生积极影响。一方面, 大数据发展可以丰富企业的信息资源供给, 扩大信息搜索空间, 提高信息处理效率^[13], 有利于企业低成本、高效率地收集和利用碎片化的信息资源^[6], 增强从现有信息和技术中提炼创新成果的能力, 进而提高企业的过程创新水平^[13]; 同时, 数据资源在研发环节的应用有助于提高企业的产品研发效率^[14]。另一方面, 企业在生产、运营和销售过程中产生大量的用户数据^[15]。大数据技术通过对这些数据的分析、处理

和研究, 帮助企业准确把握消费者偏好, 挖掘消费者的潜在需求^[16], 进而以消费者需求为导向提供更准确的定制产品和服务^[17], 增强企业研发方向和市场需求的匹配度, 降低研发过程中的不确定性风险^[18]和科技成果商业化失败的风险^[19], 提高企业创新产出回报率。同时, 大数据发展通过引领企业生产研发模式转型^[7], 为企业创新过程挖掘除因果关系界定之外的新发现, 防止根据部分信息进行因果判断导致的决策失误^[20]。此外, 大数据发展还能为研发人员带来更前沿的研究工具和研究方法^[21], 甚至催生对已有创新方法的根本性变革, 企业研发创新在利用大数据的同时也获得了新的方法, 直接优化了当前创新进程的本质^[19]。

基于此, 本文提出假设 1: 大数据发展可以提高企业创新能力。

(二) 大数据发展对企业创新能力影响的间接效应

进一步从阻碍企业创新能力提升的企业自身因素来看, 大数据发展还可以通过影响以下阻碍因素来提高企业创新能力, 主要体现在以下三个方面:

第一, 融资约束缓解效应。创新要素不足, 特别是研发资金不足会限制企业创新活动的开展。大数据发展通过将大数据技术引入到决策流程中: 一方面, 企业可以以较低成本掌握海量信息, 对这些数据信息进行充分处理和筛选, 有针对性地披露金融机构所需的信息, 更好满足信息使用者的需求^[22], 缓解和金融机构之间的信息不对称。另一方面, 金融机构也可以通过大数据仓库挖掘企业的海量数据资源, 特别是非结构化数据, 从而准确研判企业财务情况, 缓解和企业之间的信息不对称^[11], 有利于增加金融供给和提高金融资源配置效率。此外, 大数据发展将大数据技术和金融服务结合, 降低了金融服务的门槛^[23], 缓解了传统金融服务中的信贷配给歧视行为^[24], 扩大了金融服务的可得性, 为企业创新提供充足的资金保障。

第二, 人力资本升级效应。作为企业创新的关键投入要素, 高素质劳动力直接决定了企业创新产出的质量^[25]。大数据发展使得大数据技术和企业生产、运营、管理、销售等活动深度融合,

实现企业生产经营活动的数字化、智能化和自动化,直接提升了企业对大数据技术的应用程度^[13]。进一步地,企业对大数据相关专业高技能劳动力的需求也会增加,推动企业人力资本升级^[26]。高水平劳动力间可以通过学习、交流和共享知识信息,产生知识外溢效应^[27],增强企业创新能力;还可以通过对关键技术的学习消化吸收,提高技术和技能劳动力之间的匹配度^[28],优化创新资源配置,提高企业自主创新水平。

第三,管理效率提升效应。企业组织管理效率低下也会影响企业创新活动的开展,而大数据发展能够缓解企业在研发过程中面临的信息碎片化和信息不对称问题,基于数据的管理决策为企业创新管理实践带来新范式,提高企业的研发管理效率^[29]和研发决策准确率,为企业创新带来充足的动力源泉。同时,大数据发展能够降低企业内不同部门间的信息传递成本,提高信息传递效率,优化企业内部组织的管理模式和管理流程,降低管理成本,进而提高企业的生产和研发效率^[30]。

基于此,本文提出假设 2: 大数据发展通过融资约束缓解、人力资本升级和管理效率提升三条路径提高企业创新能力。

(三) 大数据发展、要素细分市场发展与企业创新能力

充分发挥大数据发展对企业技术创新的促进作用,必须重视数据要素和其它生产要素的结合,当数据和生产资料分离时,其只是可能的生产要素,要成为现实生产要素,数据必须和劳动、技术及管理相结合^[31]。同时,在我国多年的经济体制改革过程中,要素市场相对商品市场的市场化进程落后,存在市场扭曲、市场分割和要素流动障碍等问题^[32],所以,劳动力市场、资本市场和技术市场发展状况,可能会对大数据发展提升企业创新能力的效果产生影响。

第一,劳动力市场发展的影响。发达的劳动力市场能够促进科技人员的跨地区、跨行业和跨企业流动,而高水平科研人员是前沿技术知识扩散的重要载体,其在企业间的自由流动,能够推动先进技术知识在企业间的推广和应用^[33],增强大数据发展对企业技术创新的促进作用。同时,人才要素在不同企业间的优化配置,有利于强化

数据要素和高技能劳动力的结合水平,将人本身的价值转移到新产品中,进而激励企业积极开展技术创新活动^[34]。此外,发达的劳动力市场能够吸引和集聚本地区以外的大数据技术研发、应用和管理方面的高技能劳动力,不仅能够推动本地区大数据技术的发展和 innovation,还能进一步增强大数据发展对企业管理效率的提升效应。因此,可以推论,劳动力市场发展可以强化大数据发展对企业创新能力的提升作用。

第二,资本市场发展的影响。发达的资本市场不仅可以反映资本要素的真实价格,加速资本在企业间的自由流动和配置,还能提高大数据技术研发和推广对资金的吸引力,有助于构建高效、多元的大数据发展投融资体系,推动区域内大数据发展^[35],进一步降低本地区企业和金融机构之间的信息不对称,强化大数据发展对企业融资约束问题的缓解。因此,可以推论,资本市场发展可以加强大数据发展对企业创新能力的促进作用。

第三,技术市场发展的影响。技术市场是集聚各类创新主体和创新资源,为技术要素流通配置、技术交易和转移转化提供科技服务的中介平台^[36]。发达的技术市场可以为大数据相关技术的研发、交易、推广、商业化和产业化提供高质量的技术服务,提升区域内大数据发展水平。因此,可以推论,技术市场发展可以加速大数据相关技术的流通和配置效率,推动区域内大数据发展,强化大数据发展对企业创新能力的提升作用。

基于此,本文提出假设 3: 大数据发展对企业创新能力的促进作用,会受到劳动力市场、资本市场和技术市场等要素细分市场发展的影响;换句话说,大数据和劳动力市场、资本市场、技术市场等要素细分市场的融合发展可以更好推动企业创新能力提升。

三、研究设计

(一) 模型构建

参考以往研究文献,本文构建以下回归模型检验大数据发展对企业创新活动的影响:

$$Innovation_{i,t} = w + w_1 Bigdata_{m,t} + w_2 Controls + \delta_i + \theta_{i,j} + \varepsilon_{i,t} \quad (1)$$

模型(1)中,被解释变量 $Innovation_{i,t}$ 表示企业创新活动,分别从创新投入、创新产出和创新效

率三个维度衡量; 核心解释变量 $Bigdata_{m,t}$ 为企业 i 所在省份 m 在第 t 年的大数据发展; $Controls$ 为企业个体和省份层面的控制变量; δ_i 为企业个体固定效应, 考虑到地区固定效应会被企业个体固定效应吸收, 本文实际也对地区固定效应进行了控制^[37]; $\theta_{i,j}$ 代表的年份×行业固定效应控制了行业层面随时间变动的不可观测因素, 其中, 下标 t 表示年份, 下标 j 表示行业; $\varepsilon_{i,t}$ 表示随机误差项。

(二) 变量定义

1. 企业创新能力

为更加充分和全面地衡量大数据发展对企业创新能力的影响, 本文选择从创新投入、创新产出和创新效率三个维度衡量企业创新能力。参考施建军和栗晓云 (2021)^[38] 的研究, 创新投入 ($Inni$) 采用企业研发强度 (研发支出与营业收入的比值) 衡量; 创新产出 ($Inno$) 采用企业当年独立及联合专利申请数之和加 1 后取自然对数衡量; 创新效率 ($Inne$) 采用企业当年独立及联合专利申请数与连续三年研发支出除以企业期末总资产之间的比值加 1 后取自然对数 (\ln [专利申请数 / (连续三年研发支出 / 期末总资产) + 1])

衡量。

2. 大数据发展水平

自 2014 年大数据首次被写入政府工作报告以来, 国内众多科研院所、高等院校都曾发布过我国省域大数据发展指数, 但这些机构均未编制连续的大数据发展指数, 且不同机构发布的大数据发展指数存在指标选择、统计口径和评价方法上的差异, 不宜直接采用现有大数据发展指数开展实证研究。基于此, 本文以现有大数据发展指数的编制思想和逻辑为基础, 根据数据可得性, 编制了 2012—2021 年统一指标体系、口径和评价方法的省域大数据发展指数。

本文从制度发展、商业发展和民用发展三个维度出发, 构建省域大数据发展指数的一级指标, 并基于代表性、系统性、独立性、可操作性等原则选择了 11 个二级指标 (见表 1), 并使用熵值法确定各指标所占权重, 最后通过加权平均获得省域大数据发展 ($Bigdata$) 指数。此外, 本文还通过计算上述二级指标的算术平均值 ($Bigdata1$) 来度量省域大数据发展指数, 以便进行稳健性检验。

表 1 大数据发展指数指标体系

一级指标	二级指标	具体测算的指标	指标正负	指标含义
制度发展	政策力度	大数据相关政策发布次数	+	反映地方政府推进大数据发展的积极性
	试点创新	是否存在大数据综合试验区; 是=1, 否=0	+	反映地区大数据发展在国家战略中的地位
	数据开放平台	是否存在政府数据开放平台; 是=1, 否=0	+	反映地区数据开放的范围和多样性
商业发展	科技人才投入	电子及通信设备制造业从业人员平均数	+	反映大数据发展的人才支撑力度
	科技资金投入	规模以上工业企业 R&D 经费占 GDP 比重	+	反映大数据发展的资金支持力度
	商业融合	有电子商务交易活动企业比重	+	反映大数据发展的商用成熟度
	相关产业规模	软件和信息技术服务业收入占 GDP 比重	+	反映大数据相关产业发展情况
民用发展	大数据服务通达度	在线政府指数	+	反映政府利用大数据提供公共服务的能力
	智能终端普及率	移动电话普及率	+	反映公众信息采集与市场应用的广泛度
	人均移动互联网接入流量	每百人使用计算机数	+	反映网络基础设施建设情况
	通信消费能力	居民交通通信支出占总消费支出的比重	+	反映公众对信息消费的支出能力

注: 大数据相关政策发布次数、大数据综合试验区和政府数据开放平台三个数据来自各省份地方政府官方网站; 在线政府指数用省级政府网上政务服务能力得分衡量, 数据来自国家行政学院电子政务研究中心课题组每年发布的《中国电子政务发展报告》。

3. 控制变量

为减少遗漏变量对估计结果的影响, 本文在企业 and 省份层面分别选择了相应的控制变量, 其中, 企业层面的控制变量包括: 企业规模 ($asset$)、资产负债率 ($leve$)、总资产净利润率 (roa)、固定资产比率 (fix)、无形资产比率 ($intangible$)、营业收入增长率 ($operate$)、托宾 Q 值

($tobinq$)、前十大股东持股比例 ($top10$) 和独立董事比例 ($indep$), 省份层面的控制变量为人均国内生产总值 ($pgdp$)。本文涉及的所有变量和具体测算方法如表 2 所示。

(三) 样本选择和数据来源

基于数据可得性, 本文的研究样本选自 2012—2021 年在沪深 A 股上市的制造业企业。并

按照以下标准进行处理：剔除在 2012 年后上市的企业样本，剔除非正常交易的上市企业（ST、*ST、PT），剔除出现主要变量数据缺失、财务数据异常样本，最终得到 1 043 个企业样本，10 430 个观测值。企业专利申请数据来自中国研究数据服务平台（CNRDS），财务数据来自 CSMAR 数据库，省份层面的数据来自《中国统计年鉴》，要素细分市场发展指数的测算数据来自中国分省份市场化指数数据库，由于最新数据更新至 2019 年，所以 2020 年和 2021 年数据，采用 2012—2019 年各省份要素市场的市场化进程指数的年平均增长率推算得到。此外，为减轻异常值对回归结果的影响，对所有连续变量进行 1% 缩尾处理，以下报告的均是经过处理后的数据结果。所有变量的描述性统计结果如表 3 所示。

表 2 模型变量介绍

变量名称	变量符号	计算方法
创新投入	Inni	企业研发支出与营业收入的比值
创新产出	Inno	ln (企业专利申请数+1)
创新效率	Inne	ln [企业专利申请数 / (连续三年研发支出 / 期末总资产) + 1]
大数据发展	Bigdata	基于熵值法计算的省域大数据发展指数
企业规模	asset	ln (员工人数)
资产负债率	leve	总负债 / 总资产
总资产净利润率	roa	净利润 / 总资产
固定资产比率	fix	固定资产净额 / 总资产
无形资产比率	intangible	无形资产净额 / 总资产
营业收入增长率	operate	(当年营业收入 - 上年营业收入) / 上年营业收入
托宾 Q 值	tobinq	市值 / 资本重置成本
前十大股东持股比例	top10	前十大股东持股数 / 总股数
独立董事比例	indep	独立董事人数 / 董事会人数
人均国内生产总值	pgdp	ln (省份人均 GDP)

表 3 变量的描述性统计

变量	N	均值	标准差	最小值	中位数	最大值
Inni	10 430	4.19	3.36	0	3.59	19.60
Inno	10 430	3.58	1.60	0	3.64	7.63
Inne	10 430	6.31	2.07	0	6.49	10.70
Bigdata	10 430	0.32	0.21	0.05	0.28	0.77
asset	10 430	8.00	1.13	5.52	7.92	11.03
leve	10 430	0.41	0.18	0.05	0.40	0.80
roa	10 430	0.04	0.05	-0.16	0.04	0.20
fix	10 430	0.23	0.13	0.02	0.21	0.62
intangible	10 430	0.04	0.03	0.00	0.04	0.19
operate	10 430	0.14	0.30	-0.43	0.10	1.63
tobinq	10 430	2.04	1.17	0.86	1.66	7.23
top10	10 430	55.15	14.33	23.19	55.03	86.82
indep	10 430	37.45	5.40	33.33	33.33	57.14
pgdp	10 430	11.10	0.64	6.62	11.15	12.06

四、实证结果分析

(一) 基准回归结果

表 4 给出了本文的基准回归结果，在所有回归模型中均对企业个体固定效应和年份×行业固定效应进行了控制，并在企业层面聚类对标准误进行了修正。列（1）、列（2）和列（3）为未控制企业和省份层面控制变量的回归结果，列（4）、列（5）和列（6）为在列（1）、列（2）和列（3）基础上加入企业和省份层面控制变量的结果。列（1）~（3）的结果显示，在不控制企业和省份层面控制变量的情况下，回归结果非常显著，列（4）~（6）的结果显示，进一步在模型中加入企业和省份层面的控制变量后，大数据发展（Bigdata）的估计系数仍在 1% 的显著性水平下为正，但系数值略有降低，这说明在企业

表 4 全样本回归结果

变量及统计参数	Inni (1)	Inno (2)	Inne (3)	Inni (4)	Inno (5)	Inne (6)
Bigdata	1.890*** (6.080)	2.748*** (20.088)	2.722*** (14.664)	1.207*** (3.658)	1.555*** (11.604)	1.441*** (7.481)
asset				0.377*** (3.294)	0.792*** (16.320)	0.836*** (11.288)
leve				-2.276*** (-6.138)	-0.297* (-1.892)	-0.061 (-0.243)
roa				-6.892*** (-8.773)	0.446 (1.576)	-0.203 (-0.438)
fix				-0.434 (-0.974)	-1.451*** (-6.886)	-2.209*** (-6.323)
intangible				-0.461 (-0.197)	-0.528 (-0.699)	-1.519 (-1.274)
operate				-0.655*** (-8.534)	-0.011 (-0.330)	0.041 (0.719)
tobinq				-0.036 (-1.149)	-0.027** (-2.321)	-0.049*** (-2.611)
top10				-0.004 (-0.970)	-0.005** (-2.427)	-0.001 (-0.200)
indep				-0.008 (-1.233)	-0.003 (-0.937)	-0.003 (-0.507)
pgdp				0.225*** (6.331)	0.208*** (9.867)	0.205*** (7.171)
企业个体	Yes	Yes	Yes	Yes	Yes	Yes
时间×行业	Yes	Yes	Yes	Yes	Yes	Yes
N	10 430	10 430	10 430	10 430	10 430	10 430
调整 R ²	0.782	0.730	0.607	0.798	0.773	0.638

注：括号内为基于稳健标准误计算的 t 值；***、** 和 * 分别表示 1%、5% 和 10% 的显著性水平。下同。

所以, 本文的控制变量选择是有效的。列 (4)、列 (5) 和列 (6) 的回归结果表明, 在剔除了可能的混淆因素后, 大数据发展 (*Bigdata*) 显著提高了企业的创新投入 (*Inni*)、创新产出 (*Inno*) 和创新效率 (*Inne*), 系数值分别为 1.207、1.555 和 1.441。回归结果初步表明大数据发展能有效促进企业创新能力提升, 支持了本文的假设 1。

为进一步刻画大数据发展对企业创新的动态影响, 将大数据发展 (*Bigdata*) 滞后一到三期, 回归结果见表 5。可以看出, 大数据发展对企业创新投入 (*Inni*)、创新产出 (*Inno*) 和创新效率 (*Inne*) 均具有显著的衰减特征。其中, 大数据发展 (*Bigdata*) 对企业创新投入 (*Inni*) 的影响随时间推移呈现波动递减的特征, 滞后一期的大数

据发展对企业创新投入 (*Inni*) 在 1% 的显著性水平下为正, 系数为 1.106, 滞后二期的大数据发展 (*Bigdata*) 对企业创新投入 (*Inni*) 的影响不显著, 滞后三期的大数据发展 (*Bigdata*) 对企业创新投入 (*Inni*) 的影响在 5% 的显著性水平下为正, 系数为 0.854。大数据发展 (*Bigdata*) 对企业创新产出 (*Inno*) 的影响呈现随时间推移逐年递减的特征, 在滞后一到三期均通过了 1% 的显著性检验, 系数从 1.512 下降到 0.487。大数据发展 (*Bigdata*) 对企业创新效率 (*Inne*) 的影响也呈现随时间推移逐年递减的特征, 滞后一到三期的大数据发展 (*Bigdata*) 对企业创新效率 (*Inne*) 的影响分别在 1%、5% 和 10% 的显著性水平下为正, 系数分别为 1.325、0.468 和 0.391。

表 5 大数据发展对企业创新能力的动态影响

变量及统计参数	<i>Inni</i>			<i>Inno</i>			<i>Inne</i>		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
L1. <i>Bigdata</i>	1.106*** (3.154)			1.512*** (11.572)			1.325*** (6.887)		
L2. <i>Bigdata</i>		0.351 (0.915)			0.541*** (3.704)			0.468** (2.157)	
L3. <i>Bigdata</i>			0.854** (2.541)			0.487*** (3.323)			0.391* (1.859)
控制变量	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
企业个体	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
时间×行业	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	9 387	8 344	7 301	9 387	8 344	7 301	9 387	8 344	7 301
调整 R ²	0.810	0.827	0.843	0.784	0.805	0.816	0.656	0.673	0.690

(二) 内生性问题

本文的自变量为省份层面的大数据发展水平, 因变量为该省份内微观企业的创新能力, 两者间因反向因果关系产生内生性问题的可能性较低^[39], 但仍可能存在遗漏变量问题, 本文进一步使用工具变量法来缓解可能存在的内生性问题。采用各省份 1984 年每百万人邮局数量作为大数据发展的工具变量, 并参考 NUNN 和 QIAN (2014)^[40] 的做法, 通过引入时间序列变量来构造面板工具变量。采用上一年各省份大数据发展指数 (与时间有关) 和省级层面的 1984 年每百万人邮局数量的交乘项 (用符号 *PBD* 表示) 作为地区大数据发展的工具变量。表 6 列出了工具变量的回归结果。结果显示, 工具变量 *PBD* 通过了 K-P rk LM、C-D Wald F 和 K-P rk Wald F 检验, 且对企业创新投入 (*Inni*)、创新产出 (*Inno*) 和创新效率 (*Inne*) 的

估计系数均在 1% 的显著性水平下为正。因此, 在排除可能存在的内生性问题后, 假设 1 仍然成立。

表 6 工具变量回归结果

变量及统计参数	<i>Inni</i> (1)	<i>Inno</i> (2)	<i>Inne</i> (3)
<i>Bigdata</i>	1.823*** (4.118)	2.218*** (12.858)	2.023*** (7.989)
Controls	Yes	Yes	Yes
企业个体	Yes	Yes	Yes
时间×行业	Yes	Yes	Yes
K-P rk LM 统计量	448.563 [0.0000]	448.563 [0.0000]	448.563 [0.0000]
C-D Wald F 统计量	18 000.00 {16.38}	18 000.00 {16.38}	18 000.00 {16.38}
K-P rk Wald F 统计量	37 000.00 {16.38}	37 000.00 {16.38}	37 000.00 {16.38}
N	10 430	10 430	10 430
调整 R ²	0.087	0.239	0.118

注 [] 内数值为 *P* 值; C-D Wald F 统计量和 K-P rk Wald F 统计量在 10% 偏误下的临界值为 16.38, 在 { } 内表示。

(三) 稳健性检验

以上使用工具变量法缓解了可能存在的内生性问题后,为确保回归结果的可靠性,本文进一步进行了稳健性检验。第一,替换被解释变量。对衡量企业创新能力的三个被解释变量进行替换,具体来看,采用企业研发支出金额取自然对数(*Inni_*)表示创新投入,采用企业当年独立及联合发明专利申请数加1后取自然对数(*Inno_*)表示创新产出,采用企业当年独立及联合专利申请数与当年及前一年研发支出之和除以期末总资产的比值加1后取自然对数($\ln[\text{企业专利申请数}/(\text{连续两年研发支出之和}/\text{期末总资产})+1]$)

(*Inne_*)表示创新效率,回归结果见表7的列(1)、列(2)和列(3)。第二,替换解释变量。更换省级层面大数据发展水平的度量指标,采用算术平均值方法计算得到的各省份大数据发展水平(*Bigdata1*),回归结果见表7的列(4)、列(5)和列(6)。第三,使用非国有企业样本,对占全部样本总数近65%的非国有企业进行单独回归,回归结果见表7的列(7)、列(8)和列(9)。在替换模型的被解释变量、解释变量和回归样本后,大数据发展对企业创新投入(*Inni*)、创新产出(*Inno*)和创新效率(*Inne*)的影响系数仍显著为正,进一步说明本文回归结果的稳健性。

表7 稳健性检验结果

变量及统计参数	替换被解释变量			替换解释变量			非国有企业		
	<i>Inni_</i> (1)	<i>Inno_</i> (2)	<i>Inne_</i> (3)	<i>Inni</i> (4)	<i>Inno</i> (5)	<i>Inne</i> (6)	<i>Inni</i> (7)	<i>Inno</i> (8)	<i>Inne</i> (9)
<i>Bigdata</i>	1.841*** (7.498)	1.340*** (9.599)	1.609*** (7.685)				0.884* (1.761)	1.283*** (7.653)	1.363*** (5.299)
<i>Bigdata1</i>				1.696*** (3.759)	2.477*** (12.555)	2.429*** (8.054)			
<i>Controls</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
企业个体	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
时间×行业	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<i>N</i>	10 430	10 430	10 430	10 430	10 430	10 430	6 750	6 750	6 750
调整 <i>R</i> ²	0.809	0.779	0.609	0.809	0.779	0.609	0.786	0.749	0.608

(四) 机制分析

根据前文的理论分析,大数据发展可能通过缓解融资约束、提高人力资本水平和降低管理费用三种渠道提高企业创新能力,所以,本文通过建立如下回归模型对上述三种潜在影响机制进行检验:

$$m_{m,t} = \sigma + \rho Bigdata_{m,t} + \varphi Controls + \delta_i + \theta_{i,j} + v_{i,t} \quad (2)$$

模型(2)中, $m_{m,t}$ 为机制变量,包括融资约束SA指数(*SA*)、人力资本(*HC*)和管理效率(*ME*);*Bigdata*表示大数据发展水平,同时采用算术平均值方法度量的大数据发展水平(*Bigdata1*)进行稳健性检验。其余变量与模型(1)相同。

1. 缓解企业融资约束

充足的资金支持是企业开展创新活动的重要因素。参考宋敏等(2021)^[41]的研究,采用SA指数衡量企业融资约束(*SA*),回归结果见表8列(1)和列(2)。列(1)变量*Bigdata*的系数

显著为负,系数值为-0.500,表明大数据发展有助于缓解企业的融资约束,为假设2提供了经验证据。列(2)中*Bigdata1*的系数也显著为负,进一步说明回归结果的稳健性。

2. 提高企业人力资本水平

企业创新活动离不开高素质的劳动力支持。借鉴任英华等(2023)^[16]的研究,用企业硕士及以上学历员工占总员工比例衡量人力资本水平(*HC*),回归结果见表8列(3)和列(4)。列(3)中*Bigdata*的系数显著为正,系数值为2.111,说明大数据发展优化了人力资本结构,显著提升了企业的人力资本水平,为假设2提供了经验证据。列(4)中*Bigdata1*的系数也显著为正,证实了回归结果的稳健性。

3. 降低企业管理费用

高效率的组织管理决策、管理流程和管理模式对企业创新非常重要。借鉴戴魁早等(2023)^[35]的研究,采用企业管理费用率衡量企业的管理效率(*ME*),回归结果见表8列(5)和列(6)。列

(5) 中 *Bigdata* 的系数显著为负, 系数值为 -0.051, 说明大数据发展显著降低了企业的管理成本, 即通过提高管理效率这一机制促进了企业技术创新水平提升, 为假设 2 提供了经验证据。同时, 列 (6) 的回归结果也验证了大数据发展对企业管理效率提升的促进作用。

表 8 中介机制检验

变量及统计参数	SA		HC		ME	
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Bigdata</i>	-0.500*** (-33.149)		2.111*** (5.038)		-0.051*** (-5.105)	
<i>Bigdata</i> 1		-0.783*** (-34.285)		3.269*** (5.529)		-0.062*** (-4.614)
<i>Controls</i>	Yes	Yes	Yes	Yes	Yes	Yes
企业个体	Yes	Yes	Yes	Yes	Yes	Yes
时间×行业	Yes	Yes	Yes	Yes	Yes	Yes
<i>N</i>	10 430	10 430	10 430	10 430	10 430	10 430
调整 <i>R</i> ²	0.871	0.875	0.877	0.877	0.309	0.308

(五) 进一步分析

1. 要素细分市场发展的影响

为进一步验证大数据发展是否会受到劳动力、资本和技术要素市场发展的影响, 采用学术界通用做法, 在模型 (1) 基础上加入大数据发展与要素细分市场发展指数的交乘项, 回归模型如下:

$$Innovation_{i,t} = \beta + \beta_1 Bigdata_{m,t} + \beta_2 Bigdata_{m,t} \times X_{i,t} + \beta_3 Controls + \delta_i + \theta_{i,j} + \varepsilon_{i,t} \quad (3)$$

模型 (3) 中, $X_{i,t}$ 表示要素市场发展指数, 包括劳动力市场发展指数 (*lmd*)、资本市场发展指数 (*cmd*) 和技术市场发展指数 (*tmd*)。 $Bigdata \times X$ 表示大数据发展和要素市场发展指数的交乘项, β_2 为交乘项的系数, 其余符号含义和模型 (1) 一致。若交乘项的系数 β_2 显著为正, 说明要素细

分市场和大数据发展的联合作用显著推动了企业创新能力提升。

表 9 列出了模型 (3) 的回归结果, 从列 (1) ~ (3) 的回归结果可以看出, $Bigdata \times lmd$ 的系数在 1% 的显著性水平下都为正, 说明劳动力市场发展增强了大数据发展对企业创新投入 (*Inni*)、创新产出 (*Inno*) 和创新效率 (*Inne*) 的正向影响。从列 (4) ~ (6) 的回归结果可以看出, $Bigdata \times cmd$ 的系数显著为正, 说明资本市场发展强化了大数据发展对企业创新投入 (*Inni*)、创新产出 (*Inno*) 和创新效率 (*Inne*) 的积极影响。同样地, 从列 (7) ~ (9) 的回归结果可以看出, 技术市场发展也显著推动了大数据发展对企业创新能力的提升作用。假设 3 得到验证。

表 9 要素细分市场发展的影响估计结果

变量及统计参数	劳动力市场发展指数 (<i>lmd</i>)			资本市场发展指数 (<i>cmd</i>)			技术市场发展指数 (<i>tmd</i>)		
	<i>Inni</i> (1)	<i>Inno</i> (2)	<i>Inne</i> (3)	<i>Inni</i> (4)	<i>Inno</i> (5)	<i>Inne</i> (6)	<i>Inni</i> (7)	<i>Inno</i> (8)	<i>Inne</i> (9)
<i>Bigdata</i>	0.903** (2.580)	1.017*** (5.310)	1.406*** (4.638)	1.010** (2.335)	0.964** (2.549)	1.118** (2.223)	0.903*** (3.477)	0.910*** (5.365)	0.967*** (3.906)
<i>Bigdata</i> × <i>lmd</i>	0.072*** (3.455)	0.041*** (3.690)	0.073*** (3.910)						
<i>Bigdata</i> × <i>cmd</i>				0.058** (2.484)	0.038** (2.342)	0.037* (1.691)			
<i>Bigdata</i> × <i>tmd</i>							0.072*** (5.486)	0.049*** (4.999)	0.036*** (2.609)
<i>Controls</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
企业个体	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
时间×行业	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<i>N</i>	10 430	10 430	10 430	10 430	10 430	10 430	6 750	6 750	6 750
调整 <i>R</i> ²	0.810	0.774	0.616	0.793	0.757	0.621	0.810	0.774	0.639

2. 大数据发展对企业创新能力的异质性分析

鉴于各省份资源禀赋、地方政府数据开放水平、企业规模和所有制差异,使得大数据发展可能存在异质性影响。本文进一步从地区宏观环境和企业微观特征的异质性分析大数据发展与企业创新能力之间的关系,挖掘大数据发展对企业创新能力的影响是否因企业所处地区、地方政府数据开放水平、企业所有制和企业规模的不同而存在明显差异,以加深对大数据发展影响企业创新能力基本规律的认识和理解。

(1) 企业所处地区的异质性。基于企业所处不同地区,大数据发展对企业创新能力可能存在差异。一方面,我国东部沿海地区开放程度高,制造业发展基础好,技术创新能力更突出;另一方面,我国大数据发展也存在较为明显的空间差

异性,总体呈现东部强于中西部的特征^[42]。因此,大数据发展对东部地区企业创新能力的提升作用可能更明显。为对上述逻辑进行验证,根据企业所处地区,将样本划分为东部和中西部后进行分组回归,结果见表10列(1)~(6)。可以看出,在东部地区分组中,大数据发展(*Bigdata*)对企业创新投入(*Inni*)、创新产出(*Inno*)和创新效率(*Inne*)的估计系数均在1%的显著性水平下为正,说明大数据发展显著促进了东部地区企业创新能力提升;在中西部企业分组中,大数据发展(*Bigdata*)的估计系数均不显著,说明大数据发展对中西部地区的企业创新能力并未产生显著促进作用。回归结果和推论一致:大数据发展对企业创新能力的促进作用在东部地区企业中更加明显。

表10 基于企业所处地区的异质性分析

变量及统计参数	东部	中西部	东部	中西部	东部	中西部
	<i>Inni</i> (1)	<i>Inni</i> (2)	<i>Inno</i> (3)	<i>Inno</i> (4)	<i>Inne</i> (5)	<i>Inne</i> (6)
<i>Bigdata</i>	1.007*** (2.590)	0.958 (1.520)	1.424*** (10.213)	0.613 (1.358)	1.370*** (6.610)	0.145 (0.257)
<i>Controls</i>	Yes	Yes	Yes	Yes	Yes	Yes
企业个体	Yes	Yes	Yes	Yes	Yes	Yes
时间×行业	Yes	Yes	Yes	Yes	Yes	Yes
<i>N</i>	7 091	3 339	7 091	3 339	7 091	3 339
调整 <i>R</i> ²	0.783	0.831	0.787	0.763	0.645	0.636

(2) 地方政府数据开放水平的异质性。地方政府数据开放水平不同的地区,大数据发展对企业创新能力的影响可能存在差异。地方政府拥有的数据资源能提高所在地区的信息整合能力和数据资源配置能力,进而促进地区创新水平提升。地方政府数据开放水平越高,企业利用大数据技术得到的优质数据资源越多,对创新资源的优化配置能力越强,该地区企业能够更加高效、快捷地获取开展技术创新所需的信息资源和要素资源,有利于提高自身的技术创新水平。因此,大数据发展对企业创新能力的提升作用在地方政府数据开放水平高的地区更为显著。为对上述推论进行验证,基于复旦大学数字与移动治理实验室和国家信息中心数字中国研究院联合发布的《中国地方政府数据开放报告》(2022年)中的“数林指数”划分地区数据开放水平高低(前25%为高,后75%为低),将样本企业按照所在地划分为两

组并分别进行回归。回归结果见表11列(1)~(6)。结果显示,在地方政府数据开放水平高的地区,大数据发展(*Bigdata*)对企业创新投入(*Inni*)、创新产出(*Inno*)和创新效率(*Inne*)的估计系数均在1%的显著性水平下为正。在地方政府数据开放水平低的地区,大数据发展对企业创新投入(*Inni*)没有显著影响,对创新产出(*Inno*)和创新效率(*Inne*)的影响分别在1%和5%的显著性水平下为正,但系数明显低于地方政府数据开放水平高的地区。回归结果与推论一致:大数据发展对企业创新能力的提升作用在地方政府数据开放水平高的地区更加突出。

(3) 企业所有制的异质性。大数据发展对不同所有制企业的影响可能存在差异性。和非国有企业相比,国有企业拥有更强的资源优势、知识产权储备和社会地位等,受到的资源支持和政策支持力度更大,更有能力落实国家和地方的大数

据发展相关政策。同时, 创新资源如技术、资本、高素质劳动力等更容易流入到国有企业^[43]。因此, 大数据发展对国有企业创新能力的促进作用可能更为明显。为对上述推论进行验证, 从企业所有制角度出发, 将样本企业划分为国有企业和非国有企业后进行分组回归, 结果见表 12 列 (1) ~ (6)。可以看出, 无论是国有企业还是非

国有企业, 大数据发展 (*Bigdata*) 均显著提升了企业的创新投入 (*Inni*)、创新产出 (*Inno*) 和创新效率 (*Inne*), 且在创新投入和创新产出方面对国有企业的促进作用更强, 在创新效率方面对非国有企业的提升作用更大。回归结果与推论基本一致: 大数据发展对企业创新能力的提升作用在国有企业样本中更为明显。

表 11 基于地方政府数据开放水平的异质性分析

变量及统计参数	数据开放水平高	数据开放水平低	数据开放水平高	数据开放水平低	数据开放水平高	数据开放水平低
	<i>Inni</i> (1)	<i>Inni</i> (2)	<i>Inno</i> (3)	<i>Inno</i> (4)	<i>Inne</i> (5)	<i>Inne</i> (6)
<i>Bigdata</i>	1.355 *** (2.669)	-0.029 (-0.064)	1.413 *** (8.009)	0.665 *** (2.746)	1.399 *** (5.248)	0.640 ** (1.991)
<i>Controls</i>	Yes	Yes	Yes	Yes	Yes	Yes
企业个体	Yes	Yes	Yes	Yes	Yes	Yes
时间×行业	Yes	Yes	Yes	Yes	Yes	Yes
<i>N</i>	5 479	4 951	5 479	4 951	5 479	4 951
调整 <i>R</i> ²	0.778	0.818	0.775	0.787	0.615	0.665

表 12 基于企业产权性质的异质性分析

变量及统计参数	国有企业	非国有企业	国有企业	非国有企业	国有企业	非国有企业
	<i>Inni</i> (1)	<i>Inni</i> (2)	<i>Inno</i> (3)	<i>Inno</i> (4)	<i>Inne</i> (5)	<i>Inne</i> (6)
<i>Bigdata</i>	1.520 *** (4.052)	0.884 * (1.761)	1.632 *** (7.597)	1.356 *** (8.137)	1.243 *** (4.393)	1.363 *** (5.299)
<i>Controls</i>	Yes	Yes	Yes	Yes	Yes	Yes
企业个体	Yes	Yes	Yes	Yes	Yes	Yes
时间×行业	Yes	Yes	Yes	Yes	Yes	Yes
<i>N</i>	3 680	6 750	3 680	6 590	3 680	6 750
调整 <i>R</i> ²	0.836	0.786	0.815	0.760	0.686	0.608

(4) 企业规模的异质性。大数据发展对不同规模企业的影响可能存在差异。和中小型企业相比, 大型企业拥有更强的人才、技术、资金等资源优势, 获取创新资源能力强, 能有效整合技术创新活动中所需的各种创新资源和关键信息, 提高研发效率, 进而促进技术创新水平提升。因此,

大数据发展对大规模企业创新能力提升的促进作用可能更加显著。为对上述逻辑进行验证, 根据员工人数的中位数将样本企业划分为大型企业和中小型企业进行分组回归, 结果见表 13 列 (1) ~ (6)。可以看出, 在大型企业分组中, 大数据发展 (*Bigdata*) 显著提高了企业的创新投入 (*Inni*)、

表 13 基于企业规模的异质性分析

变量及统计参数	大型企业	中小型企业	大型企业	中小型企业	大型企业	中小型企业
	<i>Inni</i> (1)	<i>Inni</i> (2)	<i>Inno</i> (3)	<i>Inno</i> (4)	<i>Inne</i> (5)	<i>Inne</i> (6)
<i>Bigdata</i>	2.132 *** (7.027)	0.505 (0.839)	1.924 *** (10.312)	1.052 *** (5.883)	1.443 *** (6.091)	1.295 *** (4.488)
<i>Controls</i>	Yes	Yes	Yes	Yes	Yes	Yes
企业个体	Yes	Yes	Yes	Yes	Yes	Yes
时间×行业	Yes	Yes	Yes	Yes	Yes	Yes
<i>N</i>	5 214	5 216	5 214	5 216	5 214	5 216
调整 <i>R</i> ²	0.864	0.781	0.802	0.651	0.679	0.532

创新产出 (*Inno*) 和创新效率 (*Inne*); 在中小型企业分组中, 大数据发展 (*Bigdata*) 对企业创新产出 (*Inno*) 和创新效率 (*Inne*) 的影响均在 1% 的显著性水平下为正, 但系数都小于大型企业, 对企业创新投入 (*Inni*) 的影响则不显著。回归结果与推论基本一致: 大数据发展对企业创新能力的促进作用在大型企业中更加明显。

五、结论与建议

(一) 研究结论

提高企业创新能力不仅是加快建设科技强国和实现高水平科技自立自强的重要手段, 也是促进中国经济高质量发展的重要举措, 对我国经济社会长期可持续发展至关重要。鉴于此, 本文着重探讨了大数据发展对企业创新能力的影响, 在理论归纳大数据发展影响企业创新能力的基础上, 结合 2012—2021 年 A 股上市企业数据, 运用计量模型法对理论预期进行了实证分析。进一步地, 探究了劳动力市场、资本市场、技术市场在大数据发展推动制造业企业创新能力提升中的重要作用, 并基于企业所处地区、地方政府数据开放水平、企业所有制和企业规模的不同验证了大数据发展对企业创新能力影响的异质性。得到如下结论: 第一, 大数据发展可以显著提高企业创新能力, 包括创新投入、创新产出和创新效率。第二, 在影响机制上, 大数据发展可以通过缓解企业融资约束、优化人力资本水平和提高企业管理效率三种渠道推动企业创新能力提升。进一步研究还发现: 劳动力市场、资本市场和技术市场等要素细分市场 and 大数据发展的联合作用显著促进了企业创新能力提升。同时, 大数据发展对企业创新能力的促进作用在东部地区、地方政府数据开放水平高的地区、国有企业和大型企业中的作用效果更加明显。因此, 以大数据为支点撬动企业生产方式和创新模式的变革势在必行, 政策层面应为企业创新提供充裕的数字化转型动能。

(二) 政策建议

第一, 夯实大数据发展基础, 强化大数据政策引领企业技术创新的制度设计, 充分释放大数据发展的红利效应。一方面, 要优化数据资源管理和大数据标准体系建设。企业要及时出台数据要素管理规章制度, 让数据要素在企业研发前、

中、后各环节得到充分应用, 优化创新资源配置效率。政府要完善数据要素市场体系、大数据标准体系和大数据安全体系的体制机制, 打破数据要素流通壁垒, 构建政府数据和企业数据间的共享机制, 让大数据更好地服务于企业的创新活动。另一方面, 要重视完善大数据相关基础设施建设。结合大数据服务企业创新活动的各环节, 加快数据采集、存储设备、大数据中心等新型基础设施建设。此外, 政府要鼓励和支持企业、高校和科研院所开展大数据相关的基础研究, 加快大数据相关的核心技术突破和产业化应用。

第二, 构建支持企业技术创新的长效机制, 重视大数据在融资约束缓解、人力资本升级和管理效率提升上的重要作用, 进一步疏通大数据发展提高企业创新能力的传导机制, 利用制度保障推动企业创新发展。首先, 鼓励金融机构加快数字化转型, 高效利用本地区大数据基础设施, 建立能够充分利用大数据资源和技术的信贷评估模型, 增强信息整合、甄别、评估和利用效率, 提高风险管理水平, 进而增强企业资金的可获得性, 增加企业研发投入。其次, 从政策层面持续推进高等教育体系改革, 鼓励大中专院校开设大数据相关专业, 提倡通过校企合作等方式对接相关产业的实际需求, 培养高素质高技能人才。最后, 积极布局大数据平台和机构, 扩大服务数量和规模, 并不断完善大数据平台的服务机制, 增强大数据服务企业的效率, 提高企业管理效率和资源利用效率, 通过流程优化实现企业组织结构的扁平化。

第三, 在要素细分市场发展水平较低地区, 进一步加强大数据发展和劳动力市场、资本市场及技术市场的融合发展。具体来说, 政府部门应加快推进要素市场一体化建设, 促进数据要素和其它生产要素市场的融合与组合优化, 不断完善大数据相关人才、资金、技术自由流动的市场体系, 如在高校设置大数据相关专业和学科, 加强高校和企业的人才交流, 鼓励各创新主体积极参与大数据相关技术交易等。同时, 进一步强化大数据技术在各要素市场服务活动中的应用, 推进业务流程再造, 加快传统要素数据化, 提高各要素细分市场的服务水平, 逐步实现全国范围内要素市场数据的共享、使用和服务。

第四,进一步细化大数据发展政策,强化大数据发展对不同地区、不同类型企业创新能力提升的针对性。对中西部地区,要特别重视加快推进数据要素市场发展,不断完善数据基础设施建设,加强资金和人才投入,同时充分发挥自身的劳动密集特征,积极开展数据清洗、标注等传统数据服务。对地方政府数据开放水平低的地区,

应进一步扩大向社会提供的数据规模,提高数据质量,并加快完善数据确权工作,特别是在数据开放和数据安全上的立法与市场监管。对非国有企业和小规模企业,要在数字技术人才培养、新型基础设施推广和应用、市场竞争、环境营造等方面给予更多的政策支持,降低企业应用大数据的要素成本和交易成本,激发企业创新积极性。

[注 释]

① 数据来源于《2022 中国大数据产业发展白皮书》。

[参考文献]

- [1] 陈彦斌,刘哲希. 中国企业创新能力不足的核心原因与解决思路 [J]. 学习与探索, 2017 (10): 115-124.
- [2] 白俊红,刘宇英. 金融市场化与企业技术创新: 机制与证据 [J]. 经济管理, 2021, 43 (4): 39-54.
- [3] 夏清华,黄剑. 市场竞争、政府资源配置方式与企业创新投入——中国高新技术企业的证据 [J]. 经济管理, 2019, 41 (8): 5-20.
- [4] 胡山,余泳泽. 数字经济与企业创新: 突破性创新还是渐进性创新? [J]. 财经问题研究, 2022 (1): 42-51.
- [5] 郑帅,王海军. 数字化转型何以影响枢纽企业创新绩效? ——基于模块化视角的实证研究 [J]. 科研管理, 2022, 43 (11): 73-82.
- [6] 许宪春,任雪,常子豪. 大数据与绿色发展 [J]. 中国工业经济, 2019 (4): 5-22.
- [7] 杨俊,李小明,黄守军. 大数据、技术进步与经济增长——大数据作为生产要素的一个内生增长理论 [J]. 经济研究, 2022, 57 (4): 103-119.
- [8] WU L, HITT L, LOU B. Data analytics, innovation, and firm productivity [J]. Management science, 2020, 66 (5): 2017-2039.
- [9] 金芳,齐志豪,梁益琳. 大数据、金融集聚与绿色技术创新 [J]. 经济与管理评论, 2021, 37 (4): 97-112.
- [10] 孙洁,李杰. 大数据应用、融资约束和企业创新效率 [J]. 证券市场导报, 2022 (11): 13-23.
- [11] 艾永芳,孔涛. 区域大数据发展能促进企业绿色创新吗? [J]. 中南财经政法大学学报, 2021 (6): 116-126.
- [12] 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.
- [13] 张叶青,陆瑶,李乐芸. 大数据应用对中国企业市场价值的影响——来自中国上市公司年报文本分析的证据 [J]. 经济研究, 2021, 56 (12): 42-59.
- [14] 李海舰,赵丽. 数据成为生产要素: 特征、机制与价值形态演进 [J]. 上海经济研究, 2021 (8): 48-59.
- [15] 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.
- [16] 任英华,刘宇钊,胡宗义,等. 大数据发展、知识产权保护对企业绿色技术创新的影响 [J]. 中国人口·资源与环境, 2023, 33 (7): 157-167.
- [17] FARBOODI M, MIHET R, PHILIPPON T, et al. Big data and firm dynamics [J]. AEA papers and proceedings, 2019, 109: 38-42.
- [18] 张益豪,郭晓辉. 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 [J]. 产业经济研究, 2023 (2): 69-82.
- [19] COCKBURN I M, HENDERSON R, STERN S. The impact of Artificial intelligence on innovation [Z]. NBER Chapters, 2018.
- [20] 何大安. 互联网应用扩张与微观经济学基础——基于未来“数据与数据对话”的理论解说 [J]. 经济研究, 2018, 53 (8): 177-192.
- [21] EINAV L, LEVIN J. Economics in the age of big data [J]. Science, 2014, 346 (6210): 1-6.
- [22] ELLIOTT W B, GRANT S M, HOBSON J L. Trader participation in disclosure: implications of interactions with management [J]. Contemporary accounting research, 2018, 37 (1): 68-100.
- [23] 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.
- [24] 张金林,董小凡,李健. 数字普惠金融能否推进共同富裕? ——基于微观家庭数据的经验研究 [J]. 财经研究, 2022, 48 (7): 4-17.
- [25] 沈国兵,袁征宇. 企业互联网化对中国企业创新及出口的影响 [J]. 经济研究, 2020, 55 (1): 33-48.
- [26] 孙早,侯玉琳. 工业智能化如何重塑劳动力就业结构 [J]. 中国工业经济, 2019 (5): 61-79.
- [27] 刘维刚,倪红福. 制造业投入服务化与企业技术进步: 效应及作用机制 [J]. 财贸经济, 2018, 39 (8): 126-140.
- [28] 刘涛,油永华. 供给侧结构性改革: 基于技术资本与技能劳动关系的视角 [M]. 北京: 经济科学出版社, 2021.
- [29] 谢雄标,吴越,严良. 数字化背景下企业绿色发展路径及政策建议 [J]. 生态经济, 2015, 31 (11): 88-91.

- [30] BRYNJOLFSSON E, HITT L M, HELLEN K H. Strength in numbers: how does data-driven decision making affect firm performance? [Z]. SSRN electronic journal, 2011.
- [31] 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.
- [32] 赵奇伟, 熊性美. 中国三大市场分割程度的比较分析: 时间走势与区域差异 [J]. 世界经济, 2009 (6): 41-53.
- [33] 戴魁早. 技术市场发展对出口技术复杂度的影响及其作用机制 [J]. 中国工业经济, 2018 (7): 117-135.
- [34] 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.
- [35] 戴魁早, 王思曼, 黄姿. 数据要素市场发展与生产率提升 [J]. 经济管理, 2023, 45 (6): 22-43.
- [36] 张汝飞, 刘超, 赵彦云. 技术市场与科技创新互动效应研究——以北京技术市场为例 [J]. 数学的实践与认识, 2016, 46 (16): 43-51.
- [37] 李春涛, 闫续文, 宋敏, 等. 金融科技与企业创新——新三板上市公司的证据 [J]. 中国工业经济, 2020 (1): 81-98.
- [38] 施建军, 栗晓云. 政府补助与企业创新能力: 一个新的实证发现 [J]. 经济管理, 2021, 43 (3): 113-128.
- [39] 张杰, 周晓艳, 李勇. 要素市场扭曲抑制了中国企业 R&D? [J]. 经济研究, 2011, 46 (8): 78-91.
- [40] NUNN N, QIAN N. US food aid and civil conflict [J]. American economic review, 2014, 104 (6): 1630-1666.
- [41] 宋敏, 周鹏, 司海涛. 金融科技与企业全要素生产率——“赋能”和信贷配给的视角 [J]. 中国工业经济, 2021 (4): 138-155.
- [42] 史丹, 孙光林. 大数据发展对制造业企业全要素生产率的影响机理研究 [J]. 财贸经济, 2022, 43 (9): 85-100.
- [43] 毛其淋, 许家云. 市场化转型、就业动态与中国地区生产率增长 [J]. 管理世界, 2015 (10): 7-23.

Big Data Development and Enhancement of Enterprises Innovation Ability

Chen Rui¹, Wang Hongwei^{2,3}

(1. Institute of Quantitative & Technological Economics, Chinese Academy of Social Sciences, Beijing 100732, China;

2. Project Evaluation and Strategic Planning Research Consulting Center, Chinese Academy of Social Sciences, Beijing 100732, China;

3. University of Chinese Academy of Social Sciences, Beijing 102488, China)

Abstract: Big data development boosts the speed and efficiency of information exchange and knowledge sharing. It has become the major driving force to construct China into a powerhouse of science and technology, and help to realize the self-reliant high-level science and technology. Based on the data of A-share listed manufacturing enterprises from 2012 to 2021, the article explores the impact and mediating mechanism of big data development on enterprises innovation ability. The finding is that big data development significantly promotes enterprises innovation ability, and this finding still holds water after controlling the endogeneity problem and robustness test. The mechanism analysis shows that big data development improves enterprises innovation ability through three ways: alleviating enterprise financing constraints, optimizing human capital structure and improving management efficiency. It is further found that the development of factor market segments such as labor market, capital market and technology market enhances the effect of big data development on enterprises innovation, and the effect of big data development on enterprises innovation ability is more obvious in the eastern region, the region with high level of data openness of the local government, the state-owned enterprises and large-scale enterprises. By verifying the important role of big data development in the enhancement of enterprises innovation ability, it provides a useful reference for big data development and enterprises core technology breakthrough.

Key words: big data development; enterprises innovation ability; factor market; manufacturing industry

(责任编辑: 李 萌)