

DOI:10.3969/j.issn.1674-8131.2025.05.004

大数据应用对企业突破性技术创新的影响

——知识资源优化路径与知识产权保护的调节作用

王 军,周效宇

(中国石油大学(华东)经济管理学院,山东 青岛 266580)

摘 要:突破性技术创新需要组织、配置、利用多领域的知识资源,大数据应用可以从知识多元化和吸收能力等方面优化企业知识资源,进而促进企业突破性技术创新。采用沪深A股上市公司2009—2023年的数据分析发现:大数据应用显著提升了企业突破性技术创新水平,并通过推动知识多元化和提高知识吸收能力两条路径促进企业突破性技术创新;地区知识产权保护强度提高弱化了大数据应用对企业突破性技术创新的促进作用,该调节效应仅在竞争性行业中显著;大数据应用对突破性技术创新的促进作用,在竞争性行业比管制性行业更显著,在国有企业、知识多元化水平较高企业、非高科技行业企业、东部和市场化水平较高地区企业中显著,在非国有企业、高科技行业企业、中西部和市场化水平较低地区企业中不显著,在知识多元化水平较低企业中则表现为负向影响。因此,应加快推进大数据应用,着力提升企业知识多元化水平和知识吸收能力,构建兼顾知识流动与创新激励的动态知识产权保护体系,有效促进企业突破性技术创新。

关键词:大数据应用;突破性技术创新;知识资源;知识多元化;知识吸收能力;知识产权保护强度

中图分类号:F273.1;F124.3 **文献标志码:**A **文章编号:**1674-8131(2025)05-0045-16

引用格式:王军,周效宇.大数据应用对企业突破性技术创新的影响——知识资源优化路径与知识产权保护的调节作用[J].西部论坛,2025,35(5):45-60.

WANG Jun, ZHOU Xiao-yu. Impact of big data applications on enterprises' breakthrough technological innovation: The path of optimizing knowledge resources and the regulatory role of intellectual property protection[J]. West Forum, 2025, 35(5): 45-60.

* 收稿日期:2025-04-14;修回日期:2025-07-06

基金项目:国家自然科学基金面上项目(72473154)

作者简介:王军(1969),男,山东单县人;教授,博士生导师,主要从事产业经济与区域发展研究。周效宇(2002),通信作者,女,山东威海人;硕士研究生,主要从事产业经济与区域发展研究。

一、引言

我国经济已由高速增长阶段转向高质量发展阶段,而技术创新是推动经济高质量发展的核心引擎。尤其是突破性技术创新,不仅能解决经济发展中的重大难题,提高生产效率,推动产业变革,还能提升全球竞争力和经济韧性,促进经济系统的健康持续发展。企业是经济运行的重要微观主体,也是技术创新的关键主体。然而,面对激烈的市场竞争和复杂多变的不确定性,企业可能为规避风险而倾向于进行潜在风险较小的渐进性技术创新(池毛毛等,2022)^[1],导致具有高投入、高风险、动态非连续性等特性的突破性技术创新进展缓慢。因此,如何激励和推动企业积极开展突破性技术创新成为社会各界关注的热点。随着新一轮科技革命的发展,数字技术的应用推动数字经济快速发展,为经济发展提供了新动能。与此同时,数据成为新的生产要素,大数据的积累和应用极大提升了经济发展质量和效率。基于此,大数据的应用能否以及怎样促进企业进行突破性技术创新是值得深入探究的重要课题。

关于大数据(技术)应用对企业技术创新的影响,除了考察地区大数据发展对企业技术创新的影响(艾永芳等,2021;戴艳娟等,2023;张志新等,2024;陈丹等,2024)^[2-5]外,现有文献主要聚焦于企业应用大数据对其技术创新的影响。比如,许芳等(2020)^[6]研究发现,大数据应用能力提高可以通过促进供应链协同对企业创新绩效产生显著的正向影响;陶小龙等(2021)^[7]分析认为,大数据应用与企业开放式创新存在紧密的双向赋能关系;孙洁和李杰等(2022)^[8]研究表明,大数据应用可以通过缓解融资约束提升企业创新效率;胡蓉宁(2024)^[9]分析发现,大数据应用能通过提高营销能力和降低成本提高企业创新质量;孙献贞(2024)^[10]研究认为,大数据技术应用通过优化创新要素配置和强化知识共享显著提升了制造业企业技术创新水平;钱秋兰等(2025)^[11]研究表明,大数据应用通过降低信息不对称和降本增效显著提升了企业创新水平和创新质量。然而,鲜有文献探讨大数据应用对企业突破性技术创新的影响,相关经验证据更为缺乏。有鉴于此,本文在已有研究的基础上,探讨大数据应用对企业突破性技术创新的影响及其机制,并采用2009—2023年沪深A股上市公司的数据进行实证检验。

本文的边际贡献主要包括:一是在研究视角上,从突破性技术创新维度拓展了大数据应用的经济效应研究。突破性技术创新比一般技术创新更具前瞻性、颠覆性和重塑性,是推动经济社会系统性变革的关键动力;而大数据作为新型生产要素,其应用是对传统发展模式下生产要素组合的重构。因此,大数据应用与突破性技术创新之间具有内在统一性。探究大数据应用对企业突破性技术创新的影响有助于深入理解数字技术驱动经济高质量发展的机理,进而寻求数字经济时代促进企业突破性技术创新的有效路径。二是在理论机制上,基于资源编排理论,从知识资源角度分析了知识多元化和知识吸收能力在大数据应用促进企业突破性技术创新中的中介作用以及知识产权保护的调节作用,为相关研究提供了新的视角和思路借鉴。三是进一步从企业(产权性质和知识多元化)、行业(竞争属性和技术属性)和区域(地理区位和市场化水平)3个层面考察了大数据应用影响企业突破性技术创新的异质性,为各类企业通过大数据应用加快突破性技术创新提供了经验借鉴和政策启示。

二、理论分析与假说提出

资源编排理论强调如何有效地组织、配置、利用资源,特别是可以被重新组合转化和创新的资源(如知识资源)。大数据本身就是一种新型知识资源,大数据应用作为一种新的技术手段,不仅可以丰富企业的数据资源,还会带动企业其他知识资源的流动和重组,而知识资源是企业进行突破性技术创新的关键要素。基于此,本文主要从大数据应用对企业知识资源的优化角度探讨其对企业突破性技术创新的

影响。同时,知识多元化是企业知识资源优化的重要表现之一,而知识吸收能力是企业优化知识资源的关键能力,因此,本文进一步探究了大数据应用通过推动企业知识多元化和提高企业知识吸收能力来促进企业突破性技术创新的传导机制。此外,考虑到在现代化经济体系中,知识产权保护发挥了重要作用,而知识产权保护会影响知识的生产和流动,进而影响企业的知识资源积累和配置,本文还分析了知识产权保护强度对大数据应用影响企业突破性技术创新的调节效应。

1. 大数据应用对企业突破性技术创新的影响

资源编排理论通过分析各种资源的可用性确定最佳资源配置,用以指导生产实践,最终达到最佳生产目标。其不但关注企业资源的独特性和稀有性,而且强调如何有效运用资源,旨在通过提升协同效应、构建动态能力等形式形成持续竞争优势。大数据应用通过数据处理技术对海量数据进行分析与挖掘,从中提取有价值的信息,实现数据与生产经营的高度融合,不仅能促进企业提升资源配置的效率,而且能驱动企业生产经营及创新模式的转型升级。突破性技术创新是企业通过引入新的技术方法显著改变现有技术轨道与产品性能的创新方式,不仅会在技术上取得突破,还会产生新思想、新模式和新形态。企业实现突破性技术创新需要学习吸收新的技术和资源,并与已有技术和资源进行融合,从而突破原有的技术框架和路径。大数据应用则有利于企业更高效地识别、获取、重组、优化内外部资源,进而助力企业实现突破性技术创新。因此,本文基于资源编排理论,主要从知识资源的视角探讨大数据应用对企业突破性技术创新的影响。

一方面,大数据应用有助于企业构建实现突破性技术创新的动态协同生态体系(王艳等,2015)^[12]。通过大数据应用搭建智能化协同创新平台,能够凭借多源数据打破企业、行业间的数据壁垒,提升信息透明度和传播效率;知识可视化模块可以将复杂的异质性知识更加直观地转化,促进企业对跨领域知识的理解与融合;协同创新则能打破时间与地域的限制,为不同的企业提供实时协作沟通的技术创新环境。突破性技术创新往往需要跨学科的知识融合与高强度的持续投入,基于大数据的协同创新平台能够通过开放耦合的知识网络汇聚技术创新要素和资源,形成强大的产学研协同创新合力(崔圣君等,2025)^[13]。此外,大数据应用还能帮助企业实现组织架构模块化,在形成内部信息共享池的同时保持与外部环境的动态匹配,提高企业应对市场变化和技术进步的敏捷性(Corsino et al., 2019)^[14]。由此,基于大数据的动态协同创新体系有效提升了企业获取、利用和整合各种创新资源的能力,使技术创新从依赖企业个体的独立创新向依托产学研体系的协同创新转变,从而为实现突破性技术创新提供了更好的条件和环境。

另一方面,大数据应用可以从根源上降低企业在突破性技术创新过程中面临的技术不确定性和市场风险,有效激励企业更加积极地开展突破性技术创新。在技术研发环节,大数据应用有助于企业通过对海量专利数据的深度挖掘与对发展趋势的智能分析,精准预判技术演进方向,减少技术研发路径错误风险和创新资源浪费(马橙等,2024)^[15];在市场拓展环节,大数据应用可以帮助企业实时抓取和分析社交媒体等多源数据,精准洞察用户的潜在需求与市场的动态变化,模拟不同市场策略下的反馈结果(张叶青等,2021)^[16],从而将模糊的市场前景转化为企业清晰的经营管理决策。大数据应用对技术和市场不确定性的双重消除,有利于企业在突破性技术创新中减少试错成本,将创新资源更集中、更高效地应用于突破性技术创新,最终实现更多更高水平的突破性技术创新。

据此,本文提出假说 H1:大数据应用能够显著提高企业突破性技术创新水平。

2. 企业知识多元化和知识吸收能力的中介作用

实现突破性技术创新需要对知识进行整合利用,大数据应用一方面有助于企业通过吸收外部知识增强知识的多样性,为企业突破性技术创新提供更丰富的知识资源;另一方面能够通过增强企业自身的知识吸收能力提高知识资源的配置和利用效率,为企业突破性技术创新提供更有效的技术路径。基于此,本文认为大数据应用能够通过推动企业知识多元化和提高企业知识吸收能力两条路径,从获取外部知识的宽度和利用内部知识的深度两个方面促进企业突破性技术创新。

(1) 推动知识多元化路径。知识多元化是企业知识领域广泛程度提高的过程,反映了企业从外部获取相关知识的多样性与差异性(Zhou et al., 2012)^[17]。知识多元化有助于企业在突破性技术创新过程中根据创新目标和战略更好地配置知识资源,尤其是将不同类型的知识进行有机组合,在外部多元化知识与内部核心知识的融合中寻求技术突破点,进而提高突破性技术创新的成功率和研发效率。大数据应用能够推动企业知识多元化,并通过知识资源协同效应为企业实现突破性技术创新补充新的知识资源。大数据应用提高了企业的知识搜索能力,拓展了外部知识的来源,能够有效促进企业内外部知识的重组和融合(Majchczak et al., 2004)^[18],从而打破企业原有技术知识体系的局限,开辟新的技术创新方向,缩短知识更新换代的时间。同时,在丰富企业多样性知识储备的基础上,大数据应用通过智能平台或社区能够进一步拓展企业接触外部信息的渠道,有利于企业通过学习和交流实现跨学科、多领域的知识资源多样化组合(王金凤等, 2020)^[19],这种跨领域知识的碰撞与融合可以为企业提供全新的技术思路与创新视角,突破原有技术框架的束缚,推动企业实现突破性技术创新。因此,大数据应用有利于企业从外部获取更多异质性知识,推动企业的知识多元化,进而助力企业突破性技术创新的实现。

据此,本文提出假说 H2:大数据应用通过推动知识多元化促进企业突破性技术创新。

(2) 提高知识吸收能力路径。知识吸收能力是企业识别和吸收新知识的价值,并将其应用于实现组织目标的能力。突破性技术创新不仅需要拥有丰富的知识,还需要科学有效地运用这些知识,企业必须具备相应的知识吸收能力才能充分利用知识多元化实现突破性技术创新。大数据应用在推动企业知识多元化的同时,还能提高企业的知识吸收能力。大数据应用有助于企业搭建良好的沟通平台,不仅能降低信息传播成本、缩短知识认知距离、提升信息沟通效率,加快企业对知识的理解与吸收,提高企业知识搜索的专业性(Benner et al., 2003)^[20],还能帮助企业识别不同组合的知识溢出,了解不同类型技术方案的潜在优劣势,为企业在突破性技术创新领域建立一定优势。大数据应用可以加速企业对外部知识的获取,增加信息碰撞和结合的可能性,进而提高企业对知识的吸收效率(余江等, 2024)^[21]。同时,大数据应用有利于形成开放共享的知识资源理念,促使企业在进行数据挖掘的同时,将多元化的知识吸收后进行商业化转化(March, 1991)^[22],不仅能提升企业的创新绩效,还能为企业突破性技术创新提供充足的资金支持。

据此,本文提出假说 H3:大数据应用通过提高知识吸收能力促进企业突破性技术创新。

3. 知识产权保护强度的调节效应

近年来,我国对于知识产权的保护越来越重视,知识产权保护强度也随之提高。适度的知识产权保护能够激励企业通过技术创新树立和维持竞争优势地位(叶云岭等, 2025)^[23]。但知识产权保护强度越高,意味着企业专利的保护期限越长,这可能抑制企业间的技术和知识交流,在一定程度上抑制企业的突破性技术创新。由于知识产权保护强度会影响企业获取外部知识的成本以及自身创新被模仿的风

险,随着知识产权保护强度的增加,大数据应用对企业突破性技术创新的影响也可能发生显著变化。当知识产权保护强度处于适度的状态时,既有利于企业通过合法渠道获取创新所需的技术信息以及知识资源,又能通过适度的专利保护降低竞争对手进行技术模仿的风险,从而激励企业积极通过大数据应用加快推进突破性技术创新。然而,当知识产权保护强度进一步增加时,大数据应用对企业突破性技术创新的促进作用可能会受到一定程度的制约。这是因为过度的知识产权保护会加剧企业的知识获取难度。在严格的知识保护制度下,企业获取外部知识需要经过较多的流程及授权,同时,拥有核心知识的企业为了保持竞争优势可能会提高知识交流门槛,阻碍知识资源的跨组织流动。这使得大数据应用对企业知识多元化的推动作用难以有效发挥,并会降低企业创新的开放性,不利于突破性技术创新的实现。同时,部分企业可能会由于难以承担相关成本,将研发资源更多地投入技术难度较小、研发周期较短的渐进性或普通技术创新,导致突破性技术创新的投入和产出减少(刘光富等,2025)^[24]。因此,由于“专利丛林”对企业市场价值的减损作用(罗恺等,2023)^[25],加上突破性技术创新比其他创新的颠覆性更强、技术难度更高、外溢效应更强,知识产权保护强度的增加可能会强化企业知识的相对封闭性,使大数据应用对知识交流的促进作用被抑制,从而抑制企业的知识多元化,并使得大数据对企业突破性技术创新的促进作用降低。

据此,本文提出假说 H4:知识产权保护强度的提高会弱化大数据应用对企业突破性技术创新的促进作用。

三、实证检验设计

1. 基准模型设定

为检验大数据技术应用对企业突破性技术创新的影响,构建如下双向固定效应模型:

$$Inno_{it} = \beta_0 + \beta_1 Data_{it} + \lambda Controls_{it} + \mu_i + \epsilon_t + \varepsilon_{it}$$

其中,下标 i 和 t 分别代表企业和年份,被解释变量 ($Inno_{it}$)“突破性技术创新”为企业 i 在 t 年的突破性技术创新水平,核心解释变量 ($Data_{it}$)“大数据应用”为企业 i 在 t 年的大数据应用水平, $Controls_{it}$ 表示控制变量集合, μ_i 和 ϵ_t 分别表示个体(企业)固定效应与时间(年份)固定效应, ε_{it} 为随机误差项。

(1)企业突破性技术创新水平的测度。若企业申请的专利属于其此前未涉及的专利分类,则表明企业在其不熟悉的技术领域开展了探索,此类专利创造了新知识,具有突破性和创造性。基于此,本文借鉴李哲等(2021)^[26]和毕晓方等(2022)^[27]的做法,按照国际专利分类法(IPC)对专利进行分类,进而通过企业专利所反映的技术领域来确定其是否属于突破性技术创新。具体来讲,将企业 t 年之前所获得的所有专利放入专利池,若 t 年申请的专利 A 主分类号前三位在专利池中从未出现过,则该专利视为突破性技术创新,对所有属于突破性技术创新的专利进行加总得到“突破性技术创新”变量。

(2)企业大数据应用水平的测度。本文采用基于企业年报的文本分析法测算样本企业的大数据应用水平。借鉴柏淑嫒(2024)^[28]的研究,采用“大数据”“数据挖掘”“文本挖掘数据可视化”“异构数据”“征信”“AR”“增强现实”“混合现实”“虚拟现实”“数据科学”“MR”等作为大数据应用相关关键词,计算其在企业年报中出现的总词频,并进行对数化处理(加1后取自然对数),以此衡量样本企业的大数据应用水平。

(3)控制变量的选取。借鉴李哲等(2021)^[26]、孙献贞(2024)^[10]的研究,从企业层面选取以下控制变量:一是“资产负债率”,以负债合计与资产总计之比衡量;二是“总资产净利润率”,以净利润与平均资

产总额之比衡量;三是“存货占比”,以存货净额与总资产之比衡量;四是“董事规模”,以董事会人数的自然对数值衡量;五是“独立董事占比”,以独立董事人数占董事总人数的比例衡量;六是“上市年限”,以当年年份减去上市年份后加一的自然对数值进行衡量;七是“大股东资金占用”,以其他应收款占总资产之比衡量;八是“托宾 Q 值”,以市场价值与资产总值之比衡量;九是“机构投资者持股比例”,以机构投资者持股总数与流通股本之比衡量。

2. 中介作用检验模型

为检验知识多元化和知识吸收能力的中介作用,参考江艇(2022)^[29]的研究,构建如下计量模型:

$$Med_{it} = \alpha_0 + \alpha_1 Data_{it} + \lambda Controls_{it} + \mu_i + \epsilon_t + \varepsilon_{it}$$

$$Inno_{it} = \gamma_0 + \gamma_1 Med_{it} + \lambda Controls_{it} + \mu_i + \epsilon_t + \varepsilon_{it}$$

其中, Med_{it} 为机制变量,其他变量与基准模型一致。根据前文理论分析,选取以下机制变量:一是“专利多元化”和“产品多元化”,用以反映企业的知识多元化程度。知识多元化程度越高,则企业涉及的技术和产品类别越多,因此本文从技术和产品两个维度衡量企业的知识多元化程度。借鉴徐露允等(2017)^[30]的做法,基于不同类别专利占比,使用 Teachman 熵指数测度企业专利所包含的知识元素多样化程度,得到变量“专利多元化”,其值越大则企业专利涉及的技术领域越广。借鉴曾春华和杨兴全(2012)^[31]的研究,构建企业产品经营是否多元化的虚拟变量(若企业经营产品的行业数大于 1,取值为 1,否则取值为 0),作为“产品多元化”变量。二是“知识吸收能力”。知识吸收能力是企业从外部获取与自身发展相关知识,对其进行消化吸收,并将外部知识与内部资源进行整合,最终转换为企业自身价值的能力(Cohen et al., 1990)^[32]。知识吸收能力越强,意味着企业将静态知识转变为动态应用的能力越强,越能为突破性技术创新提供高效的转化机制。借鉴 Wu 等(2016)^[33]的研究,同时考虑到人力资本是决定知识吸收能力的重要因素,选取研发强度(研发投入与营业收入之比)和研发人员占比两个指标,对其进行对数化处理,进而采用熵权法计算得到“知识吸收能力”变量。

3. 调节效应检验模型

为检验知识产权保护强度的调节效应,构建如下调节效应模型:

$$Inno_{it} = \delta_0 + \delta_1 Data_{it} + \delta_2 Mod_{it} + \beta_3 Data_{it} \times Mod_{it} + \lambda Controls_{it} + \mu_i + \epsilon_t + \varepsilon_{it}$$

其中, Mod_{it} 为调节变量“知识产权保护强度”。知识产权保护强度越高,意味着企业获取外部知识的成本越高,会在一定程度上对企业突破性技术创新产生不利影响。借鉴龙小宁等(2018)^[34]的研究,采用国家知识产权局发布的《全国知识产权发展状况评价报告》中披露的省级层面知识产权保护指数的三次方作为“知识产权保护强度”的代理变量。其中,知识产权保护指数 = 立法指数 × 执法指数,立法指数从立法的角度定量化知识产权保护强度,执法指数则从执法机构角度量化知识产权保护效率和力度。

4. 样本选择和数据处理

本文以沪深 A 股上市公司为研究样本,样本期间为 2009—2023 年,剔除在样本期内发生 ST、*ST、PT 的样本以及关键数据缺失严重的样本,最终得到 35 766 个观测值。被解释变量、机制变量和控制变量相关数据来自中国统计年鉴、CSMAR 数据库、中国研究数据服务平台(CNRDS)以及企业年报等,解释变量的数据来自巨潮资讯网。对所有连续变量进行上下 1% 的缩尾处理,主要变量的描述性统计结果见表 1。

表 1 主要变量的描述性统计结果

变 量	观测值	均值	标准差	最小值	中位数	最大值
被解释变量 突破性技术创新	35 766	2.968	4.010	0.000	1.000	19.000
核心解释变量 大数据应用	35 766	1.252	1.298	0.000	1.099	4.615
中介变量	专利多元化	16 668	1.360	0.780	0.082	3.718
	产品多元化	20 491	0.581	0.493	0.000	1.000
	知识吸收能力	15 906	0.750	0.097	0.074	1.219
调节变量 知识产权保护强度	18 885	-4.003	1.308	-10.517	-3.770	-1.634
控制变量	资产负债率	35 766	0.413	0.203	0.027	0.925
	总资产净利润率	35 766	0.042	0.065	-0.375	0.254
	存货占比	35 766	0.134	0.110	0.000	0.778
	独立董事占比	35 766	0.377	0.054	0.250	0.600
	董事规模	35 766	2.118	0.198	1.609	2.708
	上市年限	35 766	2.064	0.887	0.000	3.434
	大股东资金占用	35 766	0.013	0.021	0.000	0.202
	托宾 Q 值	35 766	2.020	1.280	0.789	16.647
	机构投资者持股比例	35 766	0.429	0.250	0.001	0.924

四、实证检验结果分析

1. 基准回归

基准回归结果见表 2。无论是否加入控制变量和固定效应,“大数据应用”对“突破性技术创新”的回归系数均在 1%的水平上显著为正,表明企业大数据应用水平的提高对其突破性技术创新水平产生了显著的正向影响,即大数据应用能够显著促进企业的突破性技术创新。由此,假说 H1 得到验证。

表 2 基准回归结果

变 量	突破性技术创新	突破性技术创新	突破性技术创新
大数据应用	0.103 *** (6.24)	0.150 *** (9.39)	0.083 *** (3.20)
资产负债率		2.526 *** (20.53)	1.438 *** (6.56)
总资产净利润率		5.046 *** (15.01)	2.880 *** (6.76)
存货占比		-1.104 *** (-5.99)	-0.876 ** (-2.28)
董事规模		0.683 *** (5.09)	0.323 (1.44)
独立董事占比		0.017 *** (3.76)	-0.001 (-0.13)
上市年限		-1.084 *** (-36.84)	-2.218 *** (-29.45)
大股东资金占用		-4.780 *** (-5.15)	-2.157 * (-1.66)
托宾 Q 值		-0.258 *** (-17.64)	-0.119 *** (-4.96)
机构投资者持股比例		1.469 *** (16.21)	1.627 *** (7.26)

续表 2

变 量	突破性技术创新	突破性技术创新	突破性技术创新
常数项	2.839 *** (99.34)	1.777 *** (4.45)	5.771 *** (8.59)
年份固定效应	未控制	未控制	控制
企业固定效应	未控制	未控制	控制
样本量	35 766	35 766	35 230
R^2	0.001	0.076	0.323

注: *、**、***分别代表在 10%、5%、1% 的显著性水平下显著,括号内为 t 值,下表同。

2. 内生性处理

(1) 工具变量法。为缓解基准模型可能存在的遗漏变量、逆向因果关系等内生性问题,本文采取工具变量法进行内生性处理。考虑到企业的大数据应用具有同群效应,会受到同行业同省份企业大数据应用水平的影响,参考 Lewbel (1997)^[35] 和李唐等 (2020)^[36] 的处理方法,将企业大数据应用水平与同行业同省份企业大数据应用水平均值之差作为“大数据应用”的工具变量。运用两阶段最小二乘法 (2SLS) 检验的结果见表 3 的 Panel A,模型通过了工具变量识别不足和弱工具变量检验,表明工具变量选取合理;基于工具变量拟合的“大数据应用”对“突破性技术创新”的回归系数在 1% 的水平上显著为正,表明在缓解内生性问题后,大数据应用显著促进了企业突破性技术创新的结论依然成立。

(2) 倾向得分匹配 (PSM) 检验。为缓解样本选择偏差导致的内生性问题,采用 PSM 法进行样本匹配。参照万攀兵等 (2021)^[37] 的做法,将样本期内年报中大数据应用相关关键词未出现或仅出现一次的企业作为控制组,出现过一次以上的企业作为处理组,以基准模型的控制变量为协变量,采用 1:1 近邻匹配法进行样本匹配。虽然匹配导致样本量减少,但匹配后的处理组与控制组的倾向得分分布更接近 (参见图 1 和图 2,匹配后的重叠区域扩大且曲线形态趋同),同时,匹配后各协变量的标准化偏差大幅下降 (远低于 10% 的阈值),平均偏差、中位数偏差、不平衡指标等指标均显著改善,表明匹配有效缩小了组间差异,大幅降低了样本选择偏差,匹配效果较好。采用匹配后样本的检验结果见表 3 的 Panel B,“大数据应用”的回归系数仍在 1% 的水平上显著为正。

表 3 内生性检验结果

变 量	Panel A: 工具变量法 (2SLS)		Panel B: PSM 检验
	大数据应用	突破性技术创新	突破性技术创新
大数据应用		0.079 *** (2.65)	0.083 *** (2.80)
工具变量	0.981 *** (346.55)		
Kleibergen-Paaprk LM 统计量		1 708.046 [0.000]	
Cragg-Donald Wald F 统计量	9.8e+05 {16.38}		
样本量	35 230	35 230	35 230
R^2	0.966	0.039	0.323

注: (1) [] 中为 P 值, { } 中为 10% 显著性水平的临界值。(2) 所有模型均控制了控制变量以及年份和企业固定效应,限于篇幅,控制变量及常数项的估计结果略,下表同。

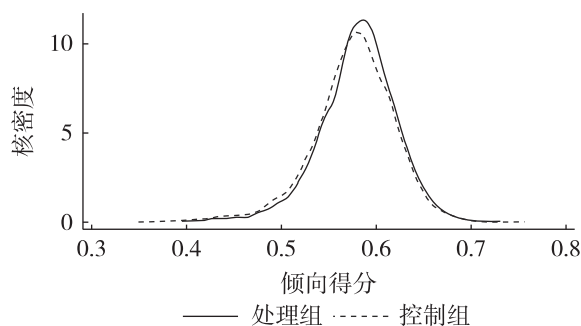


图1 样本匹配前

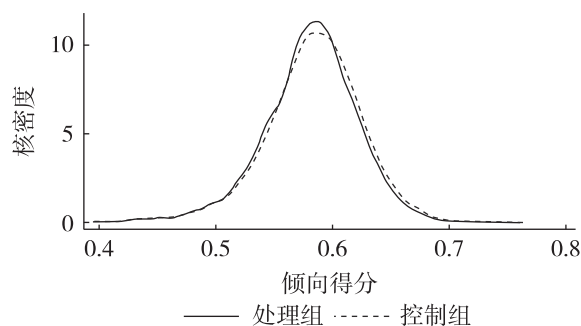


图2 样本匹配后

3. 稳健性检验

(1) 替换被解释变量。借鉴李哲等(2021)^[26]和毕晓方等(2022)^[27]的做法,基于企业突破性专利的种类数计算得到变量“突破性技术创新1”;借鉴胡山等(2022)^[38]的做法,根据三种专利的科技创新含量,通过“发明专利申请数 $\times 0.5$ +实用新型专利申请数 $\times 0.3$ +外观设计专利申请数 $\times 0.2$ ”(对数化处理)计算得到变量“突破性技术创新2”。以上述2个变量为被解释变量的检验结果见表4的Panel A。

(2) 增加行业固定效应。考虑到控制不同层面的固定效应可能会对回归结果产生影响,在基准模型中增加行业固定效应,检验结果见表4的Panel B。

(3) 剔除直辖市样本。考虑到直辖市在基础设施、数字化转型、经济发展及政策支持等方面可能存在显著优势,删除直辖市样本后重新进行基准模型检验,回归结果见表4的Panel C。

(4) 缩短样本期间。2015年9月,我国首个大数据综合试验区正式启动。基于此,剔除2016年之前的样本,采用2016—2023年的样本重新进行基准模型检验,回归结果见表4的Panel D。

(5) 聚类标准误。样本数据通常存在聚类结构,忽略组内相关性可能导致显著性水平虚高。因此,进一步对标准误按企业(股票代码)进行聚类调整,以消除可能存在的异方差和自相关问题,检验结果见表4的Panel E。

上述稳健性检验结果均显示,企业大数据应用水平的提高显著促进了企业突破性技术创新,表明本文的分析结果具有良好的稳健性。

表4 稳健性检验结果

变 量	Panel A: 替换被解释变量		Panel B: 增加固定效应
	突破性技术创新1	突破性技术创新2	突破性技术创新
大数据应用	0.033 *** (2.79)	0.012 *** (2.93)	0.079 *** (3.03)
样本量	35 230	35 230	35 229
行业固定效应	未控制	未控制	控制
R^2	0.337	0.770	0.325
变 量	Panel C: 剔除直辖市样本	Panel D: 2016—2023 年样本	Panel E: 个体层面聚类标准误
	突破性技术创新	突破性技术创新	突破性技术创新
大数据应用	0.070 ** (2.45)	0.084 ** (2.32)	0.083 *** (2.80)
样本量	28 653	24 458	35 230
R^2	0.317	0.387	0.323

4. 中介作用检验

中介作用的检验结果见表5。从推动企业知识多元化路径来看(见表5的Panel A):“大数据应用”对“专利多元化”和“产品多元化”的回归系数均显著为正,表明大数据应用能够显著提升企业的知识多元化程度;“专利多元化”和“产品多元化”对“突破性技术创新”的回归系数也显著为正,表明知识多元化程度的提高显著促进了企业突破性技术创新。知识多元化拓宽了企业知识基础的广度,有助于企业开发新知识,从而促进企业积极开展更多的突破性技术创新(王金凤等,2020)^[19]。从提高企业知识吸收能力路径来看(见表5的Panel B):“大数据应用”对“知识吸收能力”的回归系数显著为正,表明大数据应用能够显著提升企业的知识吸收能力;“知识吸收能力”对“突破性技术创新”的回归系数也显著为正,表明知识吸收能力的提高显著促进了企业突破性技术创新。由此,假说H2和H3得以验证,即大数据应用能够通过推动企业知识多元化和提高企业知识吸收能力两条路径促进企业突破性技术创新。

表5 中介作用检验结果

变量	Panel A:推动知识多元化路径				Panel B:提高知识吸收能力路径	
	专利多元化	突破性技术创新	产品多元化	突破性技术创新	知识吸收能力	突破性技术创新
大数据应用	0.018*** (3.16)		0.005** (2.01)		0.001* (1.78)	
专利多元化		0.940*** (14.41)				
产品多元化				0.539*** (5.03)		
知识吸收能力						1.980* (1.80)
样本量	16 192	16 192	20 159	20 159	15 443	15 443
R ²	0.683	0.378	0.775	0.353	0.923	0.368

5. 调节效应分析

知识产权保护会对企业获取知识资源、配置内部知识资源等产生一定影响,进而对大数据应用影响企业突破性技术创新产生调节作用。以“知识产权保护强度”为调节变量,采用全样本的检验结果见表6的Panel A。“大数据应用”的回归系数依然显著为正,虽然“知识产权保护强度”的系数不显著,但交互项(大数据应用×知识产权保护强度)的系数显著为负,表明企业所在地区知识产权保护强度的提高弱化了大数据应用对企业突破性技术创新的促进作用,假说H4得到验证。

根据前文理论分析,知识产权保护强度对大数据应用促进企业突破性技术创新具有负向调节作用,主要原因在于知识产权保护强度的提高会抑制企业间的知识交流,缩小企业的知识吸收带宽,不利于企业拓宽知识资源广度,从而抑制企业通过大数据应用进行交叉融合创新。而知识产权保护与知识流动之间的矛盾受市场竞争的影响,使得这种调节效应表现出基于竞争程度的行业异质性。在竞争性行业,一方面知识流动本身较为频繁,知识产权保护强度提升对知识流动的影响较大;另一方面,由于企业数量多、产品差异化程度低、技术更新快,知识流动对企业的突破性技术创新具有重要意义。因此,知识产权保护强度对大数据应用促进企业突破性技术创新的负向调节作用较为显著。而在管制性行业,知识流动性本身相对较弱,知识流动受大数据应用和知识产权保护强度的影响较小;同时,头部企业更容易凭借其优势地位获取更多知识资源,知识产权保护强度提高会使头部企业利用专利保护封锁部分关键

知识资源与技术路径,形成“专利丛林”(赵曦,2007)^[39],从而阻碍知识流动和突破性技术创新。因此,虽然知识产权保护强度提升对企业突破性技术创新的负面影响较大,但由于大数据应用对企业突破性技术创新的影响本身较小,知识产权保护强度对大数据应用促进企业突破性技术创新的负向调节作用较小。对此,本文根据《上市公司行业分类指引(2012年修订)》,将行业代码为B、C25、C31、C32、C36、C37、D、E48、G53、G54、G55、G56、I63、I64、K、R的行业划分为管制性行业,其余行业归为竞争性行业,进行分样本检验,回归结果见表6的Panel B和Panel C。总体上看,检验结果符合理论预期:竞争性行业的分析结果与全样本分析结果类似,知识产权保护强度的提高弱化了大数据应用对企业突破性技术创新的促进作用;在管制性行业,知识产权保护强度的提高显著抑制了企业突破性技术创新,而大数据应用对企业突破性技术创新的影响以及知识产权保护强度的调节效应均不显著。

表6 调节效应分析结果

变 量	Panel A:全样本	Panel B:竞争性行业	Panel C:管制性行业
	突破性技术创新	突破性技术创新	突破性技术创新
大数据应用	0.250 *** (2.73)	0.239 ** (2.39)	0.189 (0.80)
知识产权保护强度	-0.593 (-1.05)	0.397 (0.64)	-5.919 *** (-4.07)
交互项	-0.276 * (-1.68)	-0.312 * (-1.76)	-0.033 (-0.07)
样本量	15 879	13 435	2 416
R ²	0.349	0.358	0.365

五、拓展性研究:异质性分析

1. 企业特征异质性

(1) 产权性质异质性。将样本划分为“国有企业”和“非国有企业”两个子样本,分组检验的结果见表7的Panel A。大数据应用显著促进了国有企业突破性技术创新,但对非国有企业突破性技术创新的影响不显著。其原因可能在于这两类企业在资源获取能力和社会责任履行等方面存在差异:国有企业一方面在资源获取方面存在优势,另一方面会更加积极地响应国家政策和服务国家战略,因而在国家高度重视和大力推进创新驱动发展战略的背景下,国有企业往往会获得更多更优质的创新资源,并会更加积极地通过大数据应用加快突破性技术创新。

(2) 知识多元化异质性。根据“专利多元化”的中位数将样本划分为“知识多元化水平较高”和“知识多元化水平较低”两个子样本,分组检验的结果见表7的Panel B。大数据应用显著促进了知识多元化水平较高企业突破性技术创新,但抑制了知识多元化水平较低企业突破性技术创新。其原因可能在于企业进行突破性技术创新需要具备一定的知识条件:突破性技术创新往往涉及不同的知识领域,当企业的知识多元化水平较低时,其知识储备可能难以支撑突破性技术创新,企业会倾向于将大数据应用带来的知识资源增量用于渐进性的技术创新,导致对突破性技术创新的投入反而减少;而当企业的知识多元化水平达到一定高度后,进行突破性技术创新的知识条件得到改善,大数据应用则会促使企业利用异质性知识进行交叉融合创新,将知识多元化转化为突破性创新能力,进而促进企业的突破性技术创新。

表 7 企业特征异质性分析结果

变 量	Panel A:产权性质异质性		Panel B:知识多元化异质性	
	国有企业	非国有企业	知识多元化水平较高	知识多元化水平较低
	突破性技术创新	突破性技术创新	突破性技术创新	突破性技术创新
大数据应用	0.126 ** (2.25)	0.056 (1.50)	0.021 *** (3.57)	-0.011 ** (-2.01)
样本量	11 365	21 231	7 661	7 707
R ²	0.338	0.337	0.678	0.419

2. 行业属性异质性检验

(1) 竞争属性异质性。按照前文分类方法,将样本划分为“管制性行业”和“竞争性行业”两个子样本,分组检验的结果见表 8 的 Panel A。“大数据应用”的回归系数均显著为正,但“竞争性行业”的系数显著性和绝对值均大于“管制性行业”,表明大数据应用对于竞争性行业企业突破性技术创新的促进作用比管制性行业企业更显著。其原因可能是:竞争强度与创新频率正相关,激烈的竞争会迫使企业通过突破性技术创新建立市场优势,以获取超额利润。因此,相比管制性行业的企业,竞争性行业的企业会更加积极、充分地利用大数据来实现突破性技术创新,使得大数据应用对突破性技术创新的促进作用更为显著。

(2) 科技属性异质性。借鉴杨兴哲和周翔翼(2020)^[40]的研究,按照证监会对上市公司的分类指引,将公司分类代码属于 C25~29、C31~32、C34~41、I63~65 和 M73 的企业划归“高科技行业”子样本,其他企业划归“非高科技行业”子样本,分组检验的结果见表 8 的 Panel B。大数据应用显著促进了非高科技行业企业突破性技术创新,但对高科技行业企业突破性技术创新的影响不显著。其原因可能是:由于高科技行业的发展高度依赖技术创新,企业间通常存在较高的技术壁垒,大数据应用并不能显著促进核心技术和关键知识的流动和共享,从而对企业突破性技术创新的影响也不显著;而非高科技行业的知识流动成本较低,企业可以通过大数据应用整合分散的知识以突破技术依赖,开辟全新的技术赛道,进而实现突破性技术创新。

表 8 行业属性异质性分析结果

变量	Panel A:竞争属性异质性		Panel B:科技属性异质性	
	管制性行业	竞争性行业	高科技行业	非高科技行业
	突破性技术创新	突破性技术创新	突破性技术创新	突破性技术创新
大数据应用	0.143 * (1.82)	0.063 ** (2.27)	0.049 (1.33)	0.144 *** (2.75)
样本量	4 977	30 215	23 300	11 886
R ²	0.318	0.331	0.322	0.344

3. 区域异质性

(1) 地理区位异质性。将样本分为“东部地区”“中部地区”“西部地区”3 个子样本,分组检验的结果见表 9 的 Panel A。大数据应用显著促进了东部地区企业突破性技术创新,但对中部和西部地区企业突破性技术创新的影响不显著。其原因可能在于:东部地区的数字基础设施与数字化转型水平较高,形

成了基础设施、人才、政策、产业协同的良性循环,大数据应用的场景丰富,使大数据应用可以在多方面直接赋能企业的突破性技术创新;而中西部地区的数字化进程相对滞后,数据流通和处理成本较高,大数据应用不仅水平较低,还较难有效转化为创新驱动动力,限制了其对企业突破性技术创新的促进作用。

(2)市场化水平异质性。以各年度省级区域市场化指数的中位数为划分标准,根据企业所在省份的市场化指数将样本分为“市场化水平较高”和“市场化水平较低”两个子样本,分组检验的结果见表9的Panel B。大数据应用显著促进了市场化水平较高地区企业突破性技术创新,但对市场化水平较低地区企业突破性技术创新的影响不显著。其原因可能在于:在高市场化水平的地区,市场机制较成熟,经济主体活力较强,要素流动和配置效率较高,为大数据应用的效应发挥提供了较好的市场基础,有利于企业通过大数据应用有效优化自身知识资源,进而实现突破性技术创新;而在市场化水平较低的地区,由于市场机制不够完善,市场扭曲会制约大数据应用对知识等要素流动和配置的优化作用,最终导致大数据应用对企业突破性技术创新的影响不显著。

表9 区域异质性分析结果

变量	Panel A:地理区位异质性			Panel B:市场化水平异质性	
	东部地区	中部地区	西部	市场化水平较高	市场化水平较低
	突破性技术创新	突破性技术创新	突破性技术创新	突破性技术创新	突破性技术创新
大数据应用	0.089*** (2.90)	-0.014 (-0.18)	0.141 (1.51)	0.095** (2.01)	0.091 (0.91)
样本量	25 042	5 796	4 382	16 926	2 607
R ²	0.329	0.304	0.311	0.371	0.332

六、结论与启示

促进企业突破性技术创新是发展新质生产力、提高经济发展质量的重要路径。本文基于资源编排理论,探讨了大数据应用通过优化企业知识资源促进企业突破性技术创新的机理,进而采用沪深A股上市公司2009—2023年的数据进行实证检验,研究发现:第一,大数据应用显著促进了企业突破性技术创新,该结论在经过一系列内生性处理和稳健性检验后依然成立。第二,大数据应用能够通过推动企业知识多元化和提高企业知识吸收能力两条路径优化企业知识资源,进而促进企业突破性技术创新。第三,地区知识产权保护强度的提高会弱化大数据应用对企业突破性技术创新的促进作用,该调节效应在竞争性行业中显著,在管制性行业中不显著。第四,大数据应用显著促进了国有企业、知识多元化水平较高企业、非高科技行业企业、东部地区企业、市场化水平较高地区企业的突破性技术创新,对非国有企业、高科技行业企业、中部和西部地区企业、市场化水平较低地区企业的突破性技术创新没有显著影响,对知识多元化水平较低企业的突破性技术创新具有显著负向影响;此外,相比管制性行业企业,大数据应用对竞争性行业企业突破性技术创新的促进作用更显著。

根据上述研究结论,本文得到以下启示:第一,企业应持续推进大数据应用,并着力提升知识多元化水平和知识吸收能力,充分发挥大数据应用的知识资源优化效应,进而不断提高突破性技术创新能力和水平。同时,政府应引导企业积极对接国家创新驱动发展战略,推广大数据综合试验区,完善数据信息披露机制,推进“东数西算”,优化算力调度,降低企业算力成本,为企业加快大数据应用提供有力的政策保障和良好的发展环境。第二,要正确认识知识产权保护与知识流动之间的矛盾,合理平衡知识产权保护与突破性技术创新的关系。应避免一刀切的知识产权强保护政策,构建兼顾知识流动与创新激励的动态保护体系;制定差异化的知识产权保护周期与范围,在知识产权保护强度较高的地区加强知识流动

和共享服务;推动行业协会搭建安全可信的知识交易平台,降低中小企业获取知识资源的成本与风险。第三,实施差异化发展战略,激励知识流动,促进技术突破。比如:在高新技术研发计划和重大项目向国有企业倾斜的同时,联合高校和科研机构为非国有企业提供大数据应用技术培训和咨询服务,帮助其提升技术水平;巩固和提升东部大数据发展的基础和优势,加快中西部地区数字化转型,推动东西部算力与技术人才协同(如东部企业使用西部算力须缴纳数据资源税,反哺西部地区数字基建与人才培育);市场化水平较高的地区要充分发挥市场机制优势,市场化水平较低的地区要着力改善市场环境;鼓励行业间、企业间的知识和技术交流,提高知识资源配置效率,营造良好的知识共享与协同创新环境,有效促进企业突破技术创新。

参考文献:

- [1] 池毛毛,王俊晶,王伟军.数字化转型背景下企业创新绩效的影响机制研究——基于NCA与SEM的混合方法[J]. 科学学研究,2022,40(2):319-331.
- [2] 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):116-126,160.
- [3] 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.
- [4] 张志新,徐世超,高惠楠.大数据发展能否推动企业绿色技术创新“质效并举”——基于“国家大数据综合试验区”的准自然实验[J]. 当代经济研究,2024(4):103-115.
- [5] 陈丹,任晓刚,谢贤君.大数据发展、创新生态与企业技术创新质量——基于国家大数据综合试验区的准自然实验[J]. 中国科技论坛,2024(9):79-89.
- [6] 许芳,田萌,徐国虎.大数据应用能力对企业创新绩效的影响研究——供应链协同的中介效应与战略匹配的调节效应[J]. 宏观经济研究,2020(3):101-119.
- [7] 陶小龙,刘珊,钟雨芮,等.大数据应用与企业开放式创新的协同演化——基于扎根理论的对比性案例研究[J]. 科技进步与对策,2021,38(5):69-78.
- [8] 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.
- [9] 胡蓉宁.大数据应用对企业技术创新的效应研究[J]. 现代经济探讨,2024(3):98-108.
- [10] 孙献贞.大数据技术应用对制造业企业技术创新的影响[J]. 西部论坛,2024,34(6):64-77.
- [11] 钱秋兰,刘建江,罗双成,等.大数据应用能否促进企业创新的“增量提质”?[J]. 科学决策,2025(4):46-62.
- [12] 王艳,纪志成.基于大数据的物联网产业协同创新平台研究[J]. 贵州社会科学,2015(6):139-143.
- [13] 崔圣君,武装,于丽娜.产业合作复杂网络结构演化与资源配置研究:来自中国人工智能企业大数据的实证分析[J]. 中国软科学,2025(1):193-207.
- [14] CORSINO M, MAEIANI M, TORRISI S. Firm strategic behavior and the measurement of knowledge flows with patent citations[J]. Strategic Management Journal, 2019, 40(7):1040-1069.
- [15] 马橙,任曙明,冉晨阳.数据要素对制造业企业创新的影响研究[J]. 管理学报,2024,21(11):1642-1650.
- [16] 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.
- [17] ZHOU K Z, LI C B. How knowledge affects radical innovation: Knowledge base, market knowledge acquisition, and internal knowledge sharing[J]. Strategic Management Journal, 2012, 33(9):1090-1102.
- [18] MAJCHRZAK A, COOPER L P, NEECE O E. Knowledge reuse for innovation[J]. Management Science, 2004, 50(2):174-188.
- [19] 王金凤,王孟琪,冯立杰.外部知识异质性、知识多元化与突破式创新绩效——基于企业生命周期视角[J]. 软科学, 2020, 34(12):14-19.
- [20] BENNER M J, TUSHMAN M L. Exploitation, exploration, and process management: The productivity dilemma revisited[J].

- Academy of Management Review, 2003, 28(2): 238-256.
- [21] 余江,白宇彤,孟庆时,等.数字化转型战略对企业数字创新绩效影响研究[J].科研管理,2024,45(4):1-11.
- [22] March J G. Exploration and exploitation in organizational learning[J]. Organization Science, 1991, 2(1): 71-87
- [23] 叶云岭,单航,于海潮.知识产权保护、技术转移与企业持续创新[J/OL].当代财经,1-15(2025-05-13). <https://doi.org/10.13676/j.cnki.cn36-1030/f.20250512.002>.
- [24] 刘光富,严韵,刘嫣然.数字化转型与绿色创新——知识产权保护的调节作用[J].软科学,2025,39(7):1-7,14.
- [25] 罗恺,左祥太.专利丛林对企业市场价值影响的非均衡研究[J].科研管理,2023,44(10):168-180.
- [26] 李哲,黄静,孙健.突破式创新对分析师行为的影响——基于上市公司专利分类和引证数据的证据[J].经济管理,2021,43(5):192-208.
- [27] 毕晓方,刘晟勇,傅绍正,等.盈余平滑影响企业突破式创新吗——外部利益相关者评价的视角[J].会计研究,2022(12):91-102.
- [28] 柏淑娜,潘子成,曹伟,等.企业大数据应用对ESG评价的影响[J].世界经济,2024,47(8):133-167.
- [29] 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.
- [30] 徐露允,曾德明,李健.知识网络中心势、知识多元化对企业二元式创新绩效的影响[J].管理学报,2017,14(2):221-228.
- [31] 曾春华,杨兴全.多元化经营、财务杠杆与过度投资[J].审计与经济研究,2012,27(6):83-91.
- [32] COHEN W, LEVINTHAL D. Absorptive capacity: A new perspective on learning and innovation [J]. Administrative Science Quarterly, 1990, 35(1): 128-152
- [33] WU J, WANG C, HONG J, et al. Internationalization and innovation performance of emerging market enterprises: The role of host-country institutional development[J]. Journal of World Business, 2016, 51(2): 251-263.
- [34] 龙小宁,易巍,林志帆.知识产权保护的价值有多大?——来自中国上市公司专利数据的经验证据[J].金融研究,2018(8):120-136.
- [35] LEWBEL A. Constructing instruments for regressions with measurement error when no additional data are available, with an application to patents and R&D[J]. Econometrica, 1997, 65(5): 1202-1213.
- [36] 李唐,李青,陈楚霞.数据管理能力对企业生产率的影响效应——来自中国企业—劳动力匹配调查的新发现[J].中国工业经济,2020(6):174-192.
- [37] 万攀兵,杨冕,陈林.环境技术标准何以影响中国制造业绿色转型——基于技术改造的视角[J].中国工业经济,2021(9):118-136.
- [38] 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.
- [39] 赵曦.专利丛林法则研究[D].重庆:西南政法大学,2007.
- [40] 杨兴哲,周翔翼.治理效应抑或融资效应? 股票流动性对上市公司避税行为的影响[J].会计研究,2020(9):120-133.

Impact of Big Data Applications on Enterprises' Breakthrough Technological Innovation: The Path of Optimizing Knowledge Resources and the Regulatory Role of Intellectual Property Protection

WANG Jun, ZHOU Xiao-yu

(School of Economics and Management, China University of Petroleum (East China), Qingdao 266580, Shandong, China)

Summary: The deep integration of big data applications and breakthrough technological innovation is pivotal to advancing China's national big data strategy. Through the application of big data, enterprises can

enhance their dynamic construction capabilities, improve synergy effects, and build a sustainable competitive advantage. This creates opportunities for breakthrough technological innovation and contributes to the development of new quality productive forces, carrying significant economic and practical significance. However, few studies have explored this from the perspectives of knowledge diversity and knowledge absorption capacity.

Based on panel data from Shanghai and Shenzhen A-share listed companies spanning 2009 to 2023, this study adopts a resource orchestration perspective to theoretically and empirically investigate how big data applications affect enterprises' breakthrough technological innovation. The findings reveal that big data applications significantly drive breakthrough technological innovation, with conclusions remaining robust after comprehensive endogeneity and robustness tests. Two mediating pathways are identified: enhancing enterprises' knowledge diversity and improving knowledge absorption capabilities. Heterogeneity analyses further reveal that the positive effect is more pronounced among state-owned enterprises, enterprises with higher levels of knowledge diversity, those located in eastern China or provincial capitals, and enterprises operating in competitive or non-high-tech industries. Additionally, the intensity of intellectual property protection has a negative moderating effect in the process of big data applications promoting enterprises' breakthrough technological innovation.

Compared with previous studies, this study makes two expansions. First, in terms of research objects, it distinguishes itself from traditional innovation methods, classifies innovation into detailed categories, incorporates big data applications and breakthrough technological innovation into a unified research framework, and deeply explores the value impact and internal transmission mechanism of enterprises' big data applications, filling the research gap on breakthrough technological innovation in big data applications. Second, in terms of research level and mechanism, it focuses on the micro-level of enterprises, embeds knowledge diversity and knowledge absorption capacity into the research framework of the relationship between big data applications and breakthrough technological innovation, and explores the influence mechanism of big data applications, or the underlying mechanism of digitalization, on breakthrough technological innovation, providing a new perspective for research and deeply revealing the fundamental reasons affecting enterprises' breakthrough technological innovation.

This study sheds light on the internal logic between enterprises' big data applications and achieving breakthrough technological innovation, which is conducive to the government departments formulating more targeted big data development strategies, providing personalized policy support for different types of enterprises, and better guiding enterprises to achieve breakthrough technological innovation.

Keywords: big data application; breakthrough technological innovation; knowledge resource; knowledge diversity; knowledge absorption capacity; intensity of intellectual property protection

CLC number: F273. 1; F124. 3

Document code: A

Article ID: 1674-8131 (2025) 05-0045-16

(编辑:刘仁芳;朱 艳)