

数据要素流动如何影响企业创新持续性?

单浩男 赵 凯

摘要:数据要素的低成本流动性与可复用性为解释企业创新持续性提供了新视角。本文选取 2011—2023 年沪深 A 股上市企业与地级及以上城市数据,实证分析数据要素流动对企业创新持续性的影响。研究发现,数据要素流动促进了企业创新投入持续性和创新产出持续性。机制分析表明,数据要素流动以降低信息不对称、提升信息完备性为基础,沿着资源保障、竞争激励、信息决策和协同创新四条路径影响企业创新持续性。异质性分析结果显示,数据要素流动对企业创新持续性的促进作用在发展潜力较高的企业以及人才引进强度较高地区的企业中更强。本文的研究不仅深化了对数据要素流动影响企业创新持续性内在机制的理解,也为理解数据要素流动的微观创新效应提供了更为具体的经验证据。

关键词:数据要素流动 企业创新持续性 融资约束 市场竞争 不确定性感知 合作意识

中图分类号:F49;F273.1

文献标识码:A

文章编号:1000-7636(2026)10-0083-17

一、问题提出

第四次工业革命浪潮下,技术迭代加速,市场需求快速演变,企业经营不确定性显著上升,单次创新的竞争优势持续弱化,点状创新已难以支撑长期发展。相较于单期研发投入或专利产出,创新持续性更强调企业创新活动在时间维度的延续性与稳定性,是企业摆脱创新困境、构建长期竞争力的关键,也是破解核心技术“卡脖子”难题、落实创新驱动发展战略、实现科技自立自强的微观基础。党的二十届四中全会明确提出“提升国家创新体系整体效能,全面增强自主创新能力”的战略要求。当前中国企业创新持续性面临双重约束:一是创新活动长周期、高投入、高风险的固有属性,对企业资源整合与决策能力要求较高,部分企业持续开展高质量创新的意愿与能力不足,并可能因试错成本过高中断创新进程;二是研发投入“重应用、轻基础”的结构性问题仍然存在,2025 年全社会研发经费突破 3.9 万亿元,但基础研究占比仅约 7%,远低于发达国家水平,一定程度上制约了核心技术迭代所需的底层积累,创新易出现“断链”,长期将削弱产业韧性与国际竞争力。在此背景下,探寻提升企业创新持续性的有效路径,是理论与政策层面亟待解决的重大课题。

数据要素作为数字经济的核心生产要素,兼具非竞争性、规模报酬递增与网络效应等独特属性,可通过重构创新范式、优化资源配置为破解企业持续创新困局提供新路径,而其价值释放的重要前提是跨边界流动与反复利用。数据要素的低成本流动性与可复用性有助于突破组织与空间边界,为解释企业创新持续性

收稿日期:2026-03-24;修回日期:2026-08-26

基金项目:国家社会科学基金后期资助项目“数智化赋能高质量发展:多维解构与实践路径研究”(25FJYB016)

作者简介:单浩男 青岛理工大学商学院讲师,青岛,266520;

赵 凯 青岛大学经济学院教授、博士生导师,通信作者,青岛,266071。

作者感谢审稿快线专家的评审意见。

提供了新视角。理论上,数据要素流动可通过多重路径发挥作用:一是汇聚技术、市场、产业链与政策多维信息,降低信息不对称,提升企业机会识别与风险预判能力;二是提高合作透明度,降低搜寻与缔约成本,推动联合研发与资源互补;三是缓解银企信息摩擦,为持续研发提供外部融资支持;四是加快技术扩散、强化市场竞争,倒逼企业通过持续创新维持竞争优势。现有研究尚未明确数据要素流动对企业创新持续性的影响方向、作用强度与具体传导路径。系统考察城市间数据要素流动对企业创新持续性的影响与机制,既有助于深化对数据要素微观创新效应的认识、拓展持续创新理论,也对释放数据要素价值、服务创新驱动发展战略具有重要现实意义。

二、文献综述

回顾现有研究,与本文相关的文献主要分为以下两类。

第一类与数据要素流动有关。从现有研究的演进脉络来看,学界对数据要素流动的探讨呈现“宏观先行、微观跟进”的特征。在宏观层面,早期研究多聚焦数据要素流动的环境刻画与宏观效应。部分学者从通信基础设施、(移动)互联网发展水平及大数据政策环境三大维度构建评价体系,采用熵值法测算不同地区的数据要素流动环境指数,并从理论与实证双重维度证实该指数对人均产出的倍增效应^[1]。近年来,学界进一步突破“环境间接测度”的局限,有学者采用劳瑞(Lowry)引力模型直接测算地区数据要素流动指数,并将其作为核心解释变量,实证检验数据要素流动对区域协调发展的驱动作用^[2],推动宏观层面研究从“条件分析”向“流量效应”深化。在微观层面,考虑到企业层面的数据流动难以直接观测,部分学者以企业数字化转型指数为权重,将国家层面数据规制政策指数映射至企业维度,构建企业数据要素流动的代理变量,并基于此考察其对企业绿色创新的影响^[3]。然而,关于数据要素流动等前提条件的研究仍处于探索阶段,尚未充分关注其对企业创新活动的影响^[4]。

第二类与企业创新持续性有关。企业创新持续性的研究围绕概念界定、测度方法和影响因素展开分析,形成了较为系统的研究脉络。在概念界定方面,学界普遍认为,企业创新持续性本质上是“过往创新活动提升当前创新概率”的动态过程^[5],其核心在于强调时间连续性与动态循环性,即通过研发投入与创新产出的相互促进,为企业积累渐进性创新成果,最终支撑企业高质量发展^[6],这一界定明确了创新持续性的过程属性与价值导向。测度方法的演进呈现“国外探索-本土优化”的特征:早期国外研究采用二值变量构建动态二分类概率单位(Probit)回归模型,从投入或产出单一层面验证创新持续性^[7],但二值变量信息量有限,难以反映持续性的强度差异;后续国内研究尝试优化,通过“无形资产增量与期初总资产之比”构建动态模型,聚焦创新投入持续性的测度^[8],但未覆盖创新产出的连续特征,导致测度体系存在“投入-产出”衔接缺口。影响因素的探讨多从内部能力与外部环境双重视角展开,形成丰富实证结论。从内部能力维度来看,数字化转型作为新时期企业发展的核心方向,相关研究证实兼具创新投入与产出的双维度赋能作用^[9]。企业环境、社会与治理(ESG)表现作为衡量综合发展实力的关键指标,亦被发现对创新持续性存在正向促进作用^[10]。另外,企业规模与创新持续性之间呈现非线性的倒U型关系;同时,组织韧性和资源柔性均在企业规模与创新持续性之间发挥负向交互作用,即弱化了企业规模与创新持续性间的倒U型关系^[11]。从外部环境维度来看,企业间合作交流的影响受到关注,研究指出网络凝聚性与企业创新持续性呈倒U型关系,技术市场价值在二者关系中发挥机制作用^[12]。此外,政府行为的影响可进一步细分为政策调控与资本支持:在政策调控层面,较高的经济政策不确定性会通过延缓企业投资决策、加剧外部融资约束抑制创新持续性,而政府隐性担保能平滑创新投入波动^[13];在资本支持层面,针对中小企业的研究发现,政府风险投资可通过资源配置、风险分担机制显著增强其创新持续性,财政治理在此过程中发挥“事前服务增值-事后监督引导”的双重交互效应,且该效应受地方政府注意力与行业异质性影响^[14]。

通过上述文献梳理可以看出,现有研究围绕数据要素流动和企业创新持续性已取得较为丰富的成果,

但仍存在不足。第一,对数据要素流动的测度及其微观经济效应研究仍有待深化。一方面,关于数据要素流动的测度,现有研究未能充分反映出“流动”的动态过程与持续变化;另一方面,现有研究多聚焦于数据要素流动对区域协调发展等宏观经济问题的影响,对其企业层面微观效应的关注仍显不足。第二,鲜有研究从投入-产出双重视角,系统考察数据要素流动对企业创新投入持续性和创新产出持续性的作用效果,并从定量分析层面进行系统性分析,相关研究有待进一步补充和完善。据此,本文的边际贡献主要体现在三个方面。第一,在测度方法方面将修正引力模型与复杂网络分析法相结合,构建兼顾“流动强度”与“网络关联”的多维测度指标,在一定程度上弥补现有研究多依赖单一代理变量的不足,为数据要素流动的量化研究提供新的可行路径。第二,将数据要素流动与企业创新持续性纳入同一分析框架,拓展了数据要素流动与企业创新持续性关系的研究,同时回应企业创新持续性不足的现实问题,拓展了数据要素经济效应与企业创新持续性的研究视角。第三,将创新持续性解构为创新投入持续性与创新产出持续性两个核心维度,结合多重机制展开精细化剖析,不仅深化了对数据要素流动影响企业创新持续性内在逻辑的理解,也为认识数据要素流动的微观创新效应提供了更具体的经验证据。

三、理论分析与研究假设

(一)数据要素流动对企业创新持续性的影响

本文界定的数据要素流动有别于数字技术应用、公共数据开放与企业数字化转型,指数据要素依托数字基础设施与市场机制,跨地区、跨主体、跨场景实现传输、整合与复用的过程,其核心在于突破组织与空间边界实现数据跨域再配置。数据要素流动体现为跨区域连接、数据可达调用、要素传递再配置三类表征,对应本文选取的连通性、访问性、传递性三个刻画维度,分别反映数据联系广度、外部数据获取便利度与跨域扩散的机制作用。其影响企业创新持续性的核心逻辑并非单纯增加数据供给,而是通过跨边界网络化连接重构企业的信息、合作与资源匹配结构,将分散数据转化为可利用、可匹配的创新资源,支撑企业长期持续的创新投入与成果产出。

具体而言,首先,数据要素流动能够改善企业创新决策的信息基础。创新活动具有长周期和高不确定性特征,企业是否持续进行研发投入,很大程度上取决于其能否准确识别技术趋势、市场需求、政策变化和潜在风险。数据要素跨区域、跨主体流动后,原本分散于不同地区、产业链环节和市场主体中的信息能够被更充分地连接和传递,从而帮助企业降低信息搜寻成本,提高创新机会识别能力和风险判断能力^[15]。与一般数字化转型强调企业内部技术应用不同,数据要素流动更突出外部数据资源进入企业创新决策过程所形成的信息增量,这种信息增量有助于减少企业因环境不确定和创新路径模糊而产生的创新收缩行为。

其次,数据要素流动能够重塑企业持续创新所需的资源匹配关系。持续创新不仅需要企业内部持续投入研发资源,也需要外部资金、技术、人才和合作机会的协同支持。数据要素在不同主体之间流动,可以提升企业经营状况、创新能力和项目前景的可观察性,缓解企业与金融机构、合作伙伴之间的信息不对称,进而降低融资难度和合作搜寻成本。同时,数据流动有助于提高市场信息透明度,推动行业知识扩散和竞争格局调整,促使企业通过持续创新保持差异化优势^[16]。因此,数据要素流动不仅能够改善企业信息识别能力,还能够强化企业的外部联系与资源匹配能力。

最后,从市场层面来看,数据要素流动加速信息传递,减少企业间信息不对称,降低创新资源的搜寻与匹配成本,提高创新要素的配置效率^[3];从政策与制度层面看,完善的数据治理制度能够规范数据交易与使用行为、保障数据使用的合法性与安全性,为数据要素有序流动提供制度保障;在稳定的制度预期下,有序流动的数据要素有助于降低企业获取、整合和再利用数据的成本,缓解创新过程中的不确定性,从而强化其对企业持续创新的促进作用。

基于上述分析,本文提出假设1:数据要素流动能够提升企业创新持续性。

(二)数据要素流动影响企业创新持续性的作用机制

数据要素流动影响企业创新持续性的底层逻辑在于降低主体之间的信息不对称并提升创新决策的信息完备性。本文所讨论的各类机制并非相互割裂的并列关系,而是以数据要素流动带来的信息连接、信息扩散和信息再配置为共同基础,并分别作用于企业持续创新的不同环节。具体而言,数据要素流动通过提高企业经营状况和创新项目质量的可观察性,缓解企业与金融机构之间的信息摩擦,形成资源保障路径;通过降低行业信息壁垒、加快技术扩散和提高市场透明度,强化企业持续创新的竞争激励,形成竞争激励路径;通过跨区域、跨主体和跨场景的数据连接,使企业能够更充分地获取市场需求、技术演进、政策环境和产业链运行信息,从而降低对外部环境和创新路径的不确定性感知,形成信息决策路径;通过提升信息透明度和可验证性,降低企业识别合作对象、评估合作收益和防范机会主义行为的成本,进而提高合作创新可能性,形成协同创新路径。由此,数据要素流动并不是通过单一渠道影响创新持续性,而是沿着资源保障、竞争激励、信息决策和协同创新四条路径共同作用于企业创新投入持续性和创新产出持续性。

1. 资源保障路径:融资约束缓解

与企业数字化转型更多强调企业内部数字技术应用及治理优化不同,数据要素流动更强调提升金融机构、政府和投资者对企业经营、产业链关系和市场反馈等外部信息的可获得性,从而改善外部主体对企业创新项目质量的识别能力。信息不对称理论表明,企业与外部资金提供者之间的信息壁垒是导致融资约束的关键因素^[17]。而数据要素流动的不断增强,可以通过信号传递和风险定价优化两条路径缓解这种信息摩擦。从信号传递的角度来看,数据要素的高效流动可以帮助企业向外部投资者展示其生产经营状况,提高运营决策机制的透明度,将传统财务报表中难以量化的“软信息”转化为可验证的“硬证据”,进而拓宽融资渠道,提高企业融资可得性^[18]。同时,数据要素流动契合“数字中国”战略和“数据要素×”行动计划的政策导向,使企业在申请信贷支持或股权融资时,更易获得政府主导的风险补偿、利息补贴等政策支持,激活市场化融资与政策性扶持的协同效应。从风险定价优化的角度来看,金融机构通过整合企业内外部多元数据(如供应链分布、舆情数据等),并借助大数据分析、机器学习等手段构建多维风险评估体系,这一过程不仅能够有效降低信息不对称引致的风险溢价,还可以拓展可抵押资产的边界,破解传统信贷中过度依赖固定资产抵押的困境,进而缓解企业因信息不充分而面临的融资约束和信贷配置偏差。

创新具有长周期、高投入、高风险的特征,维持创新持续性需要稳定的资金支持,融资约束的缓解为企业持续开展创新活动提供了重要保障。一方面,外部融资可得性的提升直接扩大了企业可用于创新活动的资金来源,使得持续稳定的创新投入不再过度依赖于企业内部现金流,而企业融资渠道的拓宽也会被资本市场解读为“创新潜力认可信号”,进而吸引更多风险投资与战略合作,形成“融资—投入—再融资”的良性循环。另一方面,多元化的融资渠道有助于企业平滑创新风险,当企业面临短期内的营收波动时,外部资金的注入能够起到缓冲作用,有效规避创新活动中断的风险。另外,基于数据要素流动的融资模式还具有风险共担特征,投资者可以通过更加及时、充分的信息了解企业创新进程,与企业形成收益共享、风险共担的长期合作关系,增强企业持续开展创新活动的能力与意愿。

基于上述分析,本文提出假设2:数据要素流动通过缓解融资约束提升企业创新持续性。

2. 竞争激励路径:市场竞争增强

数据要素流动不是简单提高企业数字技术应用水平,而是通过降低行业信息壁垒、提高需求信息和技术信息的可得性,使更多主体能够识别市场机会并进入细分赛道,从而改变企业面临的竞争环境。需要说明的是,本文并不认为任何程度、任何形式的市场竞争都会无条件促进创新持续性,而是强调在数据要素流动提高市场信息透明度、降低信息壁垒和加快技术扩散的背景下,适度竞争压力可能成为推动企业持续创

新的重要外部激励。在传统市场结构中,头部企业往往会凭借其较高的市场占有率和庞大的产销体系构建信息“护城河”,形成“数据垄断租金”^[19]。然而,随着数据要素流动水平的不断提高,新进入者能够借助更加畅通的数据获取与共享渠道,快速获取行业发展态势、用户需求特征等关键数据资源,缓解数据资源禀赋差异导致的竞争劣势,增强市场竞争程度。同时,数据要素流动催生“超竞争”环境^①。数据要素的高频流动加速技术扩散并缩短产品迭代周期,迫使企业将竞争焦点从传统的产能扩张转向持续创新能力的构建。不仅如此,数据要素的流动还通过激活长尾市场,起到了扩大市场竞争范围的作用,中小企业可以利用开源的数据工具精准识别多元市场需求中的细分领域,在与头部企业的竞争中形成差异化优势,进一步促进多元竞争格局的形成。

根据熊彼特创造性破坏理论,市场竞争能够对企业创新决策形成重要外部激励^[20]。一方面,数据要素流动引致的竞争加剧与产品同质化趋势,缩短了创新的红利周期并压缩企业利润空间,推动企业突破传统竞争领域的路径依赖,转向创新赛道,通过持续开展创新活动维持竞争优势,这种市场生存压力促使企业将创新从偶发行为转化为常态化的战略选择。以智能手机行业为例,安卓系统代码开源以及供应链数据透明化,导致诸如折叠屏等硬件创新的同质化周期显著缩短,头部厂商不得不进一步提高研发投入占比来保持迭代节奏。另一方面,率先突破技术瓶颈的企业能够通过技术领先获得相应的市场收益,并将其进一步投入研发创新,形成“创新—收益—再创新”的良性循环,从而强化企业持续开展创新活动的内在动力。

基于上述分析,本文提出假设 3:数据要素流动通过提升市场竞争程度提升企业创新持续性。

3. 信息决策路径:不确定性感知降低

企业不确定性感知具体是指企业在面临外部经济政策变化时的主观感受^[21]。过高的不确定性感知会抑制企业的研发投入和专利申请,对企业创新决策产生负面影响。其产生主要源于两方面:一是信息获取渠道有限,二是信息处理能力有限^[22]。不确定性感知的降低并非来自企业内部流程数字化本身,而是来自跨区域、跨主体数据流动所提供的外部信息增量,使企业能够更及时地捕捉政策变化、市场需求和技术演进方向。从信息获取维度来看,数据要素的跨组织、跨地域流动打破了传统信息传递中的“孤岛效应”,企业通过对内整合生产经营数据,对外采集市场动态、政策导向等数据,逐步建立起覆盖范围广、响应速度快的立体化信息网络,提升信息获取的广度与实时性。从信息处理能力的维度来看,数据要素流动与企业资源计划(ERP)系统、大数据、物联网等数字技术深度融合,形成强大的信息转化能力,帮助企业将海量的碎片化信息转换为结构化知识。这种由数据驱动的处理能力突破了传统人工分析在处理效率和深度上的局限,使企业能够在复杂的外部环境中对趋势性变化做出更加及时、合理的判断,提升战略决策的科学性。

当企业通过数据要素流动获得更完整的经济环境认知时,创新决策的风险溢价会随之降低^[23]。管理者对技术路径的可行性、市场的接受度以及政策合规性的判断,能够由更多依赖经验判断转向更加注重数据支持的量化分析,减少了管理层的风险规避行为,使其更愿意持续开展创新活动。不仅如此,企业不确定性感知的降低有助于提高创新资源配置效率。一方面,企业能够借助数据要素驱动下的情景模拟技术,对不同创新路径下的风险收益分布进行预测,从而实现资金、人才等创新资源的精准配置,使其流向更具潜力的创新赛道。另一方面,数据要素的高效流动能够帮助企业建立动态反馈机制,为企业动态调整研发方向提供有力支持,使企业在遭遇技术瓶颈时能够及时调整或选择替代性创新路径,有效缓解因创新资源错配而可能引发的创新项目中断等问题。

基于上述分析,本文提出假设 4:数据要素流动通过降低企业的不确定性感知提升企业创新持续性。

① “超竞争”环境也称超强竞争环境,其主要表现为动态化、不确定性、难以预测、不断扩大和起伏不平,它的核心就是“变化”。超竞争环境产生的主要原因包括技术进步、信息化、经济全球化等。

4. 协同创新路径:合作意识增强

数据要素流动通过打破主体间的数据孤岛,使企业更容易识别潜在合作对象、评估合作价值和降低合作风险与成本,提高企业开展协同创新的可能性,因此其影响重点不仅体现在数据使用效率提升,更体现在合作关系和创新网络的重构。首先,数据要素的高效流动降低了组织间的信息不对称,使企业能够更加便捷地获取和利用外部数据资源,缓解了传统企业合作模式中因信息不透明导致的搜寻成本与议价成本过高等问题。在数据要素流动的赋能下,企业不需要投入大量专用性资产就能接入外部合作网络,参与跨组织的创新协作。其次,数据要素的非竞争性特征决定了其价值在流动过程中呈现网络递增效应^[24],即单一企业的数据往往只能反映自身的生产经营状况,但当不同企业的数据整合在一起时,就能形成覆盖整个产业链的数据网络,不同主体间的数据整合能够发挥信息互补与协同效应,使企业更加充分地认识跨主体数据共享与合作创新的潜在价值,从而增强合作意愿。另外,各企业在合作的过程中通常会留下包括履约记录、合作贡献等在内的合作信用数据。借助新一代信息技术,这些数据的真实性和可追溯性不断增强,并在流动过程中形成企业的“合作信用档案”,缓解了机会主义行为引起的合作风险,进一步推动企业从“被动合规”转向“主动合作”。

伴随着企业合作意识的不断增强,企业间知识共享程度逐渐深化。在基于数据要素流动构建的合作网络中,企业不再各自闭门造车,而是通过共享研发、生产、市场等多维度数据,促使不同领域的知识在相互碰撞中激发融合效应,从而催生出全新的技术组合^[25]。不仅如此,面对高投入、长周期、高风险的创新项目,单个企业通常难以独自承担压力,合作意识的增强推动企业通过组建创新联盟、技术实验室等方式,使各合作主体按照贡献度分摊创新成本,并依据合作协议分配创新收益。这种结构化合作形式优化了企业创新过程中的风险承担模式,为创新项目的持续推进提供了可靠保障。另外,企业间的合作有利于创新生态的构建。数据要素流动催生的合作网络本质上是一种创新生态系统,这种生态系统具有自我强化的正反馈机制,即数据流动促进合作网络扩张,合作网络扩张反推数据要素价值提升,使创新活动能够在技术迭代、市场波动乃至政策调整中保持持续性。

基于上述分析,本文提出假设 5:数据要素流动通过增强企业合作意识提升企业创新持续性。

四、实证设计

(一) 样本选取与数据来源

本文选取 2011—2023 年沪深 A 股上市公司数据和城市数据为初始样本。其中企业创新数据和原始财报数据来自深圳希施玛数据科技有限公司 CSMAR 中国经济金融研究数据库、万得(Wind)数据库与上海经禾信息技术有限公司中国研究数据服务平台(CNRDS),企业年报数据来源于巨潮资讯网;城市层面的数据来源于《中国统计年鉴》《中国城市统计年鉴》、北京大学数字金融研究中心以及各地市政府官网。在实证检验之前,对所收集到的数据进行了手工筛选与整理,具体操作如下:(1)考虑到金融行业和房地产行业的特殊性及其会计处理方法上的差异,剔除了金融保险行业与房地产行业样本;(2)考虑到存在经营异常以及退市风险的企业数据会对研究结果的稳健性造成负面影响,剔除 ST、*ST 以及 PT 企业样本;(3)剔除部分数据缺失严重的企业样本;(4)部分缺失值采用线性插值法补齐。最后得到包含 18 317 个观测值的样本数据。

(二) 模型设定

为验证数据要素流动的持续创新效应,本文设定如下双重固定效应模型:

$$\ln IP_{it} = \alpha_0 + \alpha_1 Df_{it} + X'_{it} \gamma + \theta_i + \mu_t + \varepsilon_{it} \quad (1)$$

其中,下标 i 表示企业, t 表示年份; $\ln IP_{it}$ 为被解释变量,包括 $\ln IIP_{it}$ 和 $\ln OIP_{it}$,分别表示第 t 年企业 i 的创新投入持续性和创新产出持续性; Df_{it} 为核心解释变量,表示数据要素流动; X'_{it} 表示影响企业创新持续性

的控制变量向量; θ_i 表示企业固定效应, μ_t 表示时间固定效应; ε_{it} 表示随机扰动项。

(三) 变量说明

1. 被解释变量

本文基于创新投入-产出的视角,借鉴杨仁发和杨梅君^[9]的研究,同时采用创新投入指标和创新产出指标的前后期对比来反映创新的持续程度。具体测算方式如下:

$$\ln IIP_{it} = \ln \left[\frac{RD_{it} + RD_{it-1}}{RD_{it-1} + RD_{it-2}} \times (RD_{it} + RD_{it-1}) + 1 \right] \quad (2)$$

$$\ln OIP_{it} = \ln \left[\frac{Patent_{it} + Patent_{it-1}}{Patent_{it-1} + Patent_{it-2}} \times (Patent_{it} + Patent_{it-1}) + 1 \right] \quad (3)$$

其中, $\ln IIP_{it}$ 和 $\ln OIP_{it}$ 分别表示企业 i 在第 t 年的创新投入持续性和创新产出持续性, RD_{it} 和 $Patent_{it}$ 分别表示企业 i 在第 t 年的创新投入以及创新产出。创新投入采用企业研发投入衡量,创新产出采用企业专利申请总数衡量, RD_{it-1} 和 $Patent_{it-1}$ 分别对应企业 i 在第 $t-1$ 年的创新投入及创新产出, RD_{it-2} 和 $Patent_{it-2}$ 则分别表示企业 i 在第 $t-2$ 年的创新投入及创新产出。该公式的核心设计逻辑是:第一,趋势延续维度。 $(RD_{it} + RD_{it-1}) / (RD_{it-1} + RD_{it-2})$ 捕捉创新投入的“动态增长连续性”。若该比值 ≥ 1 ,说明当期与前期的投入总和不低于前期与前前期的总和,体现投入趋势的延续性;若比值 < 1 ,则表明当期与前期投入总量低于前期与前两期投入总量,反映创新投入的跨期延续程度有所减弱;第二, $RD_{it} + RD_{it-1}$ 兼顾投入的“稳定规模基础”,避免因单期波动误判持续性;第三,加 1 后取自然对数。加 1 主要用于避免零值无法取对数的问题,取自然对数则有助于降低指标偏态和极端值的影响。同时,将“趋势+规模”的复合指标转化为可线性解释的变量。这一设计完全区别于单纯依据创新投入绝对规模衡量创新持续性的做法:即使企业某一期投入绝对值大,但相较于单纯依据创新投入绝对规模衡量创新持续性的做法,该指标同时考虑创新活动的跨期变化与规模基础。当企业前后期投入趋势减弱时,比值项随之下降,从而对综合指标形成向下调整。该计算方法一方面能够在一定程度上规避绝对值依赖,若企业仅单期投入(产出)绝对值高,但前后期趋势中断,则 $[(RD_{it} + RD_{it-1}) / (RD_{it-1} + RD_{it-2})] < 1$,即使乘以规模项,最终指标仍会低于“趋势延续+规模稳定”的企业,精准反映持续性差异;另一方面,公式核心权重落在前后期比值上,本质是测度创新活动的时间维度延续性而非截面维度规模。

2. 核心解释变量

本文从连通性、访问性和传递性三个维度刻画数据要素流动,分别由数据要素空间关联网络的度数中心度、接近中心度和中介中心度衡量^[26],采用熵值法测算数据要素流动的综合指标,作为本文的核心解释变量,具体操作如下:

首先,本文参考李昊骅等^[27]的研究,采用引力模型测算各个城市之间的数据要素流动强度。引力模型的建模思路最初源于物理学中的“万有引力定律”,之后被广泛应用于经济学中。具体计算公式如下:

$$\omega_{ijt} = K_{ijt} \frac{\sqrt{Data_{it} Data_{jt}}}{L_{ij}^2} \quad (4)$$

$$K_{ijt} = \left(\frac{Ind_{jt}}{Ind_{it}} \right)^{0.08} \left(\frac{DIF_{jt}}{DIF_{it}} \right)^{0.19} \left(\frac{SE_{jt}}{SE_{it}} \right)^{0.73} \quad (5)$$

其中, ω_{ijt} 是第 t 年城市 i 与城市 j 之间的数据要素流动强度。 $Data$ 表示城市的数据要素规模代理指标,本文设定其为电信业务总量与国内生产总值(GDP)的乘积。选择使用电信业务总量的原因在于,数据通信是电信业务的重要组成部分,具体包括固定因特网数据传送业务、国际数据通信业务以及无线数据传送业务等。另外,除数据通信以外的其他电信业务中也包含了与数据要素相关的内容,例如卫星通信业务中包

含的话音、数据以及视频图像等业务类型。可以看出,虽然电信业务总量并不能完全等同于一个城市的数据要素规模,但其能够反映城市数据通信、网络传输和信息承载能力,是数据要素跨区域流动的重要基础条件。在城市间流动阻力的刻画上,本文引入地理距离作为引力模型中的距离衰减因素。尽管数据要素具有可复制性和跨空间传输特征,能够在一定程度上突破物理边界,但这并不意味着数据要素流动完全不受空间距离影响。现实中,数据跨区域流动仍依赖通信网络、数据中心、算力设施、接口标准、产业联系和跨区域治理规则。地理距离越远,通常意味着网络互联、数据调用、制度协调、交易信任和合作匹配的综合成本越高。因此,本文在引力模型中使用地理距离,并不是假定数据要素只能按照物理距离流动,而是将其作为城市间数据要素流动综合摩擦成本的代理变量。

K_{ij} 表示引力系数,本文参考孙晋云等^[28]的研究,引入影响数据要素流动的相关指标对其进行修正,具体如式(5)所示。 Ind 代表城市产业结构的高级化水平,诸如智能制造、医疗等高端产业的发展有助于激活数据要素价值潜能并降低其流动壁垒^[29],为数据要素的高效流通提供基础平台; DIF 代表城市数字普惠金融的发展水平,现代金融业的发展以数据资源和数字技术为主要支撑,数字普惠金融便是其中的典型代表,反映了数据要素的金融价值^[30],能够在很大程度上影响数据要素的流动; SE 代表城市的科技投入水平,地方政府的科技投入水平代表了其对该地区高新技术发展的重视程度和支持力度,而数据要素的流动和应用需要以工业互联网、云计算等前沿技术为载体^[31],因此,科技投入也可以被视作影响数据要素流动的重要变量之一。三项指标的权重由熵值法测算得到。

其次,在式(4)和式(5)的基础之上,本文进一步构建数据要素流动的空间关联矩阵,然后选取各城市数据要素流动强度的均值作为阈值,对矩阵进行二值化处理,具体如下:

$$\omega_{matrix} = \begin{Bmatrix} \omega_{11} & \omega_{12} & \cdots & \omega_{1N} \\ \omega_{21} & \omega_{22} & \cdots & \omega_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \omega_{N1} & \omega_{N2} & \cdots & \omega_{NN} \end{Bmatrix}, \omega_{ij} = \begin{cases} 1, \omega_{ij} > \bar{\omega} \\ 0, \omega_{ij} \leq \bar{\omega} \end{cases} \quad (6)$$

其中, $\bar{\omega}$ 为数据要素流动强度的均值, N 为网络节点的个数。

最后,基于上述矩阵,本文利用软件 Gephi 计算每个城市数据要素流动的度数中心度、接近中心度以及中介中心度,并利用熵值法对这三项指标进行赋权,进而计算出数据要素流动的综合测度指标,具体如表 1 所示。

表 1 数据要素流动指标体系

维度	指标	计算公式	指标解释
连通性	度数中心度	$C_D = \frac{\sum_i a_{ij}}{N-1}$	C_D 表示度数中心度; a_{ij} 表示区域点 i 与区域点 j 之间是否存在有效联系; 区域点 i 与其他节点的有效联系越多, 度数中心度越高, 表明其连通性越强
访问性	接近中心度	$C_C = \frac{N-1}{\sum_{j \neq i} d_{ij}}$	C_C 表示接近中心度; d_{ij} 表示节点 i 与 j 之间的最短路径; 网络节点与其他节点的距离总和越短, 访问性越强
传递性	中介中心度	$C_B = \sum_{i \neq j \neq l} \frac{D_{jl}(i)}{D_{jl}}$	C_B 表示中介中心度; D_{jl} 表示连接区域点 j 和 l 之间的最短路径; $D_{jl}(i)$ 表示连接区域点 j 和 l 并通过点 i 的路径。中介中心度越高表示传递性越强

3. 控制变量

考虑到其他因素对企业创新持续性的影响,本文借鉴已有研究^[3,9],选取了一系列控制变量。企业层面,选取资产回报率(ROA ,税后净利润与总资产的比值)、托宾 Q 值($Tobin$,公司市值与资产重置成本的比值)、股权集中度($Top1$,第一大股东的持股比例)、企业规模($\ln Size$,企业总资产的自然对数)、营收增长率($Growthrate$)、企业年龄($\ln Age$,企业成立年数的自然对数)、现金流比率($Cashflow$,经营活动产生的现金流净

额与流动负债的比率)、审计意见(*Opinion*,由国际四大会计师事务所审计取 1,否则为 0)、政府补助(*Sub*,政府补助与营业收入的比值)、管理层创新意识(*IC*,企业管理层讨论与分析中创新相关关键词词频占比)。城市层面,选取人口规模(*lnPop*,年末常住人口数的自然对数)、消费水平(*lnConsume*,社会消费品零售总额的自然对数)、金融发展水平(*lnFin*,年末金融机构各项贷款余额的自然对数)。

(四) 变量描述性统计

主要变量的描述性统计如表 2 所示。可以看出,样本企业在创新投入持续性和创新产出持续性方面存在较明显差异,表明不同企业的创新持续性存在一定差异。核心解释变量数据要素流动的均值为 0.225 7,标准差为 0.232 3,表明不同城市的数据要素流动水平存在一定分化。其余控制变量的统计特征均未表现出明显异常,为后续实证检验提供了基本数据支撑。

表 2 描述性统计结果

变量类别	变量	观测值	均值	标准差	最小值	最大值
被解释变量	<i>IIP</i>	18 317	20 425.480 0	48 129.520 0	5.5907	372 925.900 0
	<i>OIP</i>	18 317	90.006 5	201.504 9	0.367 9	1 784.179 0
核心解释变量	<i>Df</i>	18 317	0.225 7	0.232 3	0.000 0	0.996 7
控制变量	<i>ROA</i>	18 317	0.032 0	0.080 8	-1.647 9	0.785 9
	<i>Tobin</i>	18 317	2.212 7	1.488 1	0.670 4	27.338 0
	<i>Top1</i>	18 317	31.145 0	13.945 1	1.840 0	89.090 0
	<i>Size</i>	18 317	1.12×10 ¹⁰	3.63×10 ¹⁰	9.74×10 ⁷	1.01×10 ¹²
	<i>Growthrate</i>	18 317	0.177 3	1.001 7	-0.971 3	58.841 6
	<i>Age</i>	18 317	11.192 6	6.974 6	3	34
	<i>Cashflow</i>	1831 7	0.047 7	0.067 1	-0.658 1	0.838 5
	<i>Opinion</i>	18 317	0.972 0	0.164 8	0	1
	<i>Sub</i>	18 317	0.013 0	0.124 1	-0.009 3	16.49 5
	<i>IC</i>	18 317	0.013 9	0.005 8	0.000 0	0.079 8
	<i>Pop</i>	18 317	1 124.491 0	717.642 3	24.840 0	3 213.340 0
	<i>Consume</i>	18 317	5.24×10 ⁷	4.53×10 ⁷	3.34×10 ⁵	1.85×10 ⁸
	<i>Fin</i>	18 317	2.82×10 ⁸	2.86×10 ⁸	2.00×10 ⁶	1.09×10 ⁹

五、实证结果与分析

(一) 基准回归

表 3 报告了数据要素流动对企业创新持续性影响的基准回归结果。其中,数据要素流动的回归系数均显著为正。从经济含义看,数据要素流动是基于城市间数据要素潜在流动强度和网络结构位置构建的综合指数,其数值越高,说明城市在数据要素跨区域连接、获取和传递网络中的中心性越强。数据要素流动水平提升后,企业更可能获得外部市场需求、技术趋势、产业链运行和政策环境等信息增量,从而改善创新决策质量、缓解资源匹配约束并拓展协同创新机会,最终表现为研发投入的连续性增强和创新成果的持续形成。假设 1 得到验证。

表 3 基准回归结果

变量	lnIIP	lnOIP
<i>Df</i>	0.329 3 ** (0.144 2)	0.655 1 *** (0.221 1)
<i>ROA</i>	0.142 5 (0.100 8)	0.130 4 (0.124 1)
<i>Tobin</i>	0.027 5 *** (0.006 5)	0.019 6 ** (0.009 5)
<i>Top1</i>	-0.002 1 (0.001 9)	0.000 3 (0.002 4)
<i>lnSize</i>	0.916 8 *** (0.029 9)	0.663 2 *** (0.037 1)
<i>Growthrate</i>	0.045 1 *** (0.011 0)	0.018 4 ** (0.008 1)
<i>lnAge</i>	-0.308 6 *** (0.053 3)	-0.231 2 *** (0.068 3)
<i>Cashflow</i>	0.479 6 *** (0.113 5)	-0.297 3 ** (0.138 5)
<i>Opinion</i>	-0.029 4 (0.047 7)	0.172 6 *** (0.058 3)
<i>Sub</i>	-0.005 0 (0.011 9)	-0.011 1 (0.018 9)
<i>IC</i>	16.150 3 *** (2.142 3)	14.962 8 *** (2.811 4)
<i>lnPop</i>	0.012 0 (0.143 4)	0.123 4 (0.138 5)
<i>Consume</i>	-0.089 2 (0.057 4)	-0.177 2 ** (0.079 4)
<i>Fin</i>	0.004 8 (0.087 7)	-0.041 7 (0.082 8)
常数项	-8.763 1 *** (1.346 2)	-7.386 2 *** (1.351 8)
企业固定效应	控制	控制
时间固定效应	控制	控制
观测值	18 317	18 317
$\overline{R^2}$	0.844 8	0.693 9

注：*、**、*** 分别表示在 10%、5% 和 1% 水平下显著，小括号内为稳健标准误。后表同。

(二) 内生性分析

内生性问题可能导致研究结果出现偏误。一方面,数据要素的流动在影响企业创新持续性的同时,企业的创新行为也会产生大量数据,并带动相关数据的传输与流通,从而对数据要素的流动产生影响。另一方面,影响企业创新行为持续性的因素较多,模型可能存在遗漏变量、样本选择偏误等问题。为此,本文采用工具变量法和倾向得分匹配法缓解可能的内生性问题。

1. 工具变量法

本文构造基于历史通信基础设施的工具变量 (*IV*),以缓解潜在的内生性偏误。具体而言,采用城市 1984 年每百人固定电话数量与全国城市-年份平均数据要素流动水平的交互项作为核心工具变量。其中,平均数据要素流动水平采用留一法计算,即剔除本城市后其他城市在年份 *t* 的流动水平均值,以保证冲击项不受本城市当期经济活动的干扰。相关性方面,其逻辑建立在历史路径依赖与宏观技术扩散的双重机制之上。一方面,历史通信基础较好的城市在网络建设、信息服务业与数字化人才储备上具备先发优势,路径依赖效应使其更易吸纳数字技术,助推本地数据要素有序流动。另一方面,全国平均流动水平反映数据要素发展的宏观趋势,可通过技术示范、政策引导与市场传导作用于各城市,历史通信基础扎实的城市能更快响应宏观趋势,提升本地数据要素流动水平。外生性方面,其一,1984 年固定电话普及率为历史变量,其对样本期企业创新持续性的直接影响随时间大幅弱化,主要通过通信基础传承、数字技术吸收能力作用于当期数据要素流动;其二,模型已控制企业与城市层面多维度特征,并纳入企业、年份双向固定效应,可排除历史通信基础通过经济、产业、人才等渠道直接影响企业创新的可能;其三,留一法构

造的全国均值变异源自全国共性趋势,与本城市政策及企业行为无关。综上,该交互项主要通过本地数据要素流动渠道作用于企业创新持续性,满足排他性要求。工具变量法具体回归结果如表 4 所示,核心解释变量的回归系数依然显著为正,Kleibergen-Paap rk LM 统计值在 1%水平下拒绝工具变量不可识别的原假设,Kleibergen-Paap rk Wald 统计 F 值大于 Stock-Yogo 在 10%水平下的临界值,表明弱工具变量问题不严重。上述结果表明,在利用工具变量法缓解内生性问题后,基准回归结果可靠。

2. 倾向得分匹配法

为缓解样本选择偏误带来的内生性问题,本文采用倾向得分匹配法进行处理,将数据要素流动水平高于均值的设置为处理组,其余设置为对照组,以基准回归中所有控制变量作为匹配变量,使用 1 : 1 近邻匹配为处理组寻找特征相似的对照组。倾向得分匹配法回归结果如表 4 所示,核心解释变量的回归系数依然显著为正,这说明在进一步缓解样本选择偏误后,本文基准回归结果可靠。

表 4 内生性分析回归结果

变量	工具变量法			倾向得分匹配法	
	Df	lnIIP	lnOIP	lnIIP	lnOIP
IV	1. 677 6 *** (0. 103 4)				
Df		1. 311 8 ** (0. 570 6)	3. 872 4 *** (0. 775 1)	0. 415 0 ** (0. 189 2)	0. 917 8 *** (0. 284 2)
常数项				-8. 128 4 *** (1. 802 9)	-6. 430 1 *** (1. 698 3)
Kleibergen-Paap rk LM	82. 009 *** [0. 000]				
Kleibergen-Paap rk Wald F	263. 184 { 16. 38 }				
控制变量		控制	控制	控制	控制
企业固定效应		控制	控制	控制	控制
时间固定效应		控制	控制	控制	控制
观测值		17 258	17 258	9 992	9 992
R ²		0. 545	0. 190	0. 841 3	0. 691 7

注:中括号内为不可识别检验 P 值,大括号内为 Stock-Yogo 弱识别检验在 10%水平下的临界值。

(三) 稳健性检验

为验证回归结果的稳健性,本文进行稳健性检验。一是调整研究样本。新冠疫情对各个城市的运行与企业的经营发展都产生了重大影响,据此,本文剔除了 2020—2022 年的样本,重新进行回归。二是替换核心解释变量,利用主成分分析法代替熵值法重新计算数据要素流动的指标,并将其作为核心解释变量重新进行回归。三是替换被解释变量。本文基于中国证监会公布的上市公司两位行业代码,以第 t-1 年行业研发投入和专利申请数量的中位数作为企业第 t 年的创新投入和创新产出预期目标,以企业连续达到或超过行业基准的年份数衡量企业创新投入持续性和创新产出持续性。若企业创新投入和创新产出达到或超过行业预期赋值为 1,若连续两年达到预期目标赋值为 2,依此类推。若未达到行业预期则赋值为 0。企业连续达到行业预期目标的年份越多,其创新持续性越强。四是控制高维固定效应。考虑到不同城市的经济基

础、发展目标、政府行为等因素会对当地企业的持续创新活动产生影响,本文在基准回归模型的基础上,进一步控制了城市固定效应。五是聚类层级从企业个体改变为行业。具体回归结果如表 5 所示,上述稳健性检验的主要结论与基准回归保持一致,表明本文基准回归结果稳健^①。

表 5 稳健性检验回归结果

变量	调整样本		替换解释变量		替换被解释变量	
	lnIIP	lnOIP	lnIIP	lnOIP	lnIIP _w	lnOIP _w
<i>Df</i>	0.293 9 *	0.551 8 **			0.342 6 ***	0.613 9 ***
	(0.154 4)	(0.224 6)			(0.117 6)	(0.119 4)
<i>Df</i> ²			0.052 8 ***	0.174 9 ***		
			(0.019 9)	(0.035 4)		
常数项	-9.009 9 ***	-8.350 1 ***	-8.737 5 ***	-6.651 5 ***	-9.490 0 ***	-7.097 3 ***
	(1.364 9)	(1.471 0)	(1.330 1)	(1.362 5)	(0.764 5)	(0.769 7)
控制变量	控制	控制	控制	控制	控制	控制
企业固定效应	控制	控制	控制	控制	控制	控制
时间固定效应	控制	控制	控制	控制	控制	控制
观测值	11 781	11 781	18 317	18 317	18 317	18 317
$\overline{R^2}$	0.809 6	0.672 1	0.844 8	0.694 8	0.726 3	0.619 3

(四) 机制检验

为进一步检验数据要素流动对企业创新投入持续性和创新产出持续性的作用机制,本文在基准回归模型的基础上,设定如下模型:

$$M_{it} = \beta_0 + \beta_1 Df_{it} + X'_{it} \varphi + \theta_i + \mu_t + \varepsilon_{it}$$

(7)

其中, M_{it} 表示机制变量,其余变量与前文一致。

1. 融资约束缓解机制

前文理论分析指出,数据要素流动通过缓解信息不对称、增强信号传递提升企业的融资能力。一方面,数据要素流动的增强有助于弥合企业和融资机构之间的信息鸿沟,提高企业融资可得性,进而为企业创新持续性提供资金保障;另一方面,数据要素流动还能够帮助企业向外界传递积极的信号,使得企业更容易获得相关政策扶持和投资者的关注,实现融资渠道的多样化,为强化企业创新持续性提供资金保障。本文参考方慧等^[32]的研究,采用 WW 指数衡量企业的融资约束程度,该指标越大,企业面临的融资约束就越强。表 6 第一列回归结果表明,数据要素流动可以降低企业的融资约束,进而提升企业创新持续性。假设 2 得到验证。

2. 市场竞争促进机制

根据前文的理论分析,随着数据要素流动不断增强,企业能够快速获取市场需求、行业技术动态等关键信息,这不仅能够帮助企业把握市场风向,提高战略决策的准确性,而且有助于降低市场信息壁垒和潜在进入者的信息获取成本,从而提高市场竞争程度。企业为了在竞争愈发激烈的市场上站稳脚跟,就必须持续

① 限于篇幅,稳健性检验的“控制高维固定效应”和“改变聚类层级”两部分回归结果省略,留存备索。

不断地进行创新,形成差异化优势,从而达到抢占市场份额或巩固市场地位的目的。本文采用行业集中度衡量市场竞争程度,即行业前五企业的营业收入占全行业营业收入的比重。该指标是负向指标,行业集中度越高,意味着市场竞争程度越低。表 6 第二列回归结果表明,数据要素流动有助于降低行业集中度、增强市场竞争程度,进而提升企业创新持续性。假设 3 得到验证。

3. 不确定性感知降低机制

根据前文的理论分析,数据要素的高效流动能够帮助企业更加及时地获取市场变化、技术走向,从而更精准地发现创新机遇并规避风险。通过数据支持的预测分析,企业能优化研发选择,降低试错成本,促进创新成果形成与转化。同时,跨领域数据整合带来的知识共享效应,可助力企业灵活调整创新策略,形成监测、学习、迭代的良性循环。当外部风险变得可控、路径清晰可追溯时,企业对创新收益更有信心,更愿意开展持续性创新活动。本文参考方明月等^[21]的研究,采用文本分析法构建企业的不确定性感知指数,该指标越大,企业对不确定性的敏感度越强。表 6 第三列回归结果表明,数据要素流动可以通过降低企业的不确定性感知提升企业的创新持续性。假设 4 得到验证。

4. 合作意识增强机制

根据前文的理论分析,数据要素流动增强有助于打破主体间的“信息孤岛”,为企业提供了多样的交流机会,有助于企业更好地了解潜在合作主体的需求与优势,从而找到更多的合作契合点,在共同挖掘数据价值中增进互信,提升合作意愿。企业之间的互利合作能够汇聚不同专业人才与技术,通过协同合作实现优势互补,激发创新灵感,还能营造良好氛围,鼓励冒险和探索,持续推动企业创新投入的增加和创新成果的转化。本文从 CNRDS 数据库获取企业合作意识数据作为企业创新产出持续性的机制变量,该指标越大,企业的合作意识就越强。表 6 第四列回归结果表明,数据要素流动能够促进企业合作意识的提升,进而提升企业创新持续性。假设 5 得到验证。

表 6 机制检验回归结果

变量	资源保障	竞争激励	信息决策	协同创新
<i>Df</i>	-0.011 5 ** (0.005 4)	-0.001 1 * (0.000 6)	-0.071 5 *** (0.025 5)	0.183 5 ** (0.087 0)
常数项	-0.000 01 (0.032 3)	0.007 4 (0.005 2)	0.270 6 * (0.145 5)	0.485 0 (0.576 1)
控制变量	控制	控制	控制	控制
企业固定效应	控制	控制	控制	控制
时间固定效应	控制	控制	控制	控制
观测值	15 637	18 317	18 225	13 000
$\overline{R^2}$	0.828 7	0.737 0	0.415 0	0.723 1

(五) 异质性分析

数据要素流动对企业创新持续性的影响可能因企业特征和城市环境不同而存在差异。由于本文的数据要素流动基于城市间数据联系测度,而创新持续性体现于企业微观层面,因此有必要从企业内部特征与城市外部环境两个层面考察其作用边界。企业层面选取发展潜力,主要反映企业成长预期、资源扩张需求和长期创新动力;城市层面选取人才引进强度,主要反映专业人才供给与创新要素集聚条件。前者影响企

业对外部数据资源的吸收与利用,后者则为数据处理、知识吸收和技术应用提供人力资本支撑。基于此,本文分别从企业发展潜力和城市人才引进强度两个维度开展异质性分析。

1. 企业层面:企业发展潜力

本文参考曾国安等^[33]的方法,采用企业的市账比率作为企业发展潜力的划分指标(Bm),当企业市账比率大于行业-年度中位数时,将其划分为高潜力企业,记为1,反之则划分为低潜力企业,记为0,并在此基础上构建交互项($Bm \times Df$)。表7前两列回归结果显示,交互项的回归系数均显著为正,说明在发展潜力较大的企业中,数据要素流动对企业创新投入持续性和创新产出持续性的促进作用更强。可能的原因在于,高发展潜力企业通常具有更强的成长预期、创新意愿和资源整合能力,对市场信息、技术机会及外部数据资源具有更强的识别和利用能力,因而能够更有效地识别创新方向、优化研发资源配置,并将数据流动带来的信息增量转化为持续研发投入和创新成果。相比之下,低发展潜力企业受战略取向、资源约束及吸收能力等因素影响,对外部数据资源的利用和转化能力相对有限,数据要素流动的创新赋能效应因而较弱。

2. 地区层面:人才引进强度

本文参考陈曦等^[34]的研究,利用各地市政府工作报告,采用文本分析法测度地级市人才引进强度(HR),按照中位数将样本划分为高强度地区(记为1)和低强度地区(记为0),并进一步构建与核心解释变量的交互项($HR \times Df$)。表7后两列回归结果显示,交互项的回归系数均显著为正,说明数据要素流动对企业创新持续性的作用,在人才引进强度较高的地区更强。可能的原因在于,高素质人才具备较强的数据处理、知识吸收和技术应用能力,能够帮助企业将数据流动带来的信息增量更有效地嵌入研发决策与创新过程,提高创新资源配置效率和成果转化能力。同时,人才集聚产生的知识溢出与创新网络效应,也有助于提升企业对外部数据资源的吸收和利用能力,从而强化数据要素流动对持续创新的赋能作用。

表7 异质性分析回归结果

变量	企业发展潜力异质性		人才引进强度异质性	
	$\ln IIP$	$\ln OIP$	$\ln IIP$	$\ln OIP$
Df	0.278 0 [*] (0.144 4)	0.591 6 ^{***} (0.224 6)	0.309 4 ^{**} (0.145 0)	0.520 1 ^{**} (0.215 4)
$Bm \times Df$	0.096 0 ^{**} (0.045 0)	0.118 9 [*] (0.070 1)		
$HR \times Df$			0.088 3 ^{**} (0.035 3)	0.598 5 ^{***} (0.068 0)
常数项	-8.773 3 ^{***} (1.346 2)	-7.398 7 ^{***} (1.351 7)	-8.651 4 ^{***} (1.349 1)	-6.628 8 ^{***} (1.338 5)
控制变量	控制	控制	控制	控制
企业固定效应	控制	控制	控制	控制
时间固定效应	控制	控制	控制	控制

表7(续)

变量	企业发展潜力异质性		人才引进强度异质性	
	lnIIP	lnOIP	lnIIP	lnOIP
观测值	18 317	18 317	18 317	18 317
$\overline{R^2}$	0. 844 8	0. 694 0	0. 844 8	0. 696 3

六、结论与建议

本文基于 2011—2023 年沪深 A 股上市企业数据和地级及以上城市数据,从投入-产出视角实证检验数据要素流动对企业创新持续性的影响。研究结果显示,数据要素流动能够提升企业创新投入持续性和创新产出持续性,且经过内生性处理和一系列稳健性检验后结论依然成立。机制分析表明,数据要素流动以降低信息不对称、提升信息完备性为基础,通过资源保障、竞争激励、信息决策以及协同创新四条路径影响企业创新持续性。此外,数据要素流动对企业创新持续性的影响具有明显异质性。在企业层面,其促进作用在高发展潜力企业中更强,表明较强的成长预期、创新意愿和资源整合能力有助于企业将数据流动带来的信息增量转化为持续创新动能;在城市层面,其促进作用在人才引进强度较高地区更强,说明较强的人才引进有助于改善专业人才供给,为数据处理、知识吸收与技术转化提供人力资本支撑。总体而言,数据要素流动创新效应的释放既取决于企业自身的吸收转化能力,也受到城市人才吸引与集聚条件的影响。

基于上述结论,本文提出如下政策建议:

第一,构建数据要素流动支撑体系,夯实企业创新基础。在基础设施层面,以国家级数据交易中心为核心节点,建设重点行业数据交易平台,完善数据要素市场体系。要素供给方面,有序推动科技研发、市场监管、知识产权等重点领域政务数据开放共享。在行业层面,借鉴欧盟数据空间建设经验,在新能源汽车、人工智能等战略性新兴产业和重点领域组建数据共享联盟。应用支持方面,设立专项基金补贴中小企业云服务与数据治理支出,并将数据要素管理能力纳入高新技术企业认定标准。

第二,完善创新生态系统,激活企业持续创新链式反应。融资支持方面,开发基于数据流水的供应链金融产品,探索数据资产质押融资及价值评估机制。市场竞争维度方面,建立数据垄断预警系统,动态监测平台企业数据集中度,强制垄断企业开放非核心数据接口以破除市场壁垒。协同创新方面,将数字技能纳入研发人员继续教育体系,支持龙头企业建设跨界创新实验室,对产学研数据创新联合体给予研发费用加计扣除优惠,以“技能提升-平台搭建-政策激励”组合培育协同创新生态。

第三,实施差异化扶持策略,提升政策资源配置效率。在企业层面,应根据企业发展潜力实施分类支持。对于发展潜力较高、创新意愿和资源整合能力较强的企业,可优先加强数据资源对接、应用场景开放和研发金融支持,引导其将数据要素流动带来的信息与知识优势进一步转化为持续创新成果;同时,对发展潜力相对较低的企业,应通过数据应用指导、创新资源对接等方式提升其数据吸收与转化能力,避免数据要素创新效应过度集中于优势企业。在城市层面,应强化人才对数据要素价值转化的支撑作用,完善数字技术、数据分析及研发创新人才的引进、培养与激励机制,推动高校、科研机构与企业之间的人才交流和知识共享,为企业利用流动数据识别创新机会、优化研发决策和促进成果转化提供人才保障。

参考文献:

- [1] 杨艳,王理,廖祖君. 数据要素:倍增效应与人均产出影响:基于数据要素流动环境的视角[J]. 经济问题探索,2021(12):118-135.
- [2] 刘晶,杨怡帆. 数据要素流动对区域协调发展的影响机理研究[J]. 管理学报,2025,38(3):133-148.
- [3] 高明,魏浩,王晓祺. 数据要素流动赋能企业绿色创新[J]. 中国人口·资源与环境,2024,34(11):120-129.
- [4] 梁锦凯,陈关聚,娄小亭. 数据要素市场化配置促进了企业创新“增量提质”吗[J]. 江西财经大学学报,2024(6):36-49.
- [5] Suárez D. Persistence of innovation in unstable environments: continuity and change in the firm's innovation behavior[J]. Research Policy, 2014, 43(4):726-736.
- [6] Howell S T. Financing innovation: evidence from R&D grants[J]. American Economic Review, 2017, 107(4):1136-1164.
- [7] Triguero Á, Córcoles D. Understanding innovation: an analysis of persistence for Spanish manufacturing firms[J]. Research Policy, 2013, 42(2):340-352.
- [8] 鞠晓生,卢荻,虞义华. 融资约束、营运资本管理与企业创新可持续性[J]. 经济研究,2013,48(1):4-16.
- [9] 杨仁发,杨梅君. 数字化转型的持续性创新效应研究[J]. 数量经济技术经济研究,2025,42(2):109-129.
- [10] 姚正海,王强. ESG表现对企业持续性创新的影响机制研究[J]. 软科学,2025,39(6):99-108.
- [11] 孙冰,谭钰皎,杨雪婷. 企业规模与创新持续性:基于“坚柔并济”的双调节作用[J]. 软科学,2024,38(9):101-107.
- [12] 齐念念,赵炎. 网络凝聚性、技术市场价值与企业持续性创新[J]. 管理科学,2024,37(5):21-33.
- [13] 樊霞,陈娅,张巧玲. 经济政策不确定性、政府隐性担保与企业创新持续性[J]. 管理学报,2020,17(9):1347-1354.
- [14] 杨晓璇,杨奇松. 政府风险投资、财政治理与中小企业创新持续性:新公共服务理论解读视角[J]. 现代财经(天津财经大学学报),2024,44(12):68-86.
- [15] 董旗,谭伟杰,谢家平,等. 政府数据开放与企业新质生产力发展[J]. 经济与管理研究,2025,46(3):3-23.
- [16] 黄先海,高亚兴. 数字知识存量、数实技术融合与企业实体技术创新边界[J]. 经济学动态,2025(3):36-53.
- [17] 刘青,黄荣荣,肖柏高. 超大规模市场优势与企业创新战略:基于出口引致需求的视角[J]. 数量经济技术经济研究,2025,42(5):48-70.
- [18] Hoenig D, Henkel J. Quality signals?: The role of patents, alliances, and team experience in venture capital financing[J]. Research Policy, 2015, 44(5):1049-1064.
- [19] 韩忠雪,袁翊菲. 数据资产信息披露与企业创新韧性[J]. 首都经济贸易大学学报,2026,28(3):88-102.
- [20] 张昊,张新博,樊步青. 服务化推动先进制造业 GVC 升级的作用机制研究:基于技术创新视角[J]. 中国软科学,2024(7):179-190.
- [21] 方明月,聂辉华,阮睿,等. 企业数字化转型与经济政策不确定性感知[J]. 金融研究,2023(2):21-39.
- [22] Bloom N. Fluctuations in uncertainty[J]. Journal of Economic Perspectives, 2014, 28(2):153-176.
- [23] 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.
- [24] 黄勃,李海彤,江萍,等. 战略联盟、要素流动与企业全要素生产率提升[J]. 管理世界,2022,38(10):195-212.
- [25] de Faria P, Lima F, Santos R. Cooperation in innovation activities: the importance of partners[J]. Research Policy, 2010, 39(8):1082-1092.
- [26] 李天籽,谷雨,李振涛. 高铁网络节点优势与企业生产率:基于复杂网络分析方法[J]. 产业经济研究,2023(4):99-113.
- [27] 李昊骅,方立兵,姚楚涵. 空间溢出与城投债信用风险:基于长三角城市群生产要素引力网络[J]. 管理科学学报,2023,26(1):83-104.
- [28] 孙晋云,白俊红,王钺. 数字经济如何重塑我国区域创新格局?:基于研发要素流动的视角[J]. 统计研究,2023,40(8):59-70.
- [29] 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.
- [30] 杨亚平,杨鑫,陈凯. 数据要素市场化对企业数字化创新的影响[J]. 经济管理,2025,47(3):5-25.
- [31] 王宸威,钞小静. 数据要素应用如何影响企业分布式创新?:来自上市公司微观视角下的经验证据[J]. 现代财经(天津财经大学学报),2024,44(11):3-21.
- [32] 方慧,姜春宇,李小明. 数字技术创新与产业链韧性提升:基于节点企业视角的考察[J]. 经济学动态,2025(2):127-144.
- [33] 曾国安,苏诗琴,彭爽. 企业杠杆行为与技术创新[J]. 中国工业经济,2023(8):155-173.
- [34] 陈曦,吴英巨,朱建华. 新质生产力视角下地方人才引进与全要素生产率[J]. 经济管理,2024,46(12):104-120.

How Does Data Factor Flow Affect Firm Innovation Persistence?

SHAN Haonan¹, ZHAO Kai²

(1. Qingdao University of Technology, Qingdao 266520;

2. Qingdao University, Qingdao 266071)

Abstract: Against the backdrop of the ongoing Fourth Industrial Revolution, characterized by digitalization, intelligentization, and networking, technological iteration is accelerating and market demand is becoming increasingly dynamic. Firms are therefore facing intensifying internal and external uncertainty, while the competitive advantage generated by a single innovation is steadily diminishing. Relying solely on isolated innovation is no longer sufficient to sustain long-term corporate development. Consequently, innovation persistence has become essential for firms to escape innovation traps and maintain long-term competitiveness. In this context, the low-cost, reusable nature of data factors provides a new perspective for addressing the challenge of sustaining corporate innovation.

Using data from A-share listed companies and prefecture-level and above cities in China from 2011 to 2023, this paper empirically examines the impact of data factor flow on firm innovation persistence. The results show that data factor flow enhances both innovation input and output persistence. By reducing information asymmetry and improving information completeness, data factor flow affects firm innovation persistence through four channels: resource support, competitive incentives, information-based decision-making, and collaborative innovation. Specifically, the resource support channel operates by alleviating financing constraints, the competitive incentive channel aligns with intensified market competition, the information-based decision-making channel presents as reduced uncertainty perception, and the collaborative innovation channel reflects enhanced cooperation awareness. Furthermore, the positive effect of data factor flow is stronger for firms with higher development potential and for firms located in cities with greater talent attraction. These findings suggest that firms' internal growth potential and cities' external talent support are important conditions for transforming data factor flow into sustained innovation momentum.

The marginal contributions are threefold. First, in terms of measurement, this paper combines a modified gravity model with complex network analysis to construct a multidimensional indicator of data factor flow that captures both flow intensity and network characteristics. This approach reduces reliance on a single proxy variable and offers an alternative method for quantitatively measuring cross-regional data factor flow. Second, in terms of research perspective, this paper integrates data factor flow and firm innovation persistence into a unified analytical framework. Unlike existing studies, which tend to focus on the static innovation effects of data factors, it extends the economic effects of data factors and firm innovation, paying sufficient attention to their cross-regional flow characteristics and to innovation persistence. Third, in terms of analytical dimensions, this paper classifies innovation persistence into innovation input persistence and innovation output persistence and further examines their differentiated transmission mechanisms. This distinction provides a more detailed understanding of how data factor flow supports sustained innovation and offers more specific empirical evidence on the micro-level channels through which data factors affect corporate innovation activities.

Keywords: data factor flow; firm innovation persistence; financing constraints; market competition; uncertainty perception; cooperation awareness

编校:周 斌;王冬梅