

# 数据要素与绿色创新: 基于新质生产力视角

王柯丹<sup>1, 2</sup>, 刘颖<sup>2</sup>, 汪寿阳<sup>2, 3</sup>

(1. 大连民族大学 国际商学院, 辽宁 大连 116600; 2. 中国科学院大学 经济与管理学院, 北京 100190;

3. 中国科学院 数学与系统科学研究院, 北京 100190)

**摘要:** 数据要素是形成新质生产力的动力引擎, 而绿色发展是衡量新质生产力发展水平的重要标尺, 探究数字化与绿色化之间的内在逻辑成为当前备受关注的重要议题。本文利用2010—2022年中国A股制造业上市企业数据, 实证研究了在新质生产力背景下, 数据要素对企业绿色创新质量的影响机制与实现条件。研究发现, 数据要素对企业绿色创新质量的提升具有积极影响; 数据要素的绿色赋能作用在国有企业、高新技术企业、资本与技术密集型企业及低金融约束程度的企业中更明显; 机制分析表明, 数据要素能够通过提高企业创新投入和帮助企业获得政府支持的方式提升企业绿色创新质量; 大数据技术水平和人力资本质量是数据要素驱动企业绿色创新的有效条件。本文的研究结论有助于实现新质生产力的数字化与绿色化“双化协同”, 也为推动经济高质量发展提供有益借鉴。

**关键词:** 新质生产力; 数据要素; 企业绿色创新质量; 绿色发展

**中图分类号:** F062.5 **文献标识码:** A **文章编号:** 1000-176X(2024)09-0018-16

## 一、引言

数字化与绿色化作为新质生产力最为鲜明的时代特征, 不仅引领着全球科技革命与产业转型的发展趋势, 更在重塑生产力发展蓝图和全球竞争格局中占据核心地位<sup>[1]</sup>。习近平总书记强调要“加速数字化绿色化协同发展”“数据是新的生产要素, 是基础性资源和战略性资源, 也是重要生产力”“绿色发展是高质量发展的底色, 新质生产力本身就是绿色生产力”。

数据要素作为新质生产力的引擎, 能有效实现科技创新、促进产业深度转型并推动生产要素的创新性配置<sup>[2]</sup>。绿色发展作为高质量发展的内在要求, 强调在经济增长过程中注重生态环境保护与资源高效利用, 能够实现社会经济发展与自然环境和谐共生<sup>[3]</sup>, 这不仅关乎人类社会的长远福祉, 而且是衡量新质生产力发展水平的关键标尺。因此, 探讨新质生产力背景下数字化与

**收稿日期:** 2024-07-29

**基金项目:** 国家自然科学基金基础科学中心项目“计量建模与经济政策研究”(71988101); 国家自然科学基金面上项目“基于福利效应评价的算法治理研究”(72272140); 国家自然科学基金重点项目“新型城镇化与区域协调发展的机制与治理体系研究”(72334006)

**作者简介:** 王柯丹(1997-), 女, 辽宁大连人, 讲师, 博士, 主要从事数字经济研究。E-mail: kedanwang1024@163.com  
刘颖(通讯作者)(1983-), 男, 河北唐山人, 教授, 博士, 主要从事数字经济、金融科技研究。E-mail: liuy@ucas.ac.cn  
汪寿阳(1958-), 男, 江苏盐城人, 教授, 博士, 主要从事决策理论、经济分析与预测等研究。E-mail: sywang@amss.ac.cn

绿色化之间的关系，尤其是数据要素这一核心要素在促进绿色发展中的作用机制，成为一个亟待研究的重要议题。

基于新古典经济学理论，数据作为一种生产要素，能够有效提升企业的生产效率和创新能力，进而推动产业结构的优化和升级<sup>[4]</sup>。同时，数据要素所展现的低复制成本、低流通摩擦和乘数效应<sup>[5]</sup>等新特征，使其成功突破了传统经济增长模式的限制<sup>[6]</sup>，成为推动绿色发展的关键驱动力<sup>[7, 8]</sup>。具体而言，数据要素通过提供透明、准确的信息，有效降低了企业与政府、市场之间的信息不对称程度<sup>[9]</sup>，增强了企业的市场洞察力和决策准确性。同时，数据要素还赋能传统产业智能化升级<sup>[10]</sup>，促进了绿色新兴产业的蓬勃兴起，推动了绿色技术的研发与应用，引领了绿色消费趋势的形成，并有效提升了资源利用效率，全方位地加速了社会经济向绿色可持续发展路径转型<sup>[11]</sup>。然而，已有研究多停留于理论阐释与效用构建层面，鲜有基于实证数据的深入剖析，特别是从生产力理论视角出发，对数据要素的绿色赋能的具体路径、条件及其效果的全面探讨尚显不足。同时，已有研究对数据要素与绿色发展之间的赋能关系和作用机制还未形成一致观点，亦有一些学者从污染转移和数据要素间接引致的能源压力等方面展开讨论，指出数字要素对绿色经济发展存在一定的负面效应<sup>[12]</sup>。基于以上考虑，本文聚焦于企业这一微观主体，旨在运用实证分析方法，深入探索新质生产力背景下，数据要素对企业绿色创新的赋能作用、内在机理及实现条件。

本文的边际贡献有以下三点：第一，深化了新质生产力理论，拓展了绿色发展理论边界。在现有新质生产力的理论研究基础上，本文聚焦于绿色发展这一关键维度，深入探索了数字化与绿色化融合的新路径。通过将数据要素作为新质生产力的核心要素融入绿色经济分析框架，揭示了新质生产力在绿色创新与经济转型中的作用机制。第二，紧密结合中国产业动态和社会经济背景，对新质生产力的绿色赋能作用提供了微观证据。现有新质生产力领域的研究主要停留在理论阐释与效用构建层面，本文使用企业微观数据，实证检验了数据要素对绿色创新质量的积极影响，验证了其对于国有企业、高新技术企业等特定类型企业的显著效果，为新质生产力的绿色赋能效应提供了坚实的实证支撑，并为政策制定提供了基于企业特性的精准施策参考。第三，结合生产力理论，从生产力及生产关系的视角剖析了绿色赋能作用机制与实现条件。本文验证了创新投入和政府补贴是数据要素赋能绿色创新的重要作用机制，大数据技术水平与人力资本质量是实现赋能作用的关键条件，这一研究结论丰富了新质生产力与绿色发展的理论内涵，为实践中数据要素与绿色创新的深度融合提供了新的视角。

## 二、理论分析与研究假设

新质生产力是科技创新与生产力发展的深度融合，它代表了先进生产力的质态跃升。新质生产力的核心在于“新”与“质”两个层面。其中，“新”指的是区别于传统生产力的新型生产力，它依托于新技术、新经济、新业态的深度融合与发展，展现出独特的内涵与优势<sup>[13]</sup>；在追求生产力的显著增长与提升过程中，“质”的维度凸显了创新驱动的核心地位，彰显了对创新和自主发展的深刻理解和坚定追求<sup>[14]</sup>。

数据要素在推动新质生产力的形成过程中发挥着举足轻重的作用。数据要素不仅能够直接促进经济发展方式的转变和产业结构的优化升级，更可以与传统生产要素深度交融，释放出强大的乘数效应<sup>[5]</sup>。数据要素所具备的正外部性和低复制成本的特性，使其在推动科技创新、优化资源配置及引领产业结构升级<sup>[3]</sup>等方面起到了决定性作用。数据要素的广泛应用能够有效地培育出更高素质的新质劳动力，创造出具备更高技术含量的新质劳动资料，培育出更广泛领域的新质劳动对象。这些新质要素的涌现，不仅将引发生产力的巨大跃升，更将有力推动新质生产力的形成，为社会经济发展注入新的活力和动力。

值得一提的是, 新质生产力摆脱了传统经济的增长方式和发展路径, 在促进绿色化方面发挥着至关重要的作用<sup>[15, 16]</sup>。新质生产力通过不断促进传统产业的智能化改造并催生新的产业形态, 能有效推动绿色创新和研发, 促进绿色产业的发展, 推动绿色消费和绿色生活的普及, 提升资源利用效率, 促进社会经济绿色、可持续发展。

绿色创新作为应对环境挑战和实现可持续发展的关键路径, 要求企业在技术创新的同时, 注重环境保护和资源的高效利用。依据技术创新与产业升级理论, 数据要素在推动绿色创新的过程中起到重要作用。通过数据驱动的技术革新与产业转型, 企业能够实现生产方式的绿色化、低碳化, 提升资源利用效率, 减少环境污染。通过深度挖掘和分析生产过程中的数据, 企业能够精准识别环保技术的研发方向, 洞察绿色市场的需求变化, 探索潜在的合作机会。这不仅有助于企业针对性地开展绿色技术研发, 提升企业的绿色创新质量, 还能够促进绿色技术的快速应用和市场推广, 从而增强企业在绿色领域的竞争力和影响力。因此, 本文提出以下假设:

**H1:** 数据要素对企业提升绿色创新质量具有显著的积极影响。

在新质生产力背景下, 生产要素创新性配置已成为推动经济增长与转型升级的核心动力。根据生产力理论, 生产力的提升依赖于生产要素的优化配置与技术创新。在新质生产力背景下, 数据要素作为一种全新的生产要素, 不仅丰富了生产资料的内涵, 还改变了传统生产要素的组合方式。数据要素的加入, 使得企业能够更高效地利用现有资源, 通过数据挖掘与分析, 发现潜在的创新机会, 从而增加对创新活动的投入。这种创新性的生产要素配置, 为企业实现绿色创新奠定了坚实的基础。

数据要素能够与其他生产要素深度融合, 产生强大的协同效应<sup>[17]</sup>, 形成新的生产力增长点。通过数据挖掘与分析, 企业可以洞察市场趋势、优化生产流程、提升产品质量, 进而推动技术创新与产业升级。同时, 数据要素的边际效用递增特性为企业创新投入提供了持续的动力。随着数据量的增加和数据分析技术的成熟, 企业能够不断从数据中挖掘新的价值点, 推动创新活动的深入发展。在这一过程中, 企业不仅能够更加精准地把握市场需求、生产状况与资源分布的全貌, 还能通过数据分析的精准指导, 实现创新资源的优化配置与高效利用, 确保创新投入精准聚焦于关键领域与核心环节。

尤为重要的是, 数据要素在推动企业向绿色、低碳、可持续发展方面发挥着不可替代的作用。通过数据分析, 企业能够清晰地识别出运营过程中的环保瓶颈与绿色创新潜力, 进而制定针对性的改进策略与行动计划。这种基于数据的创新决策过程不仅提升了企业的创新投入效率, 还确保了创新投入的绿色属性与可持续发展价值。因此, 本文提出以下假设:

**H2a:** 数据要素通过提高企业创新投入, 提升企业绿色创新质量。

新质生产力的形成需要政府和市场共同调节培育, 不仅需要市场“无形的手”进行资源配置, 还需要政府这一“有形的手”发挥作用, 进行方向引导与政策支持, 构建新型生产关系, 完善机制体制。根据马克思主义政治经济学原理, 生产关系对生产力具有反作用。在新质生产力背景下, 生产关系的调整与重构必须紧密契合数据要素这一新兴生产要素的特点与发展规律。数据要素以其独特的优化资源配置、促进信息共享和降低交易成本的能力, 正推动着生产关系向更加现代化、高效化的方向转型。这一过程不仅极大地增强了企业的创新活力与市场竞争力, 还为企业争取政府支持奠定了坚实的基础。

数据要素通过优化生产关系、促进企业与政府之间的良性互动, 以及提升企业绿色创新质量等方式发挥重要作用。在绿色创新领域, 由于存在外部性、公共品属性等问题, 市场可能无法充分激励企业进行绿色创新。此时, 政府可以通过提供补贴等干预手段弥补市场失灵, 激励企业增加绿色创新投入<sup>[18]</sup>。然而, 信息不对称的问题始终是政府支持过程中的一大挑战, 这可能导致决策失误和资源配置低效<sup>[19]</sup>。数据要素凭借其强大的信息整合与分析能力, 可以为政府提供更



加全面、准确的企业绿色创新信息，帮助政府更加精准地评估企业的创新潜力与风险，从而作出更为科学合理的补贴决策。

获得政府支持的企业，其绿色创新活动得到了强有力的资金支持。这不仅加速了绿色技术的研发与应用进程，提升了绿色创新的质量与效率，还为企业带来了声誉上的提升与合作伙伴的青睐。在此基础上，一个以绿色创新为核心的创新生态系统逐渐形成并不断优化，形成了一个正向循环的创新氛围。这种氛围进一步激发了企业的创新活力与积极性，推动了企业绿色创新质量的持续提升。其中，政府补贴和税收返还都是有效的政府支持方式，可以提供更多的资金支持。因此，本文提出以下假设：

**H2b:** 数据要素通过帮助企业获得政府补贴，提升企业绿色创新质量。

**H2c:** 数据要素通过帮助企业获得税收返还，提升企业绿色创新质量。

综上所述，具体作用机制如图1所示。

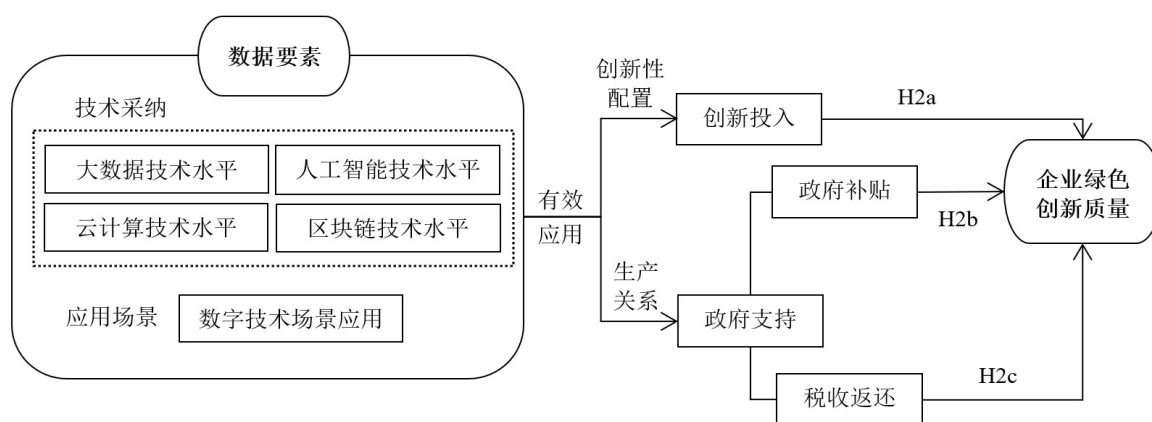


图1 新质生产力背景下数据要素影响企业绿色创新质量的机制

### 三、研究设计

#### (一) 变量选取

##### 1. 被解释变量

企业绿色创新质量 (GPR) 为本文的被解释变量。借鉴陶长琪和丁煜<sup>[20]</sup> 在评估企业创新质量方面的研究，本文采用企业绿色发明专利申请量占总绿色专利申请量的比例作为核心指标。选择这一指标具有多重优势。首先，这种方法可以有效消除企业规模对创新质量评估的潜在影响，使得不同规模的企业在创新质量上能够公平比较。其次，通过占比来度量企业绿色创新质量，还有助于排除其他不可观测因素对研究结果的干扰，提高研究的准确性和可靠性。最后，绿色发明专利申请量占比的提升可以直接反映企业创新活动的高质量、高技术发展水平。这种趋势不仅代表了企业在技术创新上的努力，也是其实现绿色发展的关键驱动力。在后续的稳健性检验中，本文采用企业独立绿色专利申请量与联合绿色专利申请量之和的自然对数作为替代变量。

##### 2. 解释变量

数据要素 (Dig) 为本文的解释变量，可以评估企业在数据处理和运用方面的能力。本文从技术应用的视角，借鉴了史青春等<sup>[21]</sup> 的研究，使用吴非等<sup>[22]</sup> 构建的关键词图谱对解释变量进行度量。该图谱包含“大数据技术水平”“人工智能技术水平”“云计算技术水平”“区块链技术水平”“数字技术场景应用”这五项指标下的细分关键词，这些关键词不仅涵盖了数据的收集、存储、清理和分析各个环节，还体现了企业在实际应用中的数据要素驱动决策的能力。

具体来说, 本文通过统计制造业企业年报中该图谱关键词的披露次数来度量企业的数据要素应用情况。这些披露次数不仅可以直接体现企业对数据要素的重视程度, 还可以间接反映企业的数据要素应用效果, 因为如果一个企业成功地将数据要素应用于实际业务中, 并取得了成效, 那么它很可能在年报中详细阐述实施情况和效果。

### 3. 控制变量

为了控制其他因素对研究的影响, 本文借鉴苏媛和李广培<sup>[23]</sup>的研究, 从企业财务特征和公司治理层面选取了9个控制变量。财务特征层面主要包括企业规模 (Size)、资产负债率 (Lev)、净资产收益率 (ROE)、成长能力 (SG)、营运能力 (TAT); 公司治理层面主要包括独立董事比例 (IBD)、高管薪酬激励 (Dpay)、股权集中度 (CR) 及两职合一 (Dual)。表1报告了文中主要变量的度量情况。

表1 变量说明

变量类型	变量名称	符 号	变量说明
被解释变量	企业绿色创新质量	GPR	绿色发明专利申请量占总绿色专利申请量的比例
解释变量	数据要素	Dig	数据要素相关关键词在年报中的出现次数
控制变量	企业规模	Size	年末总资产的自然对数
	资产负债率	Lev	总负债/总资产
	净资产收益率	ROE	净利润/平均净资产
	成长能力	SG	营业收入增长率
	营运能力	TAT	总资产周转率
	独立董事比例	IBD	独立董事会成员规模/董事会成员规模
	高管薪酬激励	Dpay	前三大高管平均薪酬的自然对数
	股权集中度	CR	第一大股东持股比例
	两职合一	Dual	董事长与总经理兼任则取值为1, 否则取值为0

## (二) 模型构建

为了考察数据要素对企业绿色创新质量的影响, 本文构建了基准模型 (1):

$$GPR_{it} = \alpha_0 + \alpha_1 Dig_{it} + cX_{it} + \sum Ind + \sum Year + \varepsilon_{it} \quad (1)$$

其中, GPR为企业绿色创新质量。Dig为数据要素。X为企业财务特征和公司治理层面的控制变量。 $\sum Ind$ 为行业虚拟变量, 用于控制行业因素的影响。 $\sum Year$ 为年份虚拟变量, 用以控制年度因素的影响。 $\varepsilon_{it}$ 为随机扰动项。为避免异方差性的影响, 本文对所有回归模型的标准误均进行了企业层面的聚类。

## (三) 数据来源

制造业涉及多个领域, 包括机械制造、电子信息、化工等, 这些行业在生产过程中往往伴随着能源消耗和环境污染。因此, 制造业企业在绿色创新方面有着迫切的需求和广阔的空间。本文以2010—2022年中国A股制造业上市企业为初始样本, 为了获得有效研究样本, 对数据进行了如下筛选: 剔除了ST类企业, 剔除了数据明显异常 (如资产负债率大于1) 的样本, 剔除了2020年之后上市的样本, 剔除了数据严重缺失的样本。经过上述筛选, 得到20794个企业一年度数据。为了避免极端值的影响, 本文对连续变量均进行了1%和99%的缩尾处理。数据要素相关数据来自上海和深圳两个证券交易所官方网站上相关企业的年报数据, 财务数据来源于CSMAR数据库, 企业绿色创新质量数据来源于CNRDS数据库。

主要变量的描述性统计结果如表2所示。由表2可知, 各变量的VIF值均小于10, 说明不存在严重的多重共线性问题。

表2 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	最大值	VIF
GPR	20 794	0. 137	0. 263	0	0. 923	—
Dig	20 794	1. 217	1. 251	0	4. 585	1. 110
Size	20 794	22. 090	1. 162	19. 973	25. 674	1. 780
Lev	20 794	0. 398	0. 192	0. 056	0. 912	1. 530
ROE	20 794	0. 060	0. 126	-0. 651	0. 342	1. 220
SG	20 794	0. 214	0. 512	-0. 635	3. 098	1. 050
TAT	20 794	0. 636	0. 354	0. 116	2. 189	1. 180
IBD	20 794	0. 385	0. 074	0. 250	0. 600	1. 020
Dpay	20 794	14. 481	0. 723	12. 758	16. 514	1. 470
CR	20 794	0. 332	0. 140	0. 087	0. 712	1. 060
Dual	20 794	0. 312	0. 463	0	1	1. 050

#### 四、实证结果与分析

##### (一) 基准回归结果分析

表3报告了基准模型(1)的回归结果。列(1)未考虑控制变量的影响,Dig的系数为0.036( $p<0.010$ ),初步验证了数据要素对企业绿色创新质量的正向影响,验证了H1。列(2)考虑了控制变量的影响,Dig的系数下降至0.026( $p<0.010$ ),再次验证了数据要素对制造业企业绿色创新质量的促进作用,这一结果还说明,如果忽略其他因素的影响,可能导致数据要素促进作用的高估。从控制变量来看,Size、ROE、SG及Dpay的系数均为正且显著,说明企业规模、盈利能力、成长能力及高管薪酬激励均有利于提高制造业企业绿色创新质量。

表3 基础回归结果

变 量	(1)	(2)
Dig	0.036*** (9.880)	0.026*** (7.540)
Size		0.047*** (10.420)
Lev		0.019 (0.970)
ROE		0.034* (1.710)
SG		0.016*** (3.080)
TAT		0.012 (1.020)
IBD		-0.019 (-0.560)
Dpay		0.016*** (2.900)
CR		-0.020 (-0.800)
Dual		0.005 (0.820)
常数项	0.093*** (19.060)	-1.174*** (-11.770)
行业/年份FE	控制	控制
观测值	20 793	20 793
R <sup>2</sup>	0.082	0.131

注:\*\*\*、\*\*和\*分别表示在1%、5%和10%的水平下显著,括号内为t值,下同。

##### (二) 内生性处理

###### 1. 控制固定效应

尽管基准模型已考虑行业和年份层面不可观测因素的影响,鉴于企业文化差异对研究的潜在

影响, 本文采用面板双向固定效应回归模型, 旨在消除不随时间变化但随企业变化的不可观测因素的影响。鉴于地区政策差异对数据要素利用效率与绿色创新质量之间的关系会产生的影响, 本文在基准模型的基础上, 进一步控制了企业所属省份的固定效应。此外, 非观测因素可能同时受到行业一年份或省份一年份共同因素变化的影响, 如政策实施有时效性, 本文考察了行业和省份时间趋势的影响。表4分别报告了控制不同因素影响下的高维固定效应回归结果。从回归结果来看, Dig的系数均为正, 分别在1%和5%的水平下显著, 说明尽可能排除非观测因素的影响后, 基本结论不变。值得注意的是, 随着控制的因素增多, Dig系数的t值呈现下降趋势, 说明考虑上述因素的影响是有必要的。

表4 控制固定效应后的回归结果

变 量	(1)	(2)	(3)	(4)	(5)	(6)
Dig	0.007*** (2.660)	0.026*** (7.450)	0.007** (2.560)	0.007** (2.360)	0.006** (2.160)	0.006** (2.010)
控制变量	控制	控制	控制	控制	控制	控制
常数项	-0.072 (-0.500)	-1.158*** (-11.640)	-0.121 (-0.820)	-0.090 (-0.610)	-0.145 (-0.990)	-0.108 (-0.730)
年份FE	控制	控制	控制	控制	控制	控制
企业FE	控制	不控制	控制	控制	控制	控制
行业FE	不控制	控制	控制	控制	控制	控制
省份FE	不控制	控制	控制	控制	控制	控制
行业/年份FE	不控制	不控制	不控制	不控制	控制	控制
省份/年份FE	不控制	不控制	不控制	控制	不控制	控制
观测值	20 773	20 785	20 764	20 754	20 759	20 759
R <sup>2</sup>	0.460	0.137	0.459	0.459	0.463	0.463

## 2.倾向得分匹配

考虑到不同数据要素应用水平的企业在绿色创新投入、盈利能力等方面呈现差异, 从而导致企业本身差异引起的绿色创新质量不同被错误归因于数据要素的影响。为了缓解上述样本自选择问题导致的内生性, 本文根据数据要素应用水平的行业中位数将样本分为数据要素应用水平高组(处理组)和数据要素应用水平低组(控制组), 并采用倾向得分匹配(PSM)的方式为处理组寻找合适的控制组, 从而消除系统差异, 最后采用PSM后的样本进行基准模型的估计。表5报告了以基准模型中的控制变量为协变量, 卡尺半径为0.010的近邻1:1匹配下的平衡性检验结果。由回归结果可知, 在PSM前, 除了股权集中度(CR)外, 处理组和控制组协变量均值差异的t检验p值均小于0.100, 说明处理组与控制组呈现显著的系统差异。在PSM后, 各协变量标准偏差的绝对值均低于5%, 并且处理组与控制组均值差异t检验的p值大于0.100, 说明PSM后的样本不存在系统差异, PSM效果较好。

为了进一步检验PSM效果, 本文还根据PSM前后处理组和控制组的倾向得分绘制了核密度曲线。<sup>①</sup>在PSM前, 处理组与控制组倾向得分核密度曲线呈现较大的差异, 在PSM后, 处理组与控制组倾向得分核密度曲线呈现明显的重合, 说明系统差异得以消除。表6报告了PSM回归结果, 考虑到PSM对匹配方法具有一定的敏感性, 作为稳健性检验, 表6还披露了近邻1:1无放回和核匹配下的回归结果。可以看到, 无论是何种匹配方式, Dig的系数均为正且显著, 说明消除样本选择偏差的影响后, 基准模型依然稳健。

① PSM前后处理组和控制组倾向得分核密度曲线未在正文中列出, 留存备案。

表5 平衡性检验结果

变量	未匹配 (U) / 匹配 (M)	均值		标准偏差 (%)	标准偏差减少幅度 (%)	t检验	
		处理组	控制组			t统计量	p值
Size	U	22.284	21.956	28.300		20.200	0.000
	M	22.284	22.266	1.500	94.600	0.960	0.337
Lev	U	0.402	0.396	3.000		2.100	0.035
	M	0.402	0.397	2.200	24.900	1.480	0.140
ROE	U	0.062	0.059	2.500		1.780	0.075
	M	0.062	0.062	-0.300	87.100	-0.210	0.835
SG	U	0.240	0.197	8.300		5.930	0.000
	M	0.240	0.254	-2.700	67.100	-1.620	0.105
TAT	U	0.642	0.633	2.600		1.870	0.061
	M	0.642	0.634	2.300	12.100	1.510	0.132
IBD	U	0.389	0.382	9.300		6.620	0.000
	M	0.389	0.389	0.100	99.100	0.050	0.959
Dpay	U	14.638	14.374	37.100		26.260	0.000
	M	14.637	14.636	0.200	99.600	0.100	0.920
CR	U	0.331	0.333	-2.000		-1.430	0.152
	M	0.331	0.331	-0.600	68.800	-0.400	0.686
Dual	U	0.331	0.298	7.100		5.080	0.000
	M	0.331	0.329	0.500	92.500	0.340	0.731

表6 PSM回归结果

变量	(1)	(2)	(3)
	近邻1:1 (有放回)	近邻1:1 (无放回)	核匹配
Dig	0.025*** (6.450)	0.027*** (6.960)	0.026*** (7.530)
Size	0.047*** (9.230)	0.045*** (8.790)	0.047*** (10.410)
Lev	0.009 (0.370)	0.026 (1.030)	0.019 (0.960)
ROE	0.025 (1.010)	0.041* (1.730)	0.034* (1.700)
SG	0.014** (2.400)	0.018*** (3.030)	0.016*** (3.090)
TAT	0.012 (0.900)	0.013 (0.980)	0.012 (1.020)
IBD	-0.026 (-0.660)	-0.016 (-0.420)	-0.019 (-0.560)
Dpay	0.015** (2.270)	0.017** (2.550)	0.016*** (2.910)
CR	-0.015 (-0.530)	-0.006 (-0.220)	-0.020 (-0.800)
Dual	0.001 (0.150)	0.002 (0.210)	0.005 (0.820)
常数项	-1.153*** (-9.820)	-1.162*** (-10.140)	-1.174*** (-11.760)
行业/年份FE	控制	控制	控制
观测值	10 702	13 953	20 790
R <sup>2</sup>	0.118	0.121	0.131

(三) 稳健性检验<sup>①</sup>

1. 替换变量

为避免研究结论受到变量度量方式的影响, 本文通过替换核心被解释变量的方式进行稳健性

① 稳健性检验结果未在正文中列出, 留存备案。



检验。本文使用企业独立绿色专利申请量与联合绿色专利申请量之和的自然对数 (GP) 作为企业绿色创新质量的替代变量。同时, 本文单独计算了独立绿色专利申请量占比 (IGPR) 和联合绿色专利申请量占比 (JGPR) 作为被解释变量的替代变量。从回归结果来看, Dig 的系数均为正且显著, 说明本文的基本结论不受核心变量度量方式的影响。

## 2. 改变估计方法

其一, 考虑到企业绿色发明专利申请量存在零值, 被解释变量零值过多可能会对估计结果产生影响, 本文采用了泊松伪极大似然估计法 (PPML) 进行稳健估计。其二, 绿色发明专利申请量占比的取值范围介于 0—1 之间, 属于受限被解释变量问题, 本文采用了 Tobit 模型进行稳健估计。其三, 绿色发明专利申请量在不进行自然对数处理的情形下是非负整数, 属于离散变量, 为典型的计数模型, 并且绿色发明专利申请量存在较多的零值, 分布较为分散, 故本文还进行了零膨胀负二项回归。回归结果表明, 无论采用哪种估计方法, Dig 的系数均为正且显著, 再次验证了基准回归结果的稳健性。

## (四) 异质性分析

### 1. 基于产权性质的异质性分析

企业产权性质可能影响数据要素与绿色创新质量的关系。为了检验产权性质的差异, 本文根据实际控制人的产权属性将样本企业分为国有企业与非国有企业。如表 7 所示, Dig 的系数均为正且显著, 为了比较两组回归的系数差异, 本文借助 Bootstrap 检验进行 300 次自助抽样得到经验 p 值为 0, 说明两组回归的系数呈现显著差异。通过比较两组系数可知, 数据要素对国有企业绿色创新质量的影响尤为显著。一方面, 国有企业和非国有企业在激励机制和风险承担方面存在差异<sup>[24]</sup>。国有企业往往受到更多的政策引导和监管, 能更有效地运用数据要素提升绿色创新质量。另一方面, 不同产权性质的企业在资源获取能力上也有所不同。国有企业通常更容易从政府获得支持和资源, 有助于其在绿色创新方面取得更大的进展; 而非国有企业可能面临更多的市场竞争, 因而在运用数据要素提升企业绿色创新质量时可能面临更大的挑战。

表 7 基于产权性质的异质性回归结果

变 量	(1)	(2)
	国有企业	非国有企业
Dig	0.034*** (3.890)	0.024*** (6.910)
控制变量	控制	控制
常数项	-1.326*** (-7.400)	-0.950*** (-8.070)
行业/年份 FE	控制	控制
观测值	6 538	14 255
R <sup>2</sup>	0.169	0.117

### 2. 基于行业性质的异质性分析

为了考察行业性质差异所带来的影响, 本文借鉴《战略性新兴产业分类目录》《战略性新兴产业分类 (2012) (试行)》确定高科技上市企业行业代码为 C25、C26、C27、C28、C29、C31、C32、C34、C35、C36、C37、C38、C39、C40、C41, 并进行分样本回归。如表 8 所示, 从列 (1) 和列 (2) 的回归结果来看, Dig 的系数虽然均为正, 但只有在高新技术行业中通过了显著性检验, 说明数据要素对企业绿色创新质量的影响在高新技术行业中更加明显。

本文借鉴鲁桐和党印<sup>[25]</sup>的研究, 把固定资产比重和研发支出比重作为分类指标, 采用聚类分析法将上述行业分为劳动密集型、资本密集型及技术密集型三类, 并在此基础上借鉴 2012 年修订的《上市公司行业分类指引》将本文的样本分为劳动密集型和资本与技术密集型, 并进行分

样本回归,从表8列(3)和列(4)的回归结果来看,Dig的系数仅在列(4)中为正且显著,说明数据要素对企业绿色创新质量的影响在资本与技术密集型行业中更加明显。

综上所述,在高新技术行业或资本与技术密集型行业的企业中,数据要素对企业绿色创新质量的促进作用尤为显著。这背后的深层次原因可能在于数据在此类行业中扮演的双重核心角色:既是战略决策不可或缺的基石,又是技术创新与产品开发活力涌动的源泉。这些企业往往具有先进的数据分析工具和技术,能够从海量数据中提取出更高质量、更具价值的新信息。这些新信息不仅有助于企业精准识别市场趋势和客户需求,而且能够为企业的价值创造提供有力支持。同时,通过数据的深度挖掘和分析,企业还能够更好地激活和利用冗余资源,实现资源的优化配置和高效利用。基于数据作出的科学决策提升了绿色创新质量,使企业能够更快速地响应市场变化,共同推动数据要素在绿色创新中的深度应用并取得显著成效。

表8 基于行业性质的异质性回归结果

变 量	(1)	(2)	(3)	(4)
	高新技术	非高新技术	劳动密集型	资本与技术密集型
Dig	0.030*** (7.580)	0.003 (0.710)	-0.000 (-0.050)	0.036*** (8.240)
控制变量	控制	控制	控制	控制
常数项	-1.263*** (-11.100)	-0.723*** (-4.210)	-0.723*** (-4.210)	-1.490*** (-11.580)
行业/年份FE	控制	控制	控制	控制
观测值	17 146	3 647	8 069	12 724
R <sup>2</sup>	0.129	0.071	0.080	0.137

### 3.基于金融约束程度的异质性分析

为了考察金融约束程度对数据要素与绿色创新质量关系的影响,本文借鉴陶长琪和丁煜<sup>[20]</sup>的研究,采用(应收账款-应付账款+存货+现金持有量)/总资产作为企业金融约束程度的替代变量,该比值越高,意味着现金流越充裕,面临的金融约束程度越低。本文根据金融约束程度的行业年度均值将样本分为低金融约束程度企业和高金融约束程度企业,并进行分样本回归。如表9所示,Dig的系数均为正且显著,并且Bootstrap检验的p值为0,数据要素对企业绿色创新质量的影响呈现金融约束程度上的差异,通过比较系数可知,数据要素对企业绿色创新质量的影响在低金融约束程度企业中更加明显。

表9 基于金融约束的分样本回归结果

变 量	(1)	(2)
	低金融约束程度企业	高金融约束程度企业
Dig	0.030*** (6.680)	0.021*** (4.730)
控制变量	控制	控制
常数项	-1.165*** (-8.810)	-1.165*** (-8.810)
行业/年份FE	控制	控制
观测值	9 995	10 798
R <sup>2</sup>	0.135	0.139

基于金融发展理论,高金融约束程度企业可能面临更多的融资难题<sup>[26]</sup>,这会使其应用数据要素提升绿色创新质量的能力。金融约束情况还会影响企业的投资策略和风险偏好。在金融约束较紧的情况下,企业可能更加注重短期收益和风险控制,而在绿色创新方面的投入可能相对较少;在金融约束较松的情况下,企业可能更加注重长期发展和创新投入,从而更有动力应用数据要素提升绿色创新质量。

### (五) 机制检验

本文假设数据要素可能通过提高企业创新投入和帮助企业获得政府支持的方式影响企业的绿色创新质量。为了验证上述机制, 本文借鉴邵帅等<sup>[27]</sup>的研究构建了模型(2)和模型(3):

$$\text{Med}_{it} = \theta_0 + \theta_1 \text{Dig}_{it} + cX_{it} + \sum \text{Ind} + \sum \text{Year} + \varepsilon_{it} \quad (2)$$

$$\text{GPR}_{it} = \varphi_0 + \varphi_1 \text{Dig}_{it} + \varphi_2 \text{Med}_{it} + cX_{it} + \sum \text{Ind} + \sum \text{Year} + \varepsilon_{it} \quad (3)$$

其中, Med为机制变量, 文中分别为创新投入和政府支持。在创新投入的度量方面, 本文参考现有研究<sup>[28]</sup>, 选用企业研发强度(RDS)作为替代变量, 具体以研究与试验发展(R&D)投入在营业收入中的占比进行度量。同时, 本文采用企业获得的政府补贴(Sub)及收到的税收返还(Tax)度量政府支持<sup>[29]</sup>, 并对上述两个变量进行自然对数处理。模型(2)用以考察数据要素对机制变量的影响, 模型(3)用以考察纳入机制变量后, 数据要素对绿色创新质量的影响, 并与基准模型进行比较, 若考虑机制变量影响后, Dig的系数呈现较大的变化, 则意味着机制得到了验证。

#### 1. 创新投入的机制检验

表10报告了创新投入的机制检验结果, 列(1)为基准回归结果, 作为模型(3)的对照。列(2)中Dig的系数为0.367且通过了1%显著性水平的检验, 说明数据要素应用水平的提高促进了企业创新投入。列(3)中RDS的系数为正且通过了1%显著性水平的检验, 说明企业加大创新投入有利于其绿色创新质量的提升。Dig的系数为正且通过了1%显著性水平的检验, 并且影响系数由0.025下降至0.024, 说明创新投入发挥了部分机制效应, 进一步可以计算得到间接效应占总效应的比例为4.0%, 说明数据要素不仅会直接影响绿色创新质量还会通过提高创新投入间接提升绿色创新质量, 验证了H2a。

表10 创新投入的机制检验结果

变 量	(1)	(2)	(3)
	GPR	RDS	GPR
Dig	0.025*** (6.740)	0.367*** (4.600)	0.024*** (6.460)
RDS			0.003*** (2.670)
控制变量	控制	控制	控制
常数项	-1.218*** (-11.170)	2.954 (1.580)	-1.227*** (-11.300)
行业/年份FE	控制	控制	控制
观测值	16 673	16 673	16 673
R <sup>2</sup>	0.131	0.176	0.135

#### 2. 政府支持的机制检验

如表11所示, 列(1)—列(3)报告了政府支持的机制检验结果, 列(1)为基准回归, 作为模型(3)的对照。列(2)中Dig的系数为0.037且通过了10%显著性水平的检验, 说明数据要素应用水平的提高有助于企业获得更多的政府补贴。列(3)中Sub的系数为正且通过了1%显著性水平的检验, 说明政府加大补贴力度鼓励企业绿色创新, 有利于其绿色创新质量的提升, 并且Dig的系数依然为正且显著, 说明政府补贴发挥了部分机制效应, 验证了H2b。列(4)—列(6)的结果为税收补贴的机制检验结果, 列(5)中Dig的系数虽然为正, 但未通过显著性检验, 说明税收返还的机制未得到验证, 即H2c未得到验证。

政府补贴和税收返还同样都是政府支持企业绿色发展的重要手段, 但税收返还的效果并不明显, 本文认为可能存在以下三个原因: 第一, 在资金流方面, 政府补贴直接注资且专款专用, 高效促进绿色创新。第二, 在监管评估方面, 政府补贴伴随严格监管, 确保资金有效利用, 而税收返还则监管薄弱, 资金用途难以控制。第三, 在作用机制方面, 政府补贴直接降低创新门槛与风

险，激励效果显著，税收返还的间接激励则相对较弱。

表 11 政府支持的机制检验结果

变 量	(1)	(2)	(3)	(4)	(5)	(6)
	GPR	Sub	GPR	GPR	Tax	GPR
Dig	0.025*** (6.590)	0.037* (1.720)	0.025*** (6.540)	0.026*** (7.040)	0.000 (0.010)	0.026*** (7.050)
Sub			0.009*** (5.330)			
Tax						0.002*** (3.800)
控制变量	控制	控制	控制	控制	控制	控制
常数项	-1.223*** (-11.400)	-2.947*** (-4.210)	-1.195*** (-11.340)	-1.245*** (-11.800)	-18.663*** (-11.660)	-1.202*** (-11.460)
行业/年份FE	控制	控制	控制	控制	控制	控制
观测值	18 335	18 335	18 335	18 792	18 792	18 792
R <sup>2</sup>	0.131	0.310	0.134	0.133	0.133	0.134

(六) 实现条件

在新质生产力持续演进的框架下，劳动者角色与劳动资料范畴的深化与拓展显得尤为关键，必须紧密跟随时代步伐不断革新。特别地，当数据要素日益成为推动新质生产力形成与升级的核心驱动力时，企业对大数据技术的革新速度及其在实际应用中的效能寄予了厚望。同时，企业对掌握先进数字技能与素养的劳动者提出了更为严苛与多元化的要求。

技术创新需要达到一定的门槛才能产生显著的效果<sup>[30]</sup>，而大数据技术作为一种共性技术，是现代劳动资料的重要组成部分，也是数据要素转化为生产力的关键劳动工具。大数据技术的发展水平直接决定了企业能不能从海量数据中提取有价值的信息，为绿色创新提供决策支持。也就是说，只有当企业掌握了足够成熟的大数据技术，才能确保数据要素被最大化地应用，进而推动绿色创新的发展。

同时，新质生产力强调劳动者、劳动资料、劳动对象及其优化组合的质变。在这一过程中，劳动者不再仅仅是简单劳动力的提供者，而是需要具备高素质、持续学习能力和创新精神的新型劳动者。依据人力资本理论和创新理论，数据要素的有效应用不仅依赖于先进的技术，更依赖于那些具备深厚专业知识和分析能力的人才<sup>[31]</sup>。只有当企业的人力资本质量达到一定的标准，能够熟练、准确地操作和分析数据时，数据要素才能真正发挥其赋能绿色创新的作用。鉴于此，本文深入剖析并系统探讨企业大数据技术水平与人力资本质量在数据要素赋能绿色创新质量提升中的关键作用。

具体来说，本文借鉴黄贇琳等<sup>[32]</sup>的研究，分析数据要素对绿色创新的促进效应是不是受到大数据技术水平和人力资本质量的门槛效应影响。换句话说，本文试图分析是否只有当大数据技术水平和人力资本质量达到某个特定水平时，数据要素才能对企业绿色创新质量产生显著影响。本文构建了模型(4)：

$$GPR_{it} = \theta_0 + \theta_1 Dig_{it} \times I(q_{it} \leq \gamma_1) + \theta_2 Dig_{it} \times I(\gamma_1 < q_{it} \leq \gamma_2) + \theta_3 Dig_{it} \times I(q_{it} > \gamma_2) + cX_{it}/Z_{it} + f_i + v_t + \varepsilon_{it} \quad (4)$$

其中，I(·)为示性函数，q为门槛变量，γ为门槛值。考虑到本文的样本数据为非平衡面板。因此，使用“xthregs”命令考察非平衡面板数据的门槛效应模型，该命令最多只能讨论双重门槛效应。f<sub>i</sub>为个体固定效应，v<sub>t</sub>为时间固定效应。本文采用大数据技术指标对企业的大数据技术水平进行度量(Big\_data)，即根据年报中出现大数据技术相关关键词的总词频进行度量。人力资本



质量 (HC) 指标以企业大专及以上学历员工数量占比进行度量。

进行门槛效应回归之前需要对门槛效应进行检验, 表 12 报告了门槛效应检验结果, 大数据技术 (Big\_data) 第一门槛效应检验的 F 统计量为 6.250 且通过了 5% 显著性水平的检验, 双重门槛效应检验的 F 统计量较小, 未通过显著性检验, 说明存在大数据技术的单一门槛效应。同理可知, 人力资本质量 (HC) 的单一门槛和双重门槛的 F 统计量均通过了显著性检验, 说明存在人力资本质量的双重门槛效应。表 12 不仅提供了门槛效应的存在性检验, 还给出了门槛值的估计结果及 95% 的置信区间。值得注意的是, 两个门槛变量的门槛值均落在置信区间内, 说明门槛值估计结果比较准确。

表 12 门槛效应检验结果

门槛变量	门槛类型	F 统计量	p 值	门槛值	95% 置信区间
Big_data	单一门槛	6.250**	0.043	10	[0, 20]
	双重门槛	1.630	0.370	11	[11, 20]
HC	单一门槛	14.150***	0.000	0.244	[0.015, 0.640]
	双重门槛	4.550*	0.057	0.565	[0.250, 0.782]

如表 13 所示, 列 (1) 以 Big\_data 为门槛变量, 当 Big\_data 到达门槛值 (10) 之前, Dig 的系数为 0.006 ( $p < 0.010$ ), 跨过门槛值之后, Dig 的系数上升至 0.015 ( $p < 0.010$ ), 说明随着企业大数据技术水平的提高, 数据要素对企业绿色创新质量的影响逐渐增强。这意味着, 要想提高数据要素对企业绿色创新质量的影响, 需要提高企业的大数据技术水平。通过简单的描述性统计可知, Big\_data 大于 10 的样本仅占总样本的 2.3%, 说明大部分企业的大数据技术水平较低, 应当提高企业的大数据技术水平。

从列 (2) 的回归结果来看, 当 HC 到达第一门槛值 (0.244) 之前, Dig 的系数虽然为正, 但并不显著。当 HC 跨过第一门槛值到达第二门槛值 (0.565) 之前, Dig 的系数为 0.011 ( $p < 0.010$ )。当 HC 跨过第二门槛值之后, Dig 的系数上升至 0.023 ( $p < 0.010$ )。当人力资本质量较低时, 数据要素与人力资本未能较好地匹配, 难以实现数据要素驱动制造业企业绿色创新质量的提升。只有当人力资本质量大于 0.244 时, 数据要素才能够促进企业绿色创新质量的提升, 并且随着人力资本质量的提高, 数据要素对企业绿色创新质量的正向影响逐渐增强, 说明人力资本质量是数据要素促进企业绿色创新质量的有效实现条件。

通过描述性分析可知, 人力资本质量高于 0.244 的样本占总样本的 34.6%, 说明仍然有 65.4% 样本企业的数据要素与人力资本未能较好地匹配, 仅有 5.7% 的样本企业人力资本质量高于 0.565, 说明企业亟须提高人力资本质量, 提高高质量人力资本的比重, 才能加大数据要素对企业绿色创新质量的促进作用。

表 13 门槛面板回归结果

变 量	(1)	(2)
	Big_data	HC
$Dig_{it} \times (q_{it} \leq \gamma_1)$	0.006*** (2.860)	0.003 (0.970)
$Dig_{it} \times (\gamma_1 < q_{it} \leq \gamma_2)$	0.015*** (4.030)	0.011*** (3.710)
$Dig_{it} \times (q_{it} > \gamma_2)$		0.023*** (4.900)
控制变量	控制	控制
常数项	-0.117 (-1.210)	-0.150 (-1.440)
企业/年份 FE	控制	控制
观测值	20 794	19 246

## 五、研究结论与政策启示

新质生产力是引领中国经济迈向高质量发展的必由之路,其内核深植于绿色发展的核心理念之中,其本质是绿色生产力的现代诠释。尽管现有研究已从多维度、多层次对新质生产力的理论基础与内涵外延进行了深入剖析,但在实证方面,关于新质生产力如何具体赋能绿色发展、促进绿色创新的研究尚显不足,缺乏系统而全面的量化分析。尤为关键的是,针对数据这一核心要素在绿色创新过程中所发挥的具体作用,现有研究存在分歧,尚未达成广泛共识。本文在新质生产力的理论体系下,实证检验了数据要素对企业绿色创新质量的促进效应。研究结果显示:数据要素对企业绿色创新质量的提升有显著的正向效应。在经过一系列内生性检验和稳健性检验后,基准回归结果依然成立。异质性分析结果发现,这种绿色赋能作用在国有企业、高新技术企业、资本与技术密集型企业及低金融约束程度企业中表现更为明显。机制分析结果表明,数据要素能通过提高企业创新投入和帮助企业获得政府支持的方式提升企业绿色创新质量。最后,本文验证了数据要素对绿色创新的促进效应受到企业大数据技术水平和人力资本质量的门槛效应影响。基于上述研究结论,本文得到以下政策启示:

第一,强化政企协同机制,共绘数据要素赋能新蓝图。政府需进一步深化与企业的合作,通过设立专项扶持基金、搭建科研创新合作桥梁等举措,激励企业深耕数据技术的研发前沿与应用实践,旨在全面提升企业在数据挖掘、深度分析及创新应用上的综合能力。此外,政府应不断完善数据使用的标准体系与法规框架,加速数据资源的开放共享进程,帮助企业降低数据获取门槛与成本,为数据要素的高效流通与价值释放奠定坚实的制度基础。

第二,深化金融与科技的融合创新,为绿色创新项目注入强劲动力。鉴于低金融约束程度下企业展现出更佳的绿色创新表现,政府应持续优化金融生态,简化融资流程,降低企业融资成本,特别是针对绿色创新项目,应提供定制化的专项贷款、风险投资等多元化金融支持。同时,倡导金融机构依托企业数据信用体系,创新金融产品与服务,确保资金流向绿色、高效的创新领域,加速资本与科技的深度融合。

第三,聚焦人力资本质量提升,构筑绿色创新人才高地。人才作为绿色创新的核心驱动力,为了培育人才,政府应加大对教育资源的投资,强化数据科学、环境科学及能源技术等关键领域的专业人才培养与引进。同时,激励企业构建内部培训体系,提升员工的数据思维与创新技能,为绿色创新项目培养复合型、高素质的人才队伍。此外,推动产学研用一体化发展,促进企业、高校与科研机构之间的紧密合作,加速绿色创新技术的研发与成果转化,增强企业的绿色创新能力与市场竞争力,共同推动绿色经济蓬勃发展。

### 参考文献:

- [1] 周文,何雨晴.新质生产力:中国式现代化的新动能与新路径[J].财经问题研究,2024(4):3-15.
- [2] 邱海平.新质生产力理论的科学内涵及其重大创新意义[J].财经问题研究,2024(5):3-14.
- [3] 戴翔,杨双至.数字赋能、数字投入来源与制造业绿色化转型[J].中国工业经济,2022(9):83-101.
- [4] 蔡继明,刘媛,高宏,等.数据要素参与价值创造的途径——基于广义价值论的一般均衡分析[J].管理世界,2022,38(7):108-121.
- [5] 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.
- [6] 蔡思航,翁翕.一个数据要素的经济学新理论框架[J].财经问题研究,2024(5):33-48.
- [7] 陈彦斌,张永山,刘瑞,等.新质生产力笔谈[J].财经问题研究,2024(7):3-18.
- [8] 周文,刘守英,郑红亮,等.专题笔谈:发展新质生产力的理论与实践问题[J].东北财经大学学报,2024(4):

- 3-18.
- [9] ZHAO Y, XIN L. Research on green innovation countermeasures of supporting the circular economy to green finance under big data[J]. *Journal of enterprise information management*, 2021, 35(4/5): 1305-1322.
- [10] 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.
- [11] 曹裕, 李想, 胡韩莉, 等. 数字化如何推动制造企业绿色转型? ——资源编排理论视角下的探索性案例研究[J]. *管理世界*, 2023, 39(3): 96-112+126+113.
- [12] LANGE S, POHL J, SANTARIUS T. Digitalization and energy consumption. Does ICT reduce energy demand? [J]. *Ecological economics*, 2020, 176: 106760.
- [13] 蒲清平, 黄媛媛. 习近平总书记关于新质生产力重要论述的生成逻辑、理论创新与时代价值[J]. *西南大学学报(社会科学版)*, 2023, 49(6): 1-11.
- [14] 汪寿阳, 刘颖. 揭示科技创新对发展生产力的重要作用[N]. *人民日报*, 2024-04-09(9).
- [15] 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.
- [16] 贾康. 新质生产力的概念、内涵与体系: 以“创新”促“质变”[J]. *东北财经大学学报*, 2024(4): 19-26.
- [17] 彭影, 李士梅. 创新要素流动与城市绿色创新发展——数据要素流动环境的空间调节作用[J]. *科技进步与对策*, 2023, 40(1): 30-39.
- [18] 李青原, 肖泽华. 异质性环境规制工具与企业绿色创新激励——来自上市企业绿色专利的证据[J]. *经济研究*, 2020, 55(9): 192-208.
- [19] 屈文洲, 谢雅璐, 叶玉妹. 信息不对称、融资约束与投资—现金流敏感性——基于市场微观结构理论的实证研究[J]. *经济研究*, 2011, 46(6): 105-117.
- [20] 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.
- [21] 史青春, 牛悦, 徐慧. 企业数据要素利用水平影响投资效率机理研究——利用数据要素激活冗余资源的中介作用[J]. *中央财经大学学报*, 2023(11): 105-115.
- [22] 吴非, 胡慧芷, 林慧妍, 等. 企业数字化转型与资本市场表现——来自股票流动性的经验证据[J]. *管理世界*, 2021, 37(7): 130-144+10.
- [23] 苏媛, 李广培. 绿色技术创新能力、产品差异化与企业竞争力——基于节能环保产业上市公司的分析[J]. *中国管理科学*, 2021, 29(4): 46-56.
- [24] 王朝才, 汪超, 曾令涛. 财政政策、企业性质与资本结构动态调整——基于A股上市公司的实证研究[J]. *财政研究*, 2016(9): 52-63.
- [25] 鲁桐, 党印. 公司治理与技术创新: 分行业比较[J]. *经济研究*, 2014, 49(6): 115-128.
- [26] ZHAO M, ZHANG B. Credit corruption, financial constraint and corporate innovation: evidence on China [J]. *Finance research letters*, 2023, 53: 103658.
- [27] 邵帅, 张可, 豆建民. 经济集聚的节能减排效应: 理论与中国经验[J]. *管理世界*, 2019, 35(1): 36-60+226.
- [28] 应千伟, 何思怡. 政府研发补贴下的企业创新策略: “滥竽充数”还是“精益求精”[J]. *南开管理评论*, 2022, 25(2): 57-69.
- [29] 吴怡俐, 吕长江, 倪晨凯. 增值税的税收中性、企业投资和企业价值——基于“留抵退税”改革的研究[J]. *管理世界*, 2021, 37(8): 180-194.
- [30] 万建香, 汪寿阳. 社会资本与技术创新能否打破“资源诅咒”? ——基于面板门槛效应的研究[J]. *经济研究*, 2016, 51(12): 76-89.
- [31] ERIKSSON C, LINDÉN J, PAPAHRISTODOULOU C. Human capital, innovation, and growth [J]. *International journal of economic theory*, 2022, 19(2): 343-369.
- [32] 黄贇琳, 秦淑悦, 张雨朦. 数字经济如何驱动制造业升级[J]. *经济管理*, 2022, 44(4): 80-97.

## Data Factors and Green Innovation: A Perspective of New Quality Productive Forces

WANG Ke-dan<sup>1, 2</sup>, LIU Ying<sup>2, 3</sup>, WANG Shou-yang<sup>2, 3</sup>

(1. International Business School, Dalian Minzu University, Dalian 116600, China;

2. School of Economics and Management, University of Chinese Academy of Sciences, Beijing 100190, China;

3. Academy of Mathematics and Systems Sciences, Chinese Academy of Sciences, Beijing 100190, China)

**Summary:** With the accelerated advancement of the global technological revolution and industrial transformation, digitalization and greenization have become the most distinctive features of the new quality productive forces, and the deep integration and collaborative development between the two has become the core issue of current scientific exploration. As the core factor of new quality productive forces, data factors are the key engine driving the leapfrog development and transformation and upgrading of productive forces. Green development, as the inherent requirement and background of high-quality development, is an important yardstick for measuring new quality productive forces. However, existing literature is limited to theoretical explanations and lacks empirical analysis based on enterprise data. There is still a lack of systematic and in-depth research on how data factors can empower green innovation from the perspective of new quality productive forces.

This paper uses the annual report data of listed companies, the China Stock Market & Accounting Research (CSMAR) database, and the Chinese Research Data Service Platform (CNRDS) database. It draws on the keyword mapping of Wu et al. (2021) and the measurement method of Shi et al. (2023) for the utilization of enterprise data factors. It assesses the application of data factors by enterprises by quantifying the disclosure frequency of keywords in the annual reports of manufacturing companies. Empirical research shows that data factors have a positive impact on the improvement of the quality of corporate green innovation. After considering endogeneity problems, replacing core variables, and changing estimation methods, the regression results remain robust. Further analysis shows that data factors have a more significant effect on the quality of green innovation in state-owned enterprises, high-tech enterprises, capital-and technology-intensive enterprises, and enterprises with low financial constraints. Its mechanisms cover multiple paths such as innovation investment and government support. At the same time, the progressiveness of data mining technology and the optimal allocation of human capital have become key conditions for the realization of data-driven green innovation.

This paper expands on existing literature in the following three aspects. First, it deepens the theory of new quality productive forces and expands the boundaries of green development theory. Secondly, considering the dynamics of Chinese industries and the socioeconomic background, it provides micro-evidence of the positive impact of data factors on the quality of green innovation. Thirdly, from the perspective of the theory of productive forces, it comprehensively analyzes the mechanism and conditions for realizing data-driven green innovation.

**Key words:** new quality productivity forces; data factors; quality of green innovation in enterprises; green development

(责任编辑: 邓 菁)

[DOI]10.19654/j.cnki.cjwtyj.2024.09.002

[引用格式]王柯丹,刘颖,汪寿阳.数据要素与绿色创新:基于新质生产力视角[J].财经问题研究,2024(9):18-33.