

大数据应用、融资约束和企业创新效率

孙洁 李杰

(天津财经大学会计学院, 天津 300222)

摘要: 在全球数字化浪潮下,“大数据”已经成为企业获得竞争优势的利器。本文基于沪深A股上市公司2007—2020年数据,借助爬虫技术收集企业年报中的“大数据”关键词,实证检验大数据应用对企业创新效率的影响及其作用机制。研究发现,企业大数据应用显著提高了企业创新效率,且这一结论在经过更换变量衡量方式、改变样本区间以及内生性处理等一系列稳健性检验后仍然成立。机制分析表明,大数据应用能够降低企业面临的融资约束从而提升企业的创新效率。异质性分析表明,在科技企业以及行业竞争较为激烈的企业中,大数据应用对创新效率产生的正向影响更明显。本文揭示了大数据应用提高企业创新效率的机理,扩展了企业大数据应用的相关研究。

关键词: 大数据; 创新效率; 融资约束; 文本分析; 信息不对称

Abstract: With the development of global digitalization, “big data” has become a powerful tool for enterprises to improve their competitive advantage. Based on the data of listed firms on the Shanghai and Shenzhen stock markets from 2007 to 2020, this study crawls the keywords of “big data” in the annual reports of enterprises to explore the impact of big data implementation on the innovation efficiency of enterprises and its mechanism. Empirical evidence indicates that the implementation of big data substantially contributes to innovation efficiency. This conclusion remains valid after a series of robust analyses, such as adjusting the variable measurement methods, changing sample intervals, and endogeneity treatments. Mechanistic analysis shows that the implementation of big data can reduce the financing constraints of enterprises thereby promoting their innovation efficiency. Furthermore, heterogeneous analyses show that the positive effect of the implementation of big data on enterprise innovation efficiency is more significant for enterprises in technology and highly competitive industries. This paper presents the mechanism of big data implementation for enterprise innovation efficiency and extends the research related to big data implementation in enterprises.

Key words: big data, innovation efficiency, financial constraints, textual analysis, information asymmetry

作者简介: 孙洁,女,管理学博士,天津财经大学会计学院教授、博士生导师,研究方向:企业财务危机与资本市场、基于机器学习的企业财务危机预测、企业信用评估等。李杰(通讯作者),女,天津财经大学会计学院博士生,研究方向:企业财务危机预测、企业信用评估、机器学习等。

中图分类号: F272.3 **文献标识码:** A

一、引言

随着互联网和信息技术的高速发展,数据作为一种新型的关键生产要素,越来越受到社会各界的广泛关注:2014年“大数据”被首次写入《政府工作报告》,2015年国务院发布的《促进大数据发展的行动纲要》将“大数据”上升至国家战略,2017年“十七大”提出要推动“大数据”与实体经济的深度融合,2019年十九届

四中全会第一次将数据作为生产要素提出,2020年《关于构建更加完善的要素市场化配置体制机制的意见》正式将数据列为新型生产要素,2021年12月“十四五”数字经济发展规划指出数据要素是数字经济深化发展的核心引擎。数据爆发式增长,海量的数据中蕴含着巨大的价值,如何利用好数据要素协同推进企业技术创新,是实现数字经济发展的关键命题。实现数据技术与企业实体的深度融合,已成为我国经济由高速增长向高质量发

展转变的重要方式。

2022年《政府工作报告》提出“深入实施创新驱动发展战略”，明确“强化企业创新的主体地位，促进科技成果转移转化”。据中国国家统计局统计，我国2020年创新投入经费达24393.1亿元，较上年增长10.2%，已连续五年实现两位数增长，创新投入总量稳居世界第二。但是，一方面，我国企业的平均创新投入较低且不同企业间的投入差距较大(刘惠好和焦文姐，2021)^[32]，多数企业在创新过程中仍面临“融资难、融资贵”的现状(鞠晓生等，2013；吴翌琳和黄实磊，2021)^{[27][42]}；另一方面，我国的整体技术水平依然处于较为落后的地位，企业的技术创新效率仍然较低(肖文和林高榜，2014)^[43]。大数据是我国经济转型的重要支点(张叶青等，2021)^[49]，明确其对微观企业创新效率的影响及其作用机理，有助于为推进大数据与实体经济融合提供理论参考。

目前，一些学者对数字经济和企业创新的关系进行了探讨与分析。Ghasemaghaei and Calic(2020)^[9]、许芳等(2020)^[45]通过调研形式构建企业大数据特征和企业创新绩效的代理指标，实证研究大数据与企业创新绩效之间的关系。但是，采用问卷调查方式测量相应指标容易受到调查样本选择差异以及被调查企业主观意见的影响，难以客观地反映企业应用大数据的真实情况。之后，部分学者通过实证研究发现数字经济对企业创新绩效(黄节根等，2021)^[26]、创新投入(张叶青等，2021；Wen et al., 2022)^{[49][20]}和创新效率(杨水利等，2022)^[46]等方面存在显著促进作用。然而，这些研究并未对数字经济如何影响企业创新活动的作用机理进行深入探讨。近期，肖土盛等(2022)^[44]通过实证研究发现，数字化转型能够通过推动企业人力资本升级和改善公司治理来促进企业创新。此外，一些学者通过实证研究发现，区域数字金融的发展能够通过缓解企业面临的融资约束来提升企业的全要素生产率(王道平和刘琳琳，2021)^[37]、整体创新能力(周振江等，2021；李健等，2022)^{[53][29]}、绿色创新水平(Li et al., 2022；Liu et al., 2022)^{[13][14]}和企业财务可持续性(李宾等，2022)^[28]等。不同于上述研究，本文重点关注企业层面的大数据应用程度对企业创新效率的影响，并探讨融资约束在两者关系中起到的中介作用。

本文基于2007—2020年沪深A股上市公司样本，使

用年报中“大数据”相关的关键词词频构造大数据应用程度的代理指标，研究大数据应用对企业创新效率的影响。本文的分析结果表明：第一，大数据应用能够显著提高企业的创新效率，且该结果在更换大数据应用和创新效率衡量指标、更换研究样本区间和解决内生性问题等一系列测试后依然稳健；第二，机制分析表明，大数据应用能够通过缓解企业面临的融资约束，提高企业的创新效率；第三，异质性分析表明，大数据应用对企业创新效率的提高作用在科技企业和行业竞争程度较为激烈的企业中更为显著。

本文可能的贡献有以下几点：第一，在研究视角上，已有研究侧重于从创新投入和创新绩效等方面来研究数字化经济的影响，鲜有文献关注企业大数据应用对企业创新效率的影响。本文以企业大数据应用水平为切入点，研究大数据应用对企业创新效率的影响，在一定程度上补充和完善了企业创新效率影响因素和大数据应用经济后果的相关文献。第二，已有研究多从企业外部的角度衡量区域数字金融发展程度对融资约束的影响，鲜有文献关注企业层面的数字技术应用水平对融资约束的影响。本文探索融资约束在企业大数据应用对创新效率影响中存在的中介作用，不仅有助于厘清大数据应用对企业创新效率的作用机制与路径，为大数据应用与实体经济的深度融合提供经验证据，而且在一定程度上丰富了企业层面大数据技术应用对融资约束影响的相关文献。第三，通过探索大数据对创新效率影响的异质性，丰富了相关研究，为政府制定相关政策提供了参考。

二、理论分析与研究假设

(一)大数据应用与企业创新效率

企业创新活动具有周期长、不确定大、失败率高但潜在收益巨大等特点(陈德球等，2021)^[23]，提升创新效率是企业保持核心竞争力的重要途径。大数据是数字经济发展的重要引擎，其核心是对海量数据的生产、采集、存储、加工和分析等(张叶青等，2021)^[49]。企业应用大数据可以提高企业的创新效率。在决定创新方向的决策阶段，企业与消费者之间存在的信息不对称以及对政府政策信息与最新技术信息的把握不及时严重阻碍了企业准确判定研发方向，如此可能降低企业的创新动

机或者迫使企业做出偏离市场需求方向的研究决策，大幅增加企业创新活动的不确定性与风险。信息论创始人香农认为，信息是一种降低不确定性的东西(Shannon, 1948)^[18]，在企业决策相关领域，拓宽获取信息的渠道有助于企业做出有效的决策(Merendino et al., 2018)^[15]。通过前端埋点、爬虫等数据挖掘技术，企业在充分收集内部相关信息的基础之上，能够以较低的成本通过多个渠道(如电子商务平台、微博、知乎、政府网站、政务新媒体、新闻报道、中国开发者网络等)快速取得市场信息、政府政策信息和技术信息等与企业创新决策相关的结构化、非结构化数据信息。通过吸收海量有效信息，企业可以精准把握市场的需求动向与企业内部的现实情况，借助大数据相关的技术工具对创新活动的行进方向进行演进与推断(Barton and Court, 2012; 唐松等, 2020)^{[3][35]}，最终做出更符合客户需求的创新决策，从源头降低企业创新活动失败的风险。

在创新项目实施阶段，对知识的管理能力是企业将创新行为持续下去的关键(Aboelmaged, 2014)^[1]，大数据应用可以通过提升企业的知识管理能力来助力创新项目的实施(Sumbal et al., 2017; 曹平等, 2021)^{[19][22]}，从而提高企业创新活动的成功率。企业的知识管理是指获取、分享和有效使用知识资源的过程(Davenport et al., 1998; Lee and Kim, 2001)^{[6][11]}。首先，企业通过数据清洗、分类、预测和聚类等大数据挖掘技术手段，能够更好地利用海量数据揭示难以发现的规律，判断数据间的联系和逻辑，形成隐性知识，提升企业获取知识资源的效率，为企业的创新活动打下良好基础。其次，有效的知识共享不仅包含相关知识信息的传递，而且需要确保不同人员对知识含义的同步理解(Fisher et al., 1997)^[8]。企业利用数据可视化等技术能够将复杂的知识通过图表等简单的方式进行诠释(张瑞等, 2018)^[51]，从而实现知识在各部门中快速准确地传输与分享，加强创新人员之间的相互交流、学习与思想碰撞，激发创新思维，从而促进创新项目的成功实施。

在创新成果转化阶段，企业通过大数据可以更好地理解环境和客户需求，助力创新成果成功转化。通过运用数据管理方法，企业可以使用爬虫、神经网络、支持向量机等大数据相关技术挖掘与分析消费者的基本信

息、行为和交易数据来建设客户画像体系(Cinar et al., 2015; Nedjah et al., 2022)^{[5][17]}，从而有助于企业快速锁定企业创新成果转化的目标与对象。通过针对精准定位的潜在消费群体采取适当的营销手段，企业能够增加新产品相关的销量与收入(Bao et al., 2022)^[2]，实现创新成果的成功商业化，从而提升企业的创新效率。

基于以上分析，本文提出以下假设：

H1：企业的大数据应用程度越高，企业的创新效率越高，大数据应用可以显著提高企业的创新效率。

(二)大数据应用、融资约束和企业创新效率

企业进行创新活动通常需要较长时间才可获得创新产出，在此过程中企业需要持续的资金支持(周开国等, 2017)^[52]。一旦企业停止对创新活动的后续支出，之前的相关投资也将付诸东流(卢馨等, 2013)^[33]。已有研究表明，企业在进行创新活动时往往受到融资约束的限制(万佳彧等, 2020; Milani and Neumann, 2022)^{[36][16]}。企业在创新过程中面临的融资约束主要有两种，其一是“融资难”——由于创新活动风险过高，创新企业难以从外部获取相应融资。此时企业只能使用内部资金来支持创新活动，但企业内部现金流更易受到经营状况和内部管理等因素的影响(张传奇等, 2019)^[50]，极可能出现现金流不足而停止创新活动的情况(Li, 2011; Beladi et al., 2021)^{[12][4]}，创新投入的中断影响企业的创新产出，降低企业创新效率。其二是“融资贵”——进行创新活动使企业经营风险和未来收益的不确定性提高，因此创新企业在融资过程中往往面临较高的融资成本(吴翌琳和黄实磊, 2021)^[42]。融资成本过高容易导致企业偿还压力大，进一步增加企业创新活动的风险，造成融资成本进一步增加的恶性循环。此时，企业为了维持正常的生产经营不得不削减企业的创新活动(Xu, 2020)^[21]，最终导致创新效率低下的结果(王华等, 2020)^[38]。大数据应用能够显著缓解企业面临的融资约束，解决创新活动中遇到的资金瓶颈问题，使企业有充足的资金流进行持续创新(鞠晓生等, 2013)^[27]，从而保障企业的创新产出顺利落地，最终提高创新效率(唐松等, 2020; 王华等, 2020)^{[35][38]}。

具体地，大数据能够从以下两个角度缓解企业面临的融资约束。第一，企业能够通过大数据应用缓解企业面临的“融资难”问题。一方面，“大数据”与实体经

济融合是经济实现高质量发展的重要途径，企业应用大数据向外部投资者释放了一种积极信号，使得企业能够得到外部投资者的关注，从而在一定程度上缓解企业“融资难”的困境。另一方面，应用大数据能够使企业使用较低的成本掌握资金供给方的海量信息；通过使用特定技术工具对数据进行筛选与分析，企业可以更针对性地披露资金供给方所需的信息，有利于提高企业信息披露的针对性，从而更好地满足企业信息使用者的需求(Elliott et al., 2018)^[7]，缓解资金供需双方的信息不对称，从而缓解企业“融资难”困境。第二，大数据应用能够缓解企业面临的“融资贵”问题。一方面，伴随着国家大力发展“数字经济”的东风，企业应用大数据技术能够享受到一系列的政策优惠，从而增加企业以较低成本获得融资的可能性。另一方面，大数据相关技术能够有效提升企业内部的信息分析能力(张嘉伟等, 2022)^[48]，使得企业能够更加快速且透彻地分析企业内部的相关情况，通过提高企业的信息披露质量增强投资者对企业的信心，降低风险预期，从而降低企业的融资成本。

基于以上分析，本文提出以下假设：

H2：企业应用大数据能够缓解面临的融资约束，从而提高创新效率。

三、研究设计

(一)数据

本文以2007—2020年沪深A股上市公司作为研究样本，为保证回归分析的可靠性，对数据进行如下处理(如表1所示)：第一，剔除金融类上市公司；第二，剔除ST和*ST企业；第三，剔除变量存在缺失的企业。经处理后共得到7742个企业-年度观测值。企业的大数据应用指标是使用Python软件爬取巨潮资讯网披露的年报并进行

表1 样本处理

样本处理过程	样本量
最初样本(2007—2020年A股上市公司)	38040
减：	
金融行业	(797)
ST和*ST	(1640)
解释变量 <i>Bdata</i> 缺失	(8070)
被解释变量 <i>Effi</i> 缺失	(14674)
控制变量缺失	(5111)
最终样本	7742

文本分析取得，相关申请和获得专利数量来自中国研究数据服务平台(CNRDS)数据库，其他的相关数据均来自于国泰安(CSMAR)数据库。为缓解异常值导致的结果偏差，本文在1%和99%的水平上对连续变量分别进行缩尾处理。

(二)模型设计

为研究大数据应用对企业创新效率的影响，本文设定如下模型进行检验：

$$Effi_{i,t} = \alpha_0 + \alpha_1 Bdata_{i,t} + \alpha_2 Ctrl_{i,t} + Indcd + Year + \varepsilon_{i,t} \quad (1)$$

其中，创新效率*Effi*为被解释变量，大数据应用*Bdata*为解释变量，*Ctrl*为控制变量，*Indcd*和*Year*表示企业所属的行业和年份固定效应，分别用来控制随行业和时间变化而无法观测的影响， ε 为误差项。本文重点关注系数 α_1 的符号以及显著水平，其经济含义是企业大数据应用对创新效率的影响。

(三)变量设定

1.被解释变量

创新效率*Effi*是指企业资源投入转化为创新产生的能力(Griffin et al., 2018)^[10]。借鉴冯根福等(2017)^[24]和陈德球等(2021)^[23]的研究，本文采用企业取得授权的专利数量(创新产出)与研发投入绝对额的自然对数(创新投入)的比值来衡量企业的创新效率。此外，考虑到进行创新活动从投入创新资源到获取专利成果存在一定的时滞，本文将专利数量滞后3期($t+3$)来构建企业的创新效率代理变量。

2.解释变量

借鉴张叶青等(2021)^[49]的研究，本文将企业年报中的“大数据”词频作为企业大数据应用*Bdata*的代理指标。具体而言，与大数据相关的词汇包括“大数据、海量数据、数据中心、信息资产、数据化、算力”等，这些词汇能够有效衡量企业的大数据应用程度。借助Python爬虫技术，本文归集了沪深A股企业的全部年报并进行词频统计，获得企业大数据应用的初始指标，之后对初始指标进行了对数化处理，最终取得了实证所需的大数据应用指标。

3.控制变量

参考以往文献，本文的控制变量包括：企业规模(*Size*)、资产负债率(*Lev*)、盈利能力(*Roa*)、有形资产比

表2 变量定义

类型	名称	符号	说明
被解释变量	创新效率	<i>Effi</i>	(<i>t</i> +3) 年度获取的专利数量 / (研发投入取对数)
解释变量	大数据应用	<i>Bdata</i>	基于张叶青等(2021) ^[49] 的年报文本分析方法取得
控制变量	企业规模	<i>Size</i>	企业总资产加1取对数
	资产负债率	<i>Lev</i>	企业负债 / 企业资产
	盈利能力	<i>Roa</i>	净利润 / 企业资产
	有形资产比	<i>Fixas</i>	(企业资产-无形资产净值-商誉净值) / 企业资产
	企业成长性	<i>Growth</i>	(本年营业收入-上年营业收入) / 上年营业收入
	企业年龄	<i>Age</i>	企业年龄
	董事会规模	<i>Board</i>	董事会规模取对数
	独立董事比例	<i>Indep</i>	独立董事人数 / 董事会总人数
	管理层持股比例	<i>Manage</i>	管理层持股数量 / 总股数
	股权集中度	<i>Owner</i>	第一大股东持股数量 / 总股数
	现金持有	<i>Cash</i>	企业持有现金的数量(单位:亿元)
	流动性	<i>Liquid</i>	经营性净现金流 / 企业资产
	托宾Q	<i>Tobinq</i>	企业市值 / 企业资产
	两权分离度	<i>Sep</i>	实际控制人拥有控制权-实际控制人拥有所有权
企业性质	<i>Soe</i>	企业性质为国企时取1, 否则取0	

(*Fixas*)、企业成长性(*Growth*)、企业年龄(*Age*)、董事会规模(*Board*)、独立董事比例(*Indep*)、管理层持股比例(*Manage*)、股权集中度(*Owner*)、现金持有(*Cash*)、流动性(*Liquid*)、托宾Q(*Tobinq*)、两权分离度(*Sep*)和企业性质(*Soe*)。

变量定义如表2所示。

四、实证结果与分析

(一)变量描述性统计

表3报告了主要变量的描述性统计结果。其中,创新效率的均值为1.298,中位数为0.236,标准差为3.480,最小值为0,最大值为24.710,说明样本企业的整体创新效率偏低,且公司之间的创新效率存在较大差距。大数据应用指标的均值为0.351,中位数为0,最小值为0,最大值为4.754,说明样本中大部分企业并未意识到大数据对于自身发展的作用,大数据应用水平整体偏低;标准差为0.760,说明样本企业的大数据应用程度存在较大差异。其他变量均处于正常范围内。

(二)基准回归结果

表4报告了大数据应用-企业创新效率关系的核心回归结果。在基准回归中,本文采用递进式的回归方式。

表3 主要变量的描述性统计结果

变量	样本量	平均值	中位数	标准差	最小值	最大值
<i>Effi</i>	7742	1.298	0.236	3.480	0.000	24.710
<i>Bdata</i>	7742	0.351	0.000	0.760	0.000	4.754
<i>Size</i>	7742	22.290	22.110	1.246	20.080	26.310
<i>Lev</i>	7742	0.411	0.408	0.188	0.055	0.821
<i>Roa</i>	7742	0.070	0.060	0.046	0.002	0.237
<i>Fixas</i>	7742	0.931	0.955	0.075	0.572	0.999
<i>Growth</i>	7742	0.291	0.135	0.656	-0.556	4.241
<i>Age</i>	7742	9.016	7.000	6.400	1.000	24.000
<i>Board</i>	7742	2.161	2.197	0.191	1.609	2.708
<i>Indep</i>	7742	37.100	33.330	5.187	33.330	57.140
<i>Manage</i>	7742	12.310	0.337	19.340	0.000	69.300
<i>Owner</i>	7742	36.450	34.990	14.790	9.440	75.460
<i>Cash</i>	7742	19.450	6.371	45.550	0.545	341.400
<i>Liquid</i>	7742	0.053	0.050	0.066	-0.130	0.236
<i>Tobinq</i>	7742	2.054	1.700	1.132	0.927	7.168
<i>Sep</i>	7742	5.360	0.000	7.931	0.000	28.400
<i>Soe</i>	7742	0.129	0.000	0.336	0.000	1.000

首先进行大数据应用和企业创新效率的单变量回归,结果如第(1)列所示,大数据应用指标(*Bdata*)的系数为0.274且在1%水平下显著;其次,在原有基础上加入控制变量进行回归,结果如第(2)列所示,*Bdata*的系数为0.271且在1%水平下显著;最后,进一步控制行业和年份固定效应进行回归,结果如第(3)列所示,*Bdata*的系数为0.372且在1%水平下显著。这意味着,企业大数据应用程度越高,企业的创新效率越高,两者之间呈现显著的正相关关系。综上,本文假设1得到了经验证据支持。

(三)稳健性检验

1.更换核心变量的衡量指标

第一,更换大数据应用的衡量方式。考虑到年报中的词频数据可能存在一定的噪音,本文借鉴张叶青等(2021)^[49]的研究方法,构建大数据应用的虚拟变量*Bdata_dum*进行验证,企业当年年报中出现“大数据”相关词汇时*Bdata_dum*取值1,否则为0。此外,本文选择了更广义的关键词集合构建大数据应用的代理指标*Bdata_exp*,除前文提及的大数据应用相关词汇外,借鉴吴非等(2021)^[41]的研究,加入大数据技术相关的关键词来构建企业大数据应用指标,这些关键词包括:“大数据、数据挖掘、文本挖掘、数据可视化、异构数据、征信、增强现实、混合现实、虚拟现实”。更换大数据应用指标的回归结果如表5第(1)(2)列所示,*Bdata_dum*和*Bdata_exp*的系数分别

表4 大数据应用与企业创新效率

	(1)	(2)	(3)
	<i>Effi</i>	<i>Effi</i>	<i>Effi</i>
<i>Bdata</i>	0.274*** (5.282)	0.271*** (5.244)	0.372*** (6.371)
<i>Size</i>		0.508*** (9.267)	0.602*** (10.339)
<i>Lev</i>		0.141 (0.542)	0.192 (0.713)
<i>Roa</i>		3.982*** (3.954)	5.686*** (5.520)
<i>Fixas</i>		2.840*** (5.539)	1.466*** (2.857)
<i>Growth</i>		-0.240*** (-4.118)	-0.156*** (-2.617)
<i>Age</i>		-0.050*** (-6.579)	-0.037*** (-4.850)
<i>Board</i>		1.625*** (6.879)	1.874*** (8.059)
<i>Indep</i>		0.016* (1.938)	0.024*** (2.977)
<i>Manage</i>		0.004 (1.428)	0.002 (0.892)
<i>Owner</i>		-0.003 (-1.155)	-0.000 (-0.069)
<i>Cash</i>		0.011*** (10.033)	0.011*** (9.274)
<i>Liquid</i>		2.005*** (3.072)	2.897*** (4.444)
<i>Tobinq</i>		0.057 (1.425)	0.073* (1.721)
<i>Sep</i>		0.006 (1.191)	0.003 (0.566)
<i>Soe</i>		0.106 (0.895)	0.088 (0.743)
截距项	1.201*** (27.625)	-17.118*** (-12.215)	-19.604*** (-12.980)
年度/行业	未控制	未控制	控制
样本量	7742	7742	7742
调整 R^2	0.003	0.108	0.172

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

为0.628和0.378，且均在1%水平下显著，说明本文的主要结论是稳健的。

第二，更换创新效率的衡量方式。首先，使用申请专利数量衡量企业的创新产出，重新构建创新效率的代理变量 *Effi1* 进行验证；其次，我国的专利分为发明专利、外观设计专利和实用新型专利三种，其中发明专利的质量最高，因此，本文以发明专利衡量企业的创新产出，构建创新效率的代理指标 *Effi2* 进行验证；最后，分别使用滞后1期、2期的企业取得授权专利数量与研发投入绝对值的自然对数的比值构建创新效率的代理指标 *Effi3* 和 *Effi4* 进行验证。更换创新效率衡量方式的回归结果如表5第(3)~(6)列所示，*Bdata* 的系数分别为0.725、0.262、0.428和0.272，且均在1%水平下显著，说明本文

表5 稳健性检验：更换核心变量、样本区间和回归模型

	更换解释变量		更换被解释变量				更换样本区间	更换回归模型
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	<i>Effi</i>	<i>Effi</i>	<i>Effi1</i>	<i>Effi2</i>	<i>Effi3</i>	<i>Effi4</i>	<i>Effi</i>	<i>Effi</i>
<i>Bdata</i>			0.725*** (4.806)	0.262*** (4.672)	0.428*** (4.985)	0.272*** (5.988)	0.695*** (6.034)	0.187*** (2.750)
<i>Bdata_dum</i>	0.628*** (6.470)							
<i>Bdata_exp</i>		0.378*** (6.656)						
<i>L.Eff</i>								0.532*** (6.015)
<i>Size</i>	0.599*** (10.278)	0.597*** (10.246)	0.729*** (4.839)	0.195*** (3.492)	1.091*** (13.931)	0.479*** (11.773)	0.646*** (8.627)	0.316*** (3.238)
<i>Lev</i>	0.229 (0.850)	0.192 (0.712)	0.553 (0.794)	-0.084 (-0.326)	-0.554 (-1.288)	-0.505** (-2.292)	0.190 (0.552)	0.098 (0.413)
<i>Roa</i>	5.696*** (5.530)	5.733*** (5.568)	5.760** (2.162)	-1.284 (-1.298)	-0.312 (-0.203)	-2.078*** (-2.640)	5.358*** (3.918)	3.400*** (3.154)
<i>Fixas</i>	1.370*** (2.674)	1.483*** (2.889)	3.735*** (2.814)	1.130** (2.292)	2.492*** (3.358)	1.144*** (3.008)	0.634 (0.729)	1.011*** (2.815)
<i>Growth</i>	-0.149** (-2.500)	-0.156*** (-2.618)	-0.196 (-1.273)	-0.034 (-0.586)	-0.018 (-0.778)	-0.001 (-0.123)	-0.092 (-1.154)	-0.095*** (-2.605)
<i>Age</i>	-0.036** (-4.790)	-0.036*** (-4.832)	-0.055*** (-2.804)	-0.007 (-0.960)	-0.049*** (-4.313)	-0.010 (-1.631)	-0.032*** (-3.003)	-0.020* (-1.805)
<i>Board</i>	1.874*** (8.063)	1.880*** (8.089)	3.540*** (5.887)	1.534*** (6.873)	1.646*** (4.376)	0.985*** (5.157)	1.922*** (6.418)	1.002*** (2.782)
<i>Indep</i>	0.025*** (3.003)	0.025*** (3.004)	0.023 (1.090)	0.016** (1.992)	0.004 (0.334)	0.004 (0.591)	0.021** (2.034)	0.008 (0.920)
<i>Manage</i>	0.002 (0.879)	0.002 (0.866)	0.008 (1.225)	0.004* (1.696)	0.006 (1.638)	0.004** (2.195)	0.002 (0.617)	0.001 (0.599)
<i>Owner</i>	-0.001 (-0.220)	-0.000 (-0.004)	-0.009 (-1.250)	-0.005** (-2.047)	-0.003 (-0.730)	-0.005** (-2.053)	0.004 (1.220)	0.001 (0.363)
<i>Cash</i>	0.011*** (9.290)	0.011*** (9.301)	0.037*** (12.720)	0.016*** (14.813)	0.008*** (12.482)	0.004*** (12.293)	0.009*** (5.368)	0.005* (1.771)
<i>Liquid</i>	2.833*** (4.349)	2.889*** (4.434)	4.690*** (2.782)	0.807 (1.289)	3.287*** (3.174)	1.173** (2.238)	1.333 (1.600)	1.325** (2.558)
<i>Tobinq</i>	0.071* (1.677)	0.070* (1.651)	0.125 (1.136)	0.072* (1.755)	0.121* (1.915)	0.087*** (2.691)	0.045 (0.740)	0.037 (1.202)
<i>Sep</i>	0.003 (0.510)	0.003 (0.581)	0.028** (2.177)	0.009* (1.919)	0.007 (0.853)	0.005 (1.243)	0.004 (0.567)	0.000 (0.028)
<i>Soe</i>	0.089 (0.753)	0.089 (0.755)	0.110 (0.360)	-0.244** (-2.153)	-0.008 (-0.041)	-0.246** (-2.474)	0.150 (1.008)	0.125 (0.756)
截距项	-19.504*** (-12.909)	-19.525*** (-12.928)	-25.112*** (-6.429)	-7.454*** (-5.139)	-28.967*** (-13.452)	-12.878*** (-12.254)	-19.817*** (-10.158)	-10.433*** (-3.844)
年度/行业	控制	控制	控制	控制	控制	控制	控制	控制
样本量	7742	7742	7742	7742	10275	9255	4286	5171
调整 R^2	0.172	0.173	0.108	0.094	0.109	0.088	0.171	-
AR(2) ρ 值								0.556
Hansen 检验 ρ 值								0.892

的主要结论是稳健的。

2. 更换样本区间

为避免企业由于追求热点而在年报中夸大披露大数据应用的相关信息而导致回归系数虚增，借鉴张叶青等(2021)^[49]的研究，本文仅使用2014年及以前的大数据发展早期样本进行验证。结果如表5第(7)列所示，*Bdata* 的系数为0.695，且在1%水平下显著，说明本文的主要结论是稳健的。

3. 更换回归模型

由于企业的创新效率可能存在动态效应，即前期创新效率可能会对后期创新效率产生冲击，导致静态面板模型的估计结果有偏，因此，本文基于动态面板回归模型进行验证。结果如表5第(8)列所示， $Bdata$ 的系数为0.187，且在1%水平下显著，说明本文的主要结论是稳健的。

4. 内生性处理

本文可能存在以下内生性问题：第一，遗漏变量。可能有一些难以量化的因素同时与企业的大数据应用和创新效率相关，例如企业的大数据应用和创新效率可能同时受到地区政策的激励，造成回归系数被高估。第二，反向因果。创新效率比较高的企业对新技术的关注度更高，更有动机投入大数据应用相关建设，导致回归系数被高估。第三，自选择偏误。大数据应用程度较高的企业可能本身各方面条件更优越、表现更佳，因而能够更有效率地开展创新活动，创新效率更高。

本文主要采取以下措施来缓解可能存在的内生性问题：第一，在控制行业和年份固定效应的基础上，通过分别进一步控制省份地区固定效应和城市地区固定效应来缓解可能存在的遗漏变量内生性问题。回归结果如表6第(1)(2)列所示，可以看出 $Bdata$ 的系数仍然显著为正，本文的研究结论是稳健的。第二，借鉴李唐等(2020)^[30]和宋德勇等(2022)^[34]的研究，本文为企业大数据应用水平构造工具变量进行工具变量法回归，以缓解可能存在的反向因果内生性问题。具体而言，采用大数据应用水平与按行业和省份分类的大数据应用水平均值差额的三次方作为工具变量($Bdata_IV$)。回归结果如表6第(3)(4)列所示，在第一阶段回归中 $Bdata_IV$ 的系数为0.143且在1%水平下显著，且 F 统计量值(163.950)远大于经验规则临界值10，说明本文选取的工具变量不存在弱工具变量问题。在第二阶段回归中 $Bdata$ 的系数为0.312且在1%水平下显著。以上回归结果说明本文的研究结论是稳健的。第三，本文采用倾向得分匹配法(PSM)来减轻可能存在的自选择偏误。具体而言，以样本企业是否应用大数据技术将其划分为处理组和对照组，使用Stata软件中的`pselect`命令从控制变量中筛选出有效的协变量，按照1:1近邻匹配法为处理组中的上市公司进行有放回的匹配。回归结果如表6第(5)列所示， $Bdata$ 的系数为0.300且

在1%水平下显著，这进一步说明了本文的研究结论是稳健的。

(四) 机制检验

为检验融资约束是否是大数据应用影响企业创新效率的内在机制，本文借鉴温忠麟等(2004)^[40]提出的中介效应检验程序进行机制检验，具体的检验模型如式(2)(3)和(4)所示。模型(2)反映大数据应用对企业创新效率的总效应；模型(3)中 γ_1 反映了大数据应用对融资约束的影响；模型(4)中 θ_1 反映了大数据应用对企业创新效率的直接效应；

表6 稳健性检验：内生性处理

	固定效应		工具变量法		倾向得分匹配法
	(1)	(2)	(3)	(4)	(5)
	$Effi$	$Effi$	$Bdata$	$Effi$	$Effi$
$Bdata_IV$			0.143*** (54.460)		
$Bdata$	0.331*** (4.508)	0.291*** (3.838)		0.312*** (2.848)	0.300*** (3.522)
$Size$	0.567*** (6.002)	0.545*** (5.401)	0.060*** (3.825)	0.590*** (6.280)	0.666*** (6.436)
Lev	0.028 (0.065)	0.119 (0.257)	-0.047 (-0.626)	0.139 (0.311)	0.353 (0.709)
Roa	5.530*** (3.457)	6.917*** (4.134)	0.108 (0.389)	6.652*** (4.025)	8.303*** (4.391)
$Fixas$	2.187*** (3.301)	2.256*** (3.211)	-0.215** (-1.967)	1.806*** (2.767)	2.451*** (3.026)
$Growth$	-0.165* (-1.814)	-0.131 (-1.386)	0.052*** (3.302)	-0.210** (-2.247)	-0.174* (-1.793)
Age	-0.028** (-2.550)	-0.029** (-2.410)	0.003* (1.842)	-0.039*** (-3.559)	-0.037*** (-2.832)
$Board$	1.819*** (4.897)	2.162*** (5.487)	0.063 (0.991)	1.932*** (5.124)	1.835*** (4.512)
$Indep$	0.027** (2.054)	0.026* (1.892)	0.002 (1.015)	0.027** (2.002)	0.025* (1.757)
$Manage$	0.004 (0.906)	0.002 (0.372)	0.001** (2.117)	0.003 (0.658)	0.003 (0.616)
$Owner$	-0.004 (-0.860)	-0.008* (-1.653)	0.001 (1.495)	-0.008* (-1.831)	0.002 (0.458)
$Cash$	0.013*** (7.693)	0.014*** (7.964)	-0.000 (-0.157)	0.010*** (7.533)	0.012*** (6.626)
$Liquid$	4.616*** (4.447)	4.331*** (4.057)	-0.148 (-0.835)	4.702*** (4.444)	3.279*** (2.730)
$Tobinq$	0.119* (1.904)	0.075 (1.148)	0.006 (0.606)	0.073 (1.188)	0.030 (0.423)
Sep	-0.004 (-0.526)	-0.005 (-0.539)	-0.002 (-1.201)	0.001 (0.119)	0.008 (0.842)
Soe	0.019 (0.100)	-0.092 (-0.458)	-0.087*** (-2.646)	-0.000 (-0.002)	-0.035 (-0.163)
截距项	-20.519*** (-8.402)	-20.455*** (-7.897)	-1.205*** (-3.017)	-19.861*** (-8.325)	-22.732*** (-7.892)
年度/行业	控制	控制	控制	控制	控制
省份	控制	未控制	未控制	未控制	未控制
市	未控制	控制	未控制	未控制	未控制
样本量	3456	3456	3456	3456	2874
F 值			163.950		
调整 R^2	0.192	0.223	0.636	0.175	0.213

模型(3)中 γ_1 和模型(4)中 θ_2 的乘积 $\gamma_1\theta_2$ 反映了融资约束的中介效应。

$$Effi_{i,t} = \beta_0 + \beta_1 Bdata_{i,t} + \beta_2 Ctrl_{i,t} + Indcd + Year + \delta_{i,t} \quad (2)$$

$$Fc_{i,t} = \gamma_0 + \gamma_1 Bdata_{i,t} + \gamma_2 Ctrl_{i,t} + Indcd + Year + \vartheta_{i,t} \quad (3)$$

$$Effi_{i,t} = \theta_0 + \theta_1 Bdata_{i,t} + \theta_2 Fc_{i,t} + \theta_3 Ctrl_{i,t} + Indcd + Year + \tau_{i,t} \quad (4)$$

借鉴魏志华等(2014)^[39]和唐松等(2020)^[35]的研究,

模型(3)中融资约束变量 Fc 的计算方法如下:首先,按照年度对经营性净现金流比年初总资产、现金股利比年初总资产、现金持有比年初总资产、资产负债率和托宾Q等五个变量进行分类,分别根据变量对样本进行升序排序后以中位数作为融资约束的分界点确定融资约束虚拟变量 Fc_i ,将5个变量的 Fc_i 加总后得到每个样本的 Fc_qu 指

数。其次,使用式(5)进行排序Logit回归,其中, $Cashdiv$ 表示企业当年发放的现金股利, Cf 表示企业的净经营性现金流, Ta 表示企业的资产总额,其余变量的设定与前文一致。最终得到每个样本企业在相应年份的融资约束代理变量 Fc , Fc 的值越大代表企业面临的融资约束程度越高。

$$Fc_{i,t} = \mu_1 \frac{Cf_{i,t}}{Ta_{i,t-1}} + \mu_2 Lev_{i,t} + \mu_3 \frac{Cashdiv_{i,t}}{Ta_{i,t-1}} + \mu_4 \frac{Cash_{i,t}}{Ta_{i,t-1}} + \mu_5 Tobinq_{i,t} + \omega_{i,t} \quad (5)$$

中介检验的回归结果如表7所示。第(1)列显示,大数据应用能够显著提升企业创新效率;第(2)列显示,大数据应用能够显著缓解企业面临的融资约束;第(3)列显示,融资约束的系数显著为负,且大数据应用的系数绝对值相较第(1)列有所降低,说明大数据应用通过缓解融资约束,显著提升企业的创新效率。因此,大数据应用—融资约束—创新效率的机制成立,且融资约束在大数据应用对创新效率的影响中起部分中介作用。由此,本文的假设2得到了经验证据支持。

为进一步检验融资约束中介效应回归结果的稳健性,本文通过更换变量的度量方式进行验证。首先,改变大数据应用和企业创新效率的度量方式,分别将模型中的 $Bdata$ 和 $Effi$ 替换为 $Bdata_exp$ 和 $Effi1$ 进行检验。检验结果如表8所示,融资约束的中介效应仍然成立。其次,改变融资约束的度量方式,借鉴李文秀和唐荣(2021)^[31]的研究,使用利息支出与流动负债之比衡量企业面临的融资约束($Fc1$)进行检验。检验结果如表9所示,再次表明融资约束的中介效应回归结果是稳健的。

(五)异质性分析

大数据应用对不同科技属性企业的创新效率会有不同的影响。科技企业作为技术创新活动的主力军,是我国实现可持续发展、促进经济转型的关键因素。一方面,相较于其他企业,科技企业通常面临较为严峻的技术竞争环境,因此更有动机利用大数据技术提升创新产出和创新效率(翟淑萍等,2020)^[47]。另一方面,科技企业拥有更多技术人才,使得其更能把握科学技术的发展前景,认识到大数据技术的发展方向,对大数据的应用能力与应用范围处于领先地位。基于以上分析,本文认为大数据应用对企业创新效率的影响将在科技企业中更加明显。

按照企业是否属于科技企业进行分组,回归结果如

表7 融资约束的中介机制检验

	(1)	(2)	(3)
	<i>Effi</i>	<i>Fc</i>	<i>Effi</i>
<i>Bdata</i>	0.344*** (5.828)	-0.058*** (-2.793)	0.336*** (5.693)
<i>Fc</i>			-0.142*** (-4.312)
<i>Size</i>	0.598*** (10.197)	-0.010 (-0.465)	0.597*** (10.186)
<i>Lev</i>	0.175 (0.642)	6.402*** (67.399)	1.082*** (3.150)
<i>Roa</i>	5.752*** (5.462)	-8.022*** (-21.802)	4.615*** (4.256)
<i>Fixas</i>	1.497*** (2.889)	-0.586*** (-3.235)	1.414*** (2.731)
<i>Growth</i>	-0.155*** (-2.578)	-0.110*** (-5.216)	-0.171*** (-2.834)
<i>Age</i>	-0.035*** (-4.670)	0.006** (2.205)	-0.035*** (-4.565)
<i>Board</i>	1.945*** (8.310)	0.108 (1.321)	1.960*** (8.384)
<i>Indep</i>	0.027*** (3.235)	0.007** (2.519)	0.028*** (3.363)
<i>Manage</i>	0.002 (0.870)	0.000 (0.161)	0.002 (0.879)
<i>Owner</i>	-0.001 (-0.228)	0.000 (0.430)	-0.001 (-0.207)
<i>Cash</i>	0.011*** (9.638)	-0.006*** (-14.581)	0.010*** (8.803)
<i>Liquid</i>	2.777*** (4.160)	-16.215*** (-69.508)	0.479 (0.561)
<i>Tobinq</i>	0.074* (1.717)	0.341*** (22.641)	0.122*** (2.750)
<i>Sep</i>	0.001 (0.232)	-0.007*** (-4.025)	0.000 (0.033)
<i>Soe</i>	0.080 (0.672)	0.007 (0.178)	0.081 (0.681)
截距项	-19.742*** (-13.000)	-0.545 (-1.027)	-19.819*** (-13.065)
年度/行业	控制	控制	控制
样本量	7614	7614	7614
调整R ²	0.173	0.719	0.175

表8 融资约束的中介机制稳健性检验：更换解释变量和被解释变量

	更换解释变量			更换被解释变量		
	(1)	(2)	(3)	(4)	(5)	(6)
	<i>Effi</i>	<i>Fc</i>	<i>Effi</i>	<i>Effi1</i>	<i>Fc</i>	<i>Effi1</i>
<i>Bdata_exp</i>	0.352*** (6.131)	-0.061*** (-3.063)	0.343*** (5.984)			
<i>Bdata</i>				0.659*** (4.302)	-0.058*** (-2.793)	0.650*** (4.245)
<i>Fc</i>			-0.141*** (-4.285)			-0.147* (-1.724)
<i>Size</i>	0.593*** (10.105)	-0.008 (-0.414)	0.592*** (10.096)	0.720*** (4.723)	-0.010 (-0.465)	0.718*** (4.715)
<i>Lev</i>	0.174 (0.642)	6.402*** (67.409)	1.076*** (3.132)	0.428 (0.607)	6.402*** (67.399)	1.371 (1.536)
<i>Roa</i>	5.797*** (5.507)	-8.029*** (-21.826)	4.667*** (4.305)	5.899** (2.157)	-8.022*** (-21.802)	4.718* (1.674)
<i>Fixas</i>	1.515*** (2.923)	-0.590*** (-3.258)	1.432*** (2.764)	3.814*** (2.834)	-0.586*** (-3.235)	3.727*** (2.768)
<i>Growth</i>	-0.155*** (-2.584)	-0.109*** (-5.206)	-0.171*** (-2.838)	-0.182 (-1.164)	-0.110*** (-5.216)	-0.198 (-1.265)
<i>Age</i>	-0.035*** (-4.651)	0.006** (2.197)	-0.035*** (-4.547)	-0.052*** (-2.649)	0.006** (2.205)	-0.051*** (-2.604)
<i>Board</i>	1.952*** (8.341)	0.107 (1.303)	1.967*** (8.414)	3.693*** (6.078)	0.108 (1.321)	3.709*** (6.104)
<i>Indep</i>	0.027*** (3.261)	0.007** (2.508)	0.028*** (3.387)	0.028 (1.302)	0.007** (2.519)	0.029 (1.352)
<i>Manage</i>	0.002 (0.845)	0.000 (0.177)	0.002 (0.854)	0.008 (1.202)	0.000 (0.161)	0.008 (1.205)
<i>Owner</i>	-0.000 (-0.167)	0.000 (0.399)	-0.000 (-0.148)	-0.009 (-1.335)	0.000 (0.430)	-0.009 (-1.326)
<i>Cash</i>	0.011*** (9.660)	-0.006*** (-14.591)	0.010*** (8.829)	0.039*** (12.991)	-0.006*** (-14.581)	0.038*** (12.529)
<i>Liquid</i>	2.770*** (4.151)	-16.215*** (-69.523)	0.487 (0.571)	4.954*** (2.858)	-16.215*** (-69.508)	2.567 (1.157)
<i>Tobinq</i>	0.071* (1.649)	0.342*** (22.673)	0.119*** (2.678)	0.119 (1.064)	0.341*** (22.641)	0.169 (1.464)
<i>Sep</i>	0.001 (0.245)	-0.007*** (-4.033)	0.000 (0.046)	0.026** (1.974)	-0.007*** (-4.025)	0.025* (1.892)
<i>Soe</i>	0.081 (0.684)	0.007 (0.169)	0.082 (0.693)	0.089 (0.290)	0.007 (0.178)	0.090 (0.293)
截距项	-19.665*** (-12.950)	-0.561 (-1.056)	-19.744*** (-13.015)	-25.444*** (-6.453)	-0.545 (-1.027)	-25.524*** (-6.473)
年度/行业	控制	控制	控制	控制	控制	控制
样本量	7614	7614	7614	7614	7614	7614
调整 R ²	0.173	0.719	0.175	0.108	0.719	0.109

表10第(1)(2)列所示。科技企业组*Bdata*的系数为0.413且在1%水平下显著，非科技企业组*Bdata*的系数为0.090但不显著；并且，分组回归结果的Suest检验*p*值为0.005，通过了组间系数差异检验。引入科技属性(*Tech*)和*Bdata*交互项的回归结果如表10第(3)列所示，交互项的估计系数为0.487且在1%水平下显著，说明在科技企业中大数据应用对创新效率的促进作用更为显著。

大数据应用对行业竞争程度不同的企业的创新效率会有不同影响。一方面，为了能更好地生存和发展，处于行业竞争激烈环境的企业更有动机利用大数据进行高效的创新活动(何玉润等, 2015)^[25]，从而提升企业的核

表9 融资约束的中介机制稳健性检验：更换融资约束变量

	(1)	(2)	(3)
	<i>Effi</i>	<i>Fc1</i>	<i>Effi</i>
<i>Bdata</i>	0.375*** (6.264)	-0.001** (-2.350)	0.363*** (6.084)
<i>Fc1</i>			-9.972*** (-7.340)
<i>Size</i>	0.610*** (10.198)	0.001 (1.177)	0.616*** (10.334)
<i>Lev</i>	0.211 (0.744)	0.092*** (38.348)	1.133*** (3.663)
<i>Roa</i>	5.922*** (5.532)	0.051*** (5.611)	6.431*** (6.016)
<i>Fixas</i>	1.488*** (2.815)	-0.027*** (-5.913)	1.223** (2.316)
<i>Growth</i>	-0.157** (-2.563)	-0.002*** (-3.679)	-0.176*** (-2.882)
<i>Age</i>	-0.037*** (-4.794)	-0.000 (-0.810)	-0.037*** (-4.880)
<i>Board</i>	1.890*** (7.904)	-0.001 (-0.613)	1.878*** (7.880)
<i>Indep</i>	0.025*** (2.930)	0.000 (0.753)	0.025*** (3.004)
<i>Manage</i>	0.002 (0.842)	-0.000*** (-6.469)	0.001 (0.294)
<i>Owner</i>	-0.001 (-0.261)	-0.000*** (-9.458)	-0.003 (-1.060)
<i>Cash</i>	0.011*** (9.111)	-0.000*** (-10.217)	0.010*** (8.217)
<i>Liquid</i>	2.936*** (4.396)	-0.014** (-2.496)	2.794*** (4.197)
<i>Tobinq</i>	0.079* (1.761)	-0.003*** (-7.425)	0.051 (1.132)
<i>Sep</i>	0.002 (0.470)	-0.000 (-1.512)	0.002 (0.343)
<i>Soe</i>	0.089 (0.740)	-0.005*** (-4.879)	0.039 (0.327)
截距项	-19.868*** (-12.836)	0.024* (1.853)	-19.624*** (-12.721)
年度/行业	控制	控制	控制
样本量	7495	7495	7495
调整 R ²	0.172	0.339	0.178

心竞争力。另一方面，竞争激烈的行业内往往具有更多的外部信息，而大量的信息使得企业的创新决策更加精准，进而提升企业的创新产出，增加创新效率。因此，大数据应用对企业创新效率的提升作用将在行业竞争程度更高的企业中更为明显。

本文使用赫芬达尔指数度量企业所在行业的竞争程度，把样本按照是否大于中位数分为行业竞争高和行业竞争低两组分别进行回归。结果如表10第(4)(5)列所示，行业竞争程度高组*Bdata*的系数为0.413且在1%水平下显著，行业竞争程度低组*Bdata*的系数为0.159但仅在10%水平下显著；并且，分组回归结果的Suest检验*p*值为0.054，通过了组间系数差异检验。引入行业竞争(*Mcomp*)和*Bdata*

交互项的回归结果如表10第(6)列所示，回归结果显示交互项的估计系数为0.424且在1%水平下显著，说明在竞争

行业中大数据应用对创新效率的提升作用更为显著。

五、结论与启示

本文通过爬虫技术爬取沪深A股上市公司年报，构建大数据应用相关指标，研究大数据应用对企业创新效率的影响，得到以下主要研究结论：首先，大数据应用对企业创新效率具有显著的提升作用，且该结论在更换核心变量衡量方式、改变样本区间和处理内生性问题后依然成立。其次，大数据应用能够有效缓解企业面临的融资约束问题，这有助于提高企业的创新动机与创新活动的可持续性，从而促进企业的创新效率。最后，异质性检验结果表明，大数据应用对企业创新效率的提升效应在科技企业组和行业竞争程度高组中相对更明显。

本文的研究具有以下政策启示：第一，大数据技术应用能够显著提升企业创新效率，推动企业数字化应用对实现高质量发展具有重要意义。一方面，企业应抓住数字经济的发展态势，把握机遇，落实大数据分析技术在企业决策和管理等过程中的应用，构建以大数据技术为基础的高质量创新驱动发展体系。另一方面，政府应加快数据基础设施建设，营造良好数据环境，提升数据质量，助力大数据与实体经济的高效融合。第二，应借助大数据应用技术完善企业信息披露制度，提高企业信息披露质量，从而强化企业与投资者之间的信息传递效率，降低企业内外部信息不对称程度，助力企业实现资金融通，缓解企业面临的融资约束。第三，鉴于不同属性的企业中大数据应用提高创新效率的差异，政府在制定相关政策时应当结合企业的科技属性和所处行业等因素，针对不同企业特定情境制定因势利导的政策体系。

[基金项目：国家自然科学基金面上项目“时序赋权与空间近邻双重视角下基于文本分析的多分类器动态集成企业信用评估研究”(批准号：72271174)、“多类别非平衡企业信用评估的多SVM集成建模研究”(批准号：71771162)]

表 10 异质性分析：区分科技属性和行业属性

	科技企业	非科技企业	全样本	行业竞争程度高	行业竞争程度低	全样本
	(1)	(2)	(3)	(4)	(5)	(6)
	Eff	Eff	Eff	Eff	Eff	Eff
Bdata	0.413*** (5.454)	0.090 (1.205)	0.005 (0.048)	0.413*** (5.488)	0.159* (1.708)	0.111 (1.353)
Bdata×Tech			0.487*** (3.984)			
Bdata×Mcomp						0.424*** (4.533)
Size	0.791*** (9.544)	0.198*** (3.147)	0.602*** (10.343)	0.530*** (6.699)	0.676*** (7.794)	0.607*** (10.438)
Lev	0.742** (2.022)	-1.085*** (-3.538)	0.180 (0.671)	0.901** (2.547)	-0.710* (-1.722)	0.172 (0.639)
Roa	8.350*** (5.904)	1.399 (1.202)	5.684*** (5.523)	8.080*** (5.849)	3.738** (2.428)	5.703*** (5.543)
Fixas	1.314* (1.792)	1.765*** (3.261)	1.456*** (2.840)	1.709** (2.479)	0.830 (1.061)	1.443*** (2.816)
Growth	-0.283*** (-3.049)	-0.068 (-1.228)	-0.156*** (-2.624)	-0.144* (-1.681)	-0.162* (-1.941)	-0.164*** (-2.752)
Age	-0.065*** (-6.042)	-0.003 (-0.368)	-0.037*** (-4.882)	-0.034*** (-3.320)	-0.038*** (-3.426)	-0.038*** (-5.005)
Board	2.064*** (6.345)	1.090*** (4.320)	1.874*** (8.068)	1.999*** (6.227)	1.728*** (5.151)	1.884*** (8.112)
Indep	0.040*** (3.354)	0.000 (0.023)	0.024*** (2.946)	0.048*** (4.265)	0.002 (0.137)	0.025*** (2.993)
Manage	0.002 (0.647)	0.002 (0.690)	0.002 (0.956)	0.002 (0.639)	0.001 (0.401)	0.002 (0.800)
Owner	-0.007* (-1.917)	0.014*** (4.837)	-0.000 (-0.082)	-0.007** (-1.970)	0.010** (2.553)	0.000 (0.009)
Cash	0.021*** (11.178)	0.005*** (4.818)	0.011*** (9.356)	0.019*** (11.020)	0.006*** (3.664)	0.010*** (9.190)
Liquid	3.399*** (3.642)	2.290*** (3.373)	2.957*** (4.540)	2.056** (2.241)	3.752*** (4.062)	3.025*** (4.643)
Tobinq	0.137** (2.517)	-0.067 (-1.135)	0.068 (1.601)	0.101* (1.882)	0.017 (0.237)	0.063 (1.488)
Sep	0.000 (0.055)	-0.002 (-0.379)	0.003 (0.560)	-0.002 (-0.363)	0.005 (0.674)	0.003 (0.548)
Soe	0.225 (1.358)	-0.408*** (-3.232)	0.072 (0.610)	0.096 (0.607)	0.030 (0.172)	0.069 (0.584)
截距项	-23.564*** (-11.329)	-8.337*** (-5.223)	-19.467*** (-12.899)	-20.565*** (-9.606)	-18.580*** (-8.566)	-19.605*** (-12.998)
年度/行业	控制	控制	控制	控制	控制	控制
样本量	4984	2758	7742	4068	3674	7742
调整 R ²	0.219	0.125	0.174	0.219	0.147	0.174
Suest 检验 -p 值	0.005		0.054			

参考文献：

[1] Aboelmaged M G. Linking operations performance to knowledge management capability: the mediating role of innovation performance[J]. Production Planning & Control, 2014, 25(1): 44-58.

[2] Bao J, Qu Y, Zhao S, Zheng N. The role of big data-based precision marketing in firm performance[J]. International Journal of Entertainment Technology and Management, 2022, 1(3): 246-271.

[3] Barton D, Court D. Making advanced analytics work for you[J]. Harvard Business Review, 2012, 90(10): 78-83.

[4] Beladi H, Deng J, Hu M. Cash flow uncertainty, financial constraints and R&D investment[J]. International Review of Financial Analysis, 2021, 76: 101785.

[5] Cinar Y G, Zoghbi S, Moens M F. Inferring user interests on social media from text and images[C]/2015 IEEE International Conference on Data Mining Workshop (ICDMW). IEEE, 2015: 1342-1347.

[6] Davenport T H, De Long D W, Beers M C. Successful knowledge management projects[J]. MIT Sloan Management Review, 1998, 39(2):

43-57.

- [7] 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.
- [8] Fisher R J, Maltz E, Jaworski B J. Enhancing communication between marketing and engineering: the moderating role of relative functional identification[J]. *Journal of Marketing*, 1997, 61(3): 54-70.
- [9] Ghasemaghaei M, Calic G. Assessing the impact of big data on firm innovation performance: big data is not always better data[J]. *Journal of Business Research*, 2020, 108: 147-162.
- [10] Griffin P A, Hong H A, Ryou J W. Corporate innovative efficiency: evidence of effects on credit ratings[J]. *Journal of Corporate Finance*, 2018, 51: 352-373.
- [11] Lee J H, Kim Y G. A stage model of organizational knowledge management: a latent content analysis[J]. *Expert Systems with Applications*, 2001, 20(4): 299-311.
- [12] Li D. Financial constraints, R&D investment, and stock returns[J]. *Review of Financial Studies*, 2011, 24(9): 2974-3007.
- [13] Li X, Shao X, Chang T, Albu L L. Does digital finance promote the green innovation of China's listed companies?[J]. *Energy Economics*, 2022, 114: 106254.
- [14] Liu J, Jiang Y, Gan S, He L, Zhang Q. Can digital finance promote corporate green innovation?[J]. *Environmental Science and Pollution Research*, 2022, 29(24): 35828-35840.
- [15] Merendino A, Dibb S, Meadows M, Quinn L, Wilson D, Simkin L, Canhoto A. Big data, big decisions: the impact of big data on board level decision-making[J]. *Journal of Business Research*, 2018, 93: 67-78.
- [16] Milani S, Neumann R. R&D, patents, and financing constraints of the top global innovative firms[J]. *Journal of Economic Behavior & Organization*, 2022, 196: 546-567.
- [17] Nedjah N, Azevedo V R, Mourelle L D M. Client profile prediction using convolutional neural networks for efficient recommendation systems in the context of smart factories[J]. *Enterprise Information Systems*, 2022, 16(10-11): 1653-1693.
- [18] Shannon C E. A mathematical theory of communication[J]. *Bell System Technical Journal*, 1948, 27(3): 379-423.
- [19] Sumbal M S, Tsui E, See-to E W K. Interrelationship between big data and knowledge management: an exploratory study in the oil and gas sector[J]. *Journal of Knowledge Management*, 2017, 21(1), 180-196.
- [20] Wen H, Zhong Q, Lee C C. Digitalization, competition strategy and corporate innovation: evidence from Chinese manufacturing listed companies[J]. *International Review of Financial Analysis*, 2022, 82: 102166.
- [21] Xu Z. Economic policy uncertainty, cost of capital, and corporate innovation[J]. *Journal of Banking & Finance*, 2020, 111: 105698.
- [22] 曹平, 陆松, 梁明柳. 大数据战略、知识管理能力与中国企业创新[J]. *产经评论*, 2021, 12(2): 102-119.
- [23] 陈德球, 孙颖, 王丹. 关系网络嵌入、联合创业投资与企业创新效率[J]. *经济研究*, 2021, 56(11): 67-83.
- [24] 冯根福, 刘虹, 冯照楨, 温军. 股票流动性会促进我国企业技术创新吗?[J]. *金融研究*, 2017, (3): 192-206.
- [25] 何玉润, 林慧婷, 王茂林. 产品市场竞争、高管激励与企业创新——基于中国上市公司的经验证据[J]. *财贸经济*, 2015, (2): 125-135.
- [26] 黄节根, 吉祥熙, 李元旭. 数字化水平对企业创新绩效的影响研究——来自沪深A股上市公司的经验证据[J]. *江西社会科学*, 2021, 41(5): 61-72+254-255.
- [27] 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.
- [28] 李宾, 龚爽, 曾雅婷. 数字普惠金融、融资约束与中小企业财务可持续[J]. *改革*, 2022, (5): 126-142.
- [29] 李健, 张金林, 董小凡. 数字经济如何影响企业创新能力: 内在机制与经验证据[J]. *经济管理*, 2022, 44(8): 5-22.
- [30] 李唐, 李青, 陈楚霞. 数据管理能力对企业生产率的影响效应——来自中国工业企业—劳动力匹配调查的新发现[J]. *中国工业经济*, 2020, (6): 174-192.
- [31] 李文秀, 唐荣. 融资约束、产业政策与本土企业出口行为——基于微观视角的理论与实证分析[J]. *中国软科学*, 2021, (7): 174-183.
- [32] 刘惠好, 焦文妞. 银行业竞争、融资约束与企业创新投入——基于实体经济金融化的视角[J]. *山西财经大学学报*, 2021, 43(10): 56-67.
- [33] 卢馨, 郑阳飞, 李建明. 融资约束对企业R&D投资的影响研究——来自中国高新技术上市公司的经验证据[J]. *会计研究*, 2013, (5): 51-58+96.
- [34] 宋德勇, 朱文博, 丁海. 企业数字化能否促进绿色技术创新?——基于重污染行业上市公司的考察[J]. *财经研究*, 2022, 48(4): 34-48.
- [35] 唐松, 伍旭川, 祝佳. 数字金融与企业技术创新——结构特征、机制识别与金融监管下的效应差异[J]. *管理世界*, 2020, 36(5): 52-66+9.
- [36] 万佳或, 周勤, 肖义. 数字金融、融资约束与企业创新[J]. *经济评论*, 2020, (1): 71-83.
- [37] 王道平, 刘琳琳. 数字金融、金融错配与企业全要素生产率——基于融资约束视角的分析[J]. *金融论坛*, 2021, 26(8): 28-38.
- [38] 王华, 韦欣彤, 曹青子, 廖凌韬. “营改增”与企业创新效率——来自准自然实验的证据[J]. *会计研究*, 2020, (10): 150-163.
- [39] 魏志华, 曾爱民, 李博. 金融生态环境与企业融资约束——基于中国上市公司的实证研究[J]. *会计研究*, 2014, (5): 73-80+95.
- [40] 温忠麟, 张雷, 侯杰泰, 刘红云. 中介效应检验程序及其应用[J]. *心理学报*, 2004, (5): 614-620.
- [41] 吴非, 胡慧芷, 林慧妍, 任晓怡. 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.
- [42] 吴翌琳, 黄实磊. 融资效率对企业双元创新投资的影响研究——兼论产品市场竞争的作用[J]. *会计研究*, 2021, (12): 121-135.
- [43] 肖文, 林高榜. 政府支持、研发管理与技术创新效率——基于中国工业行业的实证分析[J]. *管理世界*, 2014, (4): 71-80.
- [44] 肖土盛, 吴雨珊, 开文韬. 数字化的翅膀能否助力企业高质量发展——来自企业创新的经验证据[J]. *经济管理*, 2022, 44(5): 41-62.
- [45] 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.
- [46] 杨水利, 陈娜, 李雷. 数字化转型与企业创新效率——来自中国制造业上市公司的经验证据[J]. *运筹与管理*, 2022, 31(5): 169-176.
- [47] 翟淑萍, 张晓琳, 王鹤洁. 员工薪酬粘性改善企业创新效率吗?——基于高管-员工薪酬粘性差距的调节作用[J]. *研究与发展管理*, 2020, 32(4): 162-175.
- [48] 张嘉伟, 胡丹丹, 周磊. 数字经济能否缓解管理层短视行为?——来自真实盈余管理的经验证据[J]. *经济管理*, 2022, 44(1): 122-139.
- [49] 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.
- [50] 张传奇, 孙毅, 芦雪瑶. 现金流不确定性、管理者风险偏好和企业创新[J]. *中南财经政法大学学报*, 2019, (6): 71-81.
- [51] 张瑞, 唐旭丽, 王定峰, 潘建鹏. 基于知识关联的金融数据可视化分析[J]. *情报理论与实践*, 2018, 41(10): 131-136.
- [52] 周开国, 卢允之, 杨海生. 融资约束、创新能力与企业协同创新[J]. *经济研究*, 2017, 52(7): 94-108.
- [53] 周振江, 郑雨晴, 李剑培. 数字金融如何助力企业创新——基于融资约束和信息约束的视角[J]. *产经评论*, 2021, 12(4): 49-65.

(责任编辑: 吴林祥)