

DOI: 10.16538/j.cnki.fem.20230201.401

数据资产对企业创新投入的影响研究

李健¹, 董小凡², 张金林², 陶云清³

(1. 山东财经大学金融学院, 山东 济南 250014; 2. 中南财经政法大学金融学院, 湖北 武汉 430073;
3. 北京大学国家发展研究院/数字金融研究中心, 北京 100091)

摘要: 数据资产已成为企业和国家的重要战略资源和核心生产力。本文运用2011—2020年中国A股上市公司数据,通过量化企业数据资产,考察了数据资产对创新投入的影响及其机制。研究发现,数据资产能够显著提升企业的创新投入。在使用工具变量、双重差分等一系列稳健性检验后,研究结论仍然成立。异质性分析表明,数据资产对企业创新投入的提升作用在中小民营企业、融资依赖度较高的企业、技术密集型企业,以及金融科技水平低和创新不足的地区的企业更显著。此外,数据资产不仅可以通过缓解企业融资约束提高企业创新意愿,而且能够加强企业间研发合作拓展企业创新资源,进而有利于提升企业创新投入。本文为厘清数据价值对企业创新的作用与路径提供线索,也为我国数据资产的发展提供政策依据。

关键词: 数据资产; 创新投入; 融资约束; 研发合作

中图分类号: F270 **文献标识码:** A **文章编号:** 1001-4950(2023)12-0018-16

一、引言

伴随着大数据、区块链、人工智能等数字化技术的快速发展,数字经济成为驱动我国经济发展的关键动力。在数字经济时代,数据资源已成为关键生产要素。其原因在于,数据资源具有无限供给的禀赋,破除了传统生产要素对经济增长的供给制约,为经济可持续增长提供了可能。当前,数据作为新型生产要素,正向经济社会各领域加速渗透,并涌现出数据驱动型的产业、业态和商业模式(戚聿东和肖旭,2020)。对此,党的十九届四中全会首次提出将“数据”作为生产要素,并参与收入分配机制。2022年4月,中共中央、国务院印发《关于构建更加完善的要素市场化配置体制机制的意见》,进一步明确了数据作为第五种新型生产要素。数据作为新型的核心生产力,其资产化过程已成为必然趋势,但如何发挥数据资产价值已成为企业亟待解决的

收稿日期: 2022-09-21

基金项目: 国家社会科学基金重大项目(22&ZD115); 国家社会科学基金重大项目(21ZDA010); 国家自然科学基金地区项目(72261028)

作者简介: 李健(1992—),男,山东财经大学金融学院讲师(通讯作者,leeyks0215@163.com);

董小凡(1996—),女,中南财经政法大学金融学院博士研究生;

张金林(1964—),男,中南财经政法大学金融学院教授,博士生导师;

陶云清(1994—),男,北京大学国家发展研究院/数字金融研究中心博士后。

难题,也是学术界研究的重点问题(Leonelli,2019)。

针对数据资产的概念仍未形成共识,但数据资产化的基础是“数据”发挥生产力的过程(许宪春等,2022)。可以说,数据是所发生过的事件、时间、交易等信息的电子记录(Akoff,1989)。随着信息化技术的快速发展,数据已从小规模增长转向爆发增长、海量集聚的特点,该特点也是数据资产化过程的基础(Hannila等,2022)。根据《数据资产管理实践白皮书(5.0版)》中的定义,数据资产是指企业通过采集、存储与分析所形成的海量信息型的资源,且未来能够给企业带来经济价值(陈国青等,2020)。总的来说,数据资产在更多的情况下是企业生产经营过程中以电子形式所记录的信息,其本身不直接参与产品生产,但却蕴含着巨大的使用价值(Jones和Tonetti,2020),并能够在未来指导企业生产活动,帮助管理者进行决策。

创新是我国现代化建设的核心。“十四五”规划强调,以创新引领发展,加快建设科技强国。同时,党的二十大报告强调,创新是第一动力。客观来说,我国科技创新取得了巨大进步,但仍存在着自主创新能力相对不足、核心技术“卡脖子”的突出问题(李健等,2022)。突破创新发展瓶颈、顺利实施创新驱动战略,关键在于能够激发微观主体的创新意愿和创新资源(江轩宇,2016)。那么,数据资产作为新型生产要素能否提高企业创新投入?从已有相关研究来看,一方面,数据资产对企业的影响主要体现在成本管理、组织建设、风险管控等(Kiron,2013;Sivarajah等,2017);另一方面,激活企业数据资产能够驱动企业价值,提升企业核心竞争力(Huang和Von Mieghem,2014)。现有研究主要从理论上讨论了数据资产化的价值,但数据资产对企业创新行为的影响仍然较少,尤其是数据资产对企业创新投入影响路径有哪些?本文研究有助于回答数据资产在微观层面的价值,厘清数据资产对企业创新投入的作用。

本文可能的贡献在于:一是数据资产已成为重要的战略资源,但相关的研究并不多,且现有研究主要从宏观视角讨论了数据资产的价值、核算方式以及交易机制等方面的问题(彭刚等,2022;许宪春等,2022),而探讨数据资产对企业创新的研究较少。因此,本文从微观视角考察数据资产对企业创新投入的影响,从而丰富数据资产微观领域的研究。二是现有研究侧重于理论探讨数据资产对企业的影响(Nguyen和Paczos,2020;徐涛等,2022),缺少实证方面的研究,且数据资产测度方式仍处于探索阶段。据此,本文通过讨论企业数据资产化的过程量化数据资产,从而补充了数据资产的量化研究。三是国内外学者从理论模型上分析了数据给企业带来的变化(Tan等,2015;Hariri等,2019),少有研究分析数据资产对企业创新行为的影响路径。对此,本文将分析视角进行拓展,从创新意愿和创新资源的视角出发,揭示数据资产对企业创新投入的影响路径,有助于补充数据资产的理论研究与实证研究。

二、文献回顾与研究假说

随着大数据、人工智能、物联网在生产生活中的广泛运用,数据资产已成为数字经济时代重要的战略资产。在价值创造过程中,数据资产是能够重复应用于生产,且具有经济价值的数字资产(Rassier等,2019),有助于指导企业生产经营并创造社会效益(许宪春等,2022)。首先,作为信息与事实的载体,数据资产有助于帮助企业更好地进行生产经营与管理决策,提高信息与知识的转化效能,改进工艺流程并提升产品品质,提升劳动生产率与使用价值量,进而获得更多的经济利润(Veldkamp和Chung,2019;徐翔和赵墨非,2020),为需要大量投入的企业技术创新提供了可能。其次,数据资产相较于其他种类要素具有更强的正外部性,不仅能够提升企业的生产效率,还能借助信息流、技术流、资金流以及物资流形成数字产业集群,进而实现数字化协同联动与正反馈效应(Farboodi等,2019),有助于提升企业财务稳定性,为企业创新活动奠定物质基础。最后,在数字经济背景下,产业数字化与数字产业化已成为创新发展的必然要求。处

理运用数据资产能力较高的企业往往能够以较低的成本获取资金,借助数据资产与资本、劳动等传统生产要素的深度融合,有助于充分激发经济发展新动能,带动企业创新意愿与创新资源的提升(Begenau等,2018)。此外,随着数字化技术向各行业中的深度发展,数据资产会与劳动力、土地、资产等传统生产要素全面融合(Autor等,2020),促进了所有生产要素的协同和联动,从而释放各类生产要素在创新活动中的潜能,推动企业创新变革。据此,本文提出如下假设:

H1:数据资产有利于提升企业创新投入。

从创新意愿角度看,企业的数据资产能够提升外部资金的筹集能力,保证了企业充裕合理的创新活动流动性,有助于提升企业主观创新意愿与客观创新能力,进而以更高的财务稳定性提升创新投入。当前,我国企业普遍面临融资约束的问题,极大地削弱了企业的创新意愿(张璇等,2019)。首先,创新活动本身具有高投入、高风险、长周期等特征,对资金具有高度依赖性(Holmstrom,1989),而创新研发结果的不确定性较高,难以与传统商业银行的信贷安全性要求相匹配,导致企业技术创新容易受制于融资约束问题。然而,数据资产对传统资产构成有效的补充,催生了金融市场的新模式与新业态,有助于金融机构更充分地掌握企业生产经营状况与资信能力,缓解银企信息不对称导致的逆向选择与信贷风险,从而解决企业融资成本高等问题,缓解企业融资约束,提升信贷市场运行效率与资源供需的匹配性(Begenau等,2018; Carrière-Swallow和Haksar,2019)。其次,在产业数字化与数字产业化趋势下,拥有数据资产的企业往往被认为在生产经营管理中更具有竞争优势,更能获得投资者青睐(Kiron,2013)。最后,当企业拥有足够的现金流时,企业的创新活动会优先使用内部资金(Brown等,2009)。而数据资产作为企业战略性资产,能够提升企业运营效率,降低生产经营成本(Goodridge和Haskel,2015),以数字化供应链管理等多场景平台优化解决方案,赋能价值创造,为企业技术创新提供了稳定的现金流作为研发物质基础。据此,本文提出如下假设:

H2:企业数据资产会通过缓解融资约束提升企业创新投入。

从创新资源角度看,数据资产的虚拟性特征能够破除企业合作的时间与地域限制,带动了企业间合作形式的革新,从而有助于企业间提升创新合作质效,以更为高效的合作、共享共赢提升自主创新能力,推动企业创新资源与创新投入的提升。首先,数据资产具有共享性、交换性以及非消耗性等特征(Acquisti等,2016),可以实现企业间的资源共享与交换,进而有利于企业管理者产生新的产品创意(Tan等,2015),提升企业创新投入。其次,数字经济的发展在提升客户交互与业务运作效率的同时,也推进了市场需求的快速转变与创新技术的更新迭代,企业仅凭借自有资源开展技术创新活动将难以适应复杂多变的市场需求(李健等,2022)。然而,企业通过对共享的数据资产进行大数据分析可以更准确地刻画消费者画像,从而能够有针对性地改进现有产品,提升产品服务和质量,这也可以将数据资产看作是提高生产效率的资产(Farboodi和Veldkamp,2020)。最后,数据资产与数字技术深度融合有助于推进企业技术创新,研发类数据的采集、加工、储存、分析与运用贯穿于企业创新全流程中,以数据试错替代实物试错,有利于形成企业间的数据网络与研发合作,对于形成以数据驱动企业间创新合作的新路径具有重要意义(Akcigit和Liu,2016)。据此,本文提出如下假设:

H3:企业数据资产会通过增进企业间研发合作提升企业创新投入。

三、数据资产的量化与特征分析

(一)数据资产形成机制与量化

为量化企业数据资产,本文认为有必要探讨了解数据资产的形成机制。因此,本文借鉴Rassier等(2019)的研究,从“数据收集、数据存储、数据分析以及数据应用”四阶段的数据价值转变形态,探讨数据资产形成机制。

数据资产化过程中的第一阶段是数据收集,该阶段是企业将生产信息、交易信息等方面从手工记录转为数字化记录的过程,而这一阶段需要企业收集大量的数据才会凸显出数据的价值。第二阶段是数据存储,该阶段是企业将收集来的数据通过编码、脱敏等过程后形成标准化的数据,并存储于服务器中,以供企业进行后续分析。第三阶段是数据分析,该阶段是企业研发人员采用数据挖掘、机器学习等数字化分析技术对存储数据进行分析,其目的在于提取数据中的关键信息和知识点。第四阶段是数据应用。该阶段是数据资产实现价值的重要阶段,其目的在于通过第三阶段的数据信息挖掘辅助企业进行生产运营,帮助企业构建产品创意。

基于数据资产形成机制,本文依据数据资产形成阶段,结合企业财务报表附注中所披露的年末固定资产、无形资产明细中与数据资产化过程相关的项目占固定资产、无形资产的比例来测度企业数据资产。具体而言,当企业所披露的固定资产、无形资产支出明细中出现图1所示关键词时,则初步判定为企业数据资产化过程,再将该企业同一年度与数据资产化过程相关的固定资产、无形资产进行求和,并计算其占固定资产和无形资产总额的比例,以此作为企业数据资产的代理变量(*Digas*)。企业数据资产的形成应涵盖数据收集、数据存储、数据分析以及数据应用四个阶段。对此,本文进行人工复核,只有当企业的数据资产化过程经历完整的四个阶段才会判定为企业数据资产。此外,为保证本文数据资产变量的合理性,本文在后续的稳健性

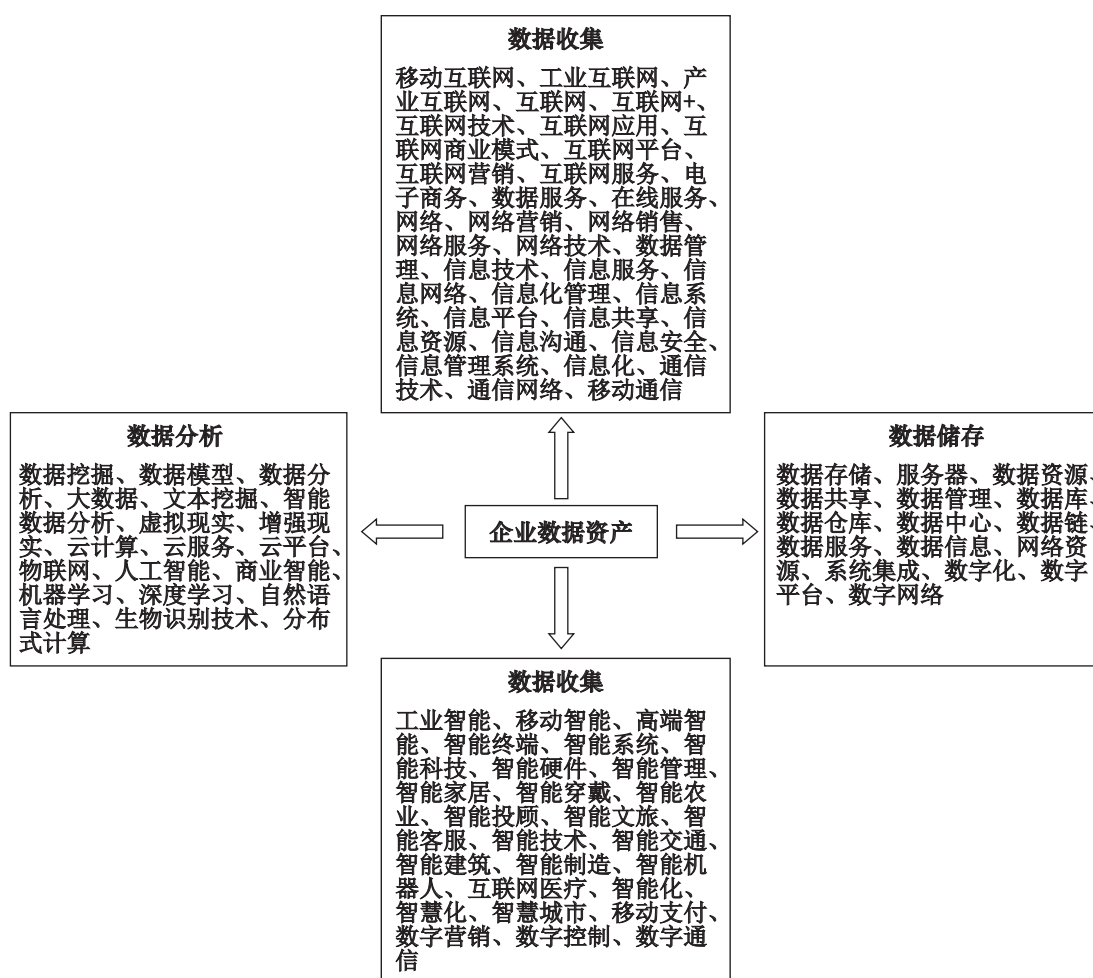


图1 企业数据资产指标特征词谱

检验中,采用文本挖掘法对图1所示关键词词频进行挖掘,并使用熵值法赋权,最终可得到企业数字资产指数。

(二)企业数据资产的典型事实

随着数字化技术的快速发展,数字经济已成为经济发展的新动能。根据国务院发布的《携手构建网络空间命运共同体》白皮书指出,截至2021年,中国数字经济规模占国内生产总值比重高达39.8%。得益于数字经济的快速发展,不仅利于了各类资源要素的加速流动,而且实现了各类微观主体的跨界融合,这为数据资产的发展奠定了基础。与之相伴的是,数字经济的快速发展也成为推动实体经济恢复发展和创新的重要手段。由此,本文绘制了企业数据资产和创新投入的2011—2020年的变化趋势,如图2所示。研究发现,数据资产一直呈现上升趋势,尤其是党的十九大后更呈现出明显的增长态势(图1所示)。从企业研发活力视角看,企业研发投入呈现缓慢的增长,但在2017年后企业创新投入增长较为明显(图2b所示)。

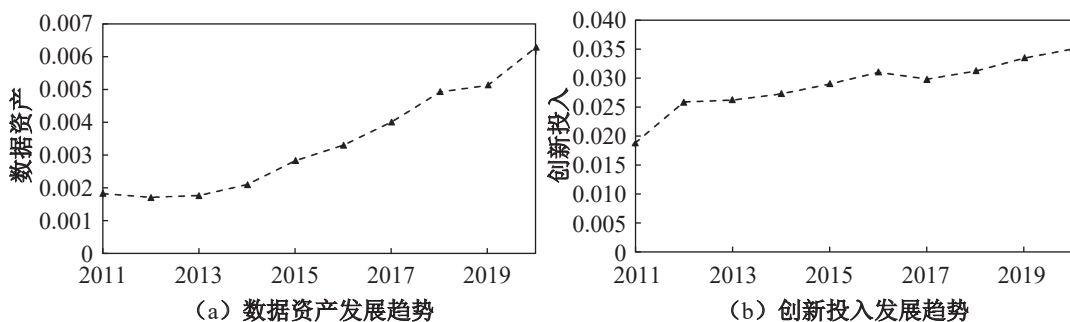


图2 企业数据资产与创新投入情况

数据资产的核心特性是虚拟性(Jones和Tonetti, 2020),数据资产必须要结合其他生产要素才能发挥价值。目前来看,企业对于数据资产的运用更多的是查询与分析等,并没有完全发挥数据资产的价值,且数据资产多存在于ICT产业中(徐涛等, 2022)。表1展示了按证监会行业分类的拥有数据资产企业的行业均值。研究发现,数据资产较多地分布于信息技术业、传播与文化产业、综合类以及批发和零售贸易等行业,而采掘业、农、林、牧、渔业以及电力、煤气及水的生产和供应业等劳动密集型行业中拥有数据资产较少。

表1 各行业数据资产均值

行业	均值	行业	均值
农、林、牧、渔业	0.0008	建筑业	0.0013
制造业	0.0021	交通运输、仓储业	0.0015
信息技术业	0.0191	采掘业	0.0005
综合类	0.0036	传播与文化产业	0.0091
批发和零售贸易	0.0035	社会服务业	0.0023
电力、煤气及水的生产和供应业	0.0010		

拥有数据资产的企业与未拥有企业可能存在明显区别,表2展示了企业是否拥有数据资产的相关变量组间差异检验结果。研究发现,拥有数据资产企业的规模显著高于未拥有数据资产企业。从盈利能力上看,拥有数据资产企业的净资产收益率明显高于未拥有数据资产企业,这表明拥有数据资产企业在市场竞争中更具有优势。从创新活动上看,拥有数据资产企业的专利数量显著高于未拥有数据资产企业,且未来成长能力更强。因此,拥有数据资产企业属于竞争优势较强、创新能力较强、成长能力较好的企业。

表2 数据资产企业的经营特征

类别	企业规模	盈利能力	成长能力	研发能力
拥有数据资产的企业	21.9296	0.0357	2.3983	1.8377
未拥有数据资产的企业	22.5346	0.0248	2.1976	0.9179
差值	0.6050***	0.0109***	0.2006***	0.9198***

四、研究设计与变量选取

(一)数据来源

本文选取2011—2020年中国A股上市公司作为研究对象。数据来源主要包括：一是上市公司数据资产指数通过收集样本期间的年报，采用手工量化的方式得到。二是上市公司的其余变量均通过Wind和CSMAR数据库进行采集。同时，本文按照已有研究文献剔除了金融行业、ST和*ST企业以及数据缺失严重的企业。通过上述处理，本文共获得8100个观测值。

(二)变量定义

1.被解释变量。本文借鉴黎文靖等(2021)的研究思路，分别采用研发支出占销售收入的比重(Inn_1)与研发支出占总资产的比重(Inn_2)作为被解释变量。同时，本文发现样本期内大部分企业未披露研发经费的支出情况。为避免被解释变量缺失引起的估计偏误，本文参考刘诗源等(2020)的研究，将企业未披露研发经费支出的情况按照0值处理。

2.核心解释变量。本文通过查找企业年报中披露的固定资产、无形资产支出明细量化数据资产的拥有程度，并在后续的稳健性检验中采用文本挖掘法挖掘特征词词谱出现的频率，采用熵值法进行赋权，具体量化方法见第三部分。

3.控制变量。参考已有研究文献(唐松等,2020;宋敏等,2021)，本文分别选取企业规模($Size$)、企业年龄(Age)、资产负债率(Lev)、总资产收益率(Roa)、股权集中度($Share$)、企业成长性($Tobinq$)、两职合一($Merge$)和审计意见($Optin$)。具体变量定义见表3所示。

表3 变量定义

变量名	变量符号	衡量方法
创新投入	Inn_1	研发支出/销售收入
	Inn_2	研发支出/资产总额
企业数据资产程度	$Digas$	见数据资产指标量化
企业规模	$Size$	总资产的对数值
企业年龄	Age	观测年份减去成立年份后取对数值
资产负债率	Lev	总负债/总资产
总资产收益率	Roa	净利润/总资产
股权集中度	$Share$	前五位大股东持股数/总股数
成长性	$Tobinq$	(股票市值+债务总额)/总资产
两职合一	$Merge$	董事长与总经理为同一人取1，否则为0
审计意见	$Optin$	标准无保留意见取0，否则为1

(三)研究设计

为验证本文的研究假设，本文的基准模型如下：

$$Inn_{it} = \alpha_0 + \alpha_1 Digas_{it} + \alpha Controls_{it} + \mu_t + \theta_i + \varepsilon_{it} \quad (1)$$

其中， Inn 代表企业创新投入，包括 Inn_1 和 Inn_2 ， $Digas$ 代表企业数据资产， $Controls$ 代表一系列控制变量， ε 代表随机扰动项。为缓解遗漏变量带来的影响，本文还加入了企业固定效应 θ_i 和年份固定效应 μ_t ，且回归结果的标准误均经过企业层面聚类调整。

(四)描述性统计

表4展示了主要变量的描述性统计结果。结果显示,两种方式计算的企业创新投入均值分别为0.0287和0.0155,标准差分别为0.0405和0.0179,这说明企业间的创新投入存在显著差异。企业数据资产的均值和标准差分别为0.0032和0.0075,这表明不同企业所拥有数据资产程度存在较大差异,也为本文的回归分析奠定了良好基础。

表4 变量描述性统计

变量	样本数	均值	标准差	最小值	最大值
<i>Inn₁</i>	8100	0.0287	0.0405	0	1.2591
<i>Inn₂</i>	8100	0.0155	0.0179	0	0.1622
<i>Digas</i>	8100	0.0032	0.0075	0	0.04775
<i>Size</i>	8100	22.4891	1.2996	19.8713	26.2241
<i>Age</i>	8100	2.9388	0.2776	1.3862	3.7376
<i>Lev</i>	8100	0.4658	0.2021	0.0626	0.9076
<i>Roa</i>	8100	0.0349	0.0578	-0.2029	0.2097
<i>Share</i>	8100	0.5439	0.1537	0	0.9614
<i>Tobinq</i>	8100	2.2127	1.5684	0.8209	9.6767
<i>Merge</i>	8100	0.0266	0.1611	0	1
<i>Optin</i>	8100	0.1928	0.3945	0	1

五、实证分析

(一)基准回归结果

表5展示了基准模型的回归结果。其中,被解释变量分别为*Inn₁*和*Inn₂*,每列回归均控制了企业和年份固定效应,且标准误均经过企业层面聚类调整。研究发现,第(1)与(2)列加入企业财务特征控制变量的实证结果显示,*Digas*的估计系数在1%水平上为正(分别为0.0358与0.0842);第(3)和(4)列加入全部的控制变量的实证结果显示,*Digas*的估计系数依然显著为正(分别为0.0354与0.0259)。实验结果表明,数据资产能够显著提高企业创新投入,假设H1得到验证。此外,控制变量的结果与已有文献基本一致。

表5 企业数据资产与创新投入

变量	(1) <i>Inn₁</i>	(2) <i>Inn₂</i>	(3) <i>Inn₁</i>	(4) <i>Inn₂</i>
<i>Digas</i>	0.0358*** (5.01)	0.0842*** (5.43)	0.0354*** (4.94)	0.0259*** (6.88)
<i>Size</i>	0.0009** (2.42)	-0.0017*** (-8.57)	0.0009** (2.19)	-0.0017*** (-7.75)
<i>Age</i>	-0.0081* (-1.81)	-0.0014 (-0.62)	-0.0091** (-1.99)	-0.0007 (-0.29)
<i>Lev</i>	-0.0111*** (-6.52)	-0.0001 (-0.11)	-0.0112*** (-6.54)	-0.0002 (-0.30)
<i>Roa</i>	-0.0357*** (-9.76)	0.0079*** (4.11)	-0.0359*** (-9.49)	0.006*** (3.05)
<i>Share</i>			-0.0036* (-1.65)	0.0023** (2.02)
<i>Tobinq</i>			-0.0001 (-0.74)	0.0001** (2.03)
<i>Merge</i>			-0.0026** (-2.37)	-0.0013** (-2.36)
<i>Optin</i>			0.0005 (0.91)	-0.0001 (-0.33)
常数项	0.0258* (1.75)	0.0529*** (6.83)	0.0306** (2.04)	0.0495*** (6.26)
时间固定效应	YES	YES	YES	YES
企业固定效应	YES	YES	YES	YES
<i>N</i>	8100	8100	8100	8100
<i>R²</i>	0.0925	0.077	0.0977	0.0930

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

本文进一步考察了数据资产指标分解回归以及滞后回归,见表6所示。表6中的PanelA展

示了数据资产指标分解回归,其中,*Digas_dco*、*Digas_dst*、*Digas_dan*以及*Digas_dap*分别代表数据资产形成机制的四个阶段,即数据收集、数据存储、数据分析以及数据应用。从中可看出,企业数据资产形成机制中的每一阶段的估计系数均在1%显著水平上显著为正(*Inn₁*与*Inn₂*的结果均一致),这表明数据资产化过程会对企业创新投入产生影响。表6中的PanelB展示了数据资产指标滞后回归。从中可看出,随着时间的推移,数据资产对创新投入的影响效力在逐渐转弱(*Inn₁*与*Inn₂*的结果均一致)。具体而言,在10%显著水平上,滞后1至3期时数据资产对创新投入的影响显著为正,但滞后4期后没有统计上的显著性影响。可见,数据资产展现出了时间序列上的动态衰减特征,这也刚好体现了数据资产的时效性特点(许宪春等,2022)。

表6 特征分析:企业数据资产与创新投入

PanelA	(1) <i>Inn₁</i>	(2) <i>Inn₂</i>	(3) <i>Inn₁</i>	(4) <i>Inn₂</i>	(5) <i>Inn₁</i>	(6) <i>Inn₂</i>	(7) <i>Inn₁</i>	(8) <i>Inn₂</i>
<i>Digas_dco</i>	0.0442*** (5.59)	0.0359*** (8.57)						
<i>Digas_dst</i>			0.0115*** (4.71)	0.0053*** (4.12)				
<i>Digas_dan</i>					0.0183*** (4.88)	0.0097*** (4.88)		
<i>Digas_dap</i>							0.0111*** (2.95)	0.0051** (2.56)
控制变量	YES	YES	YES	YES	YES	YES	YES	YES
时间固定效应	YES	YES	YES	YES	YES	YES	YES	YES
企业固定效应	YES	YES	YES	YES	YES	YES	YES	YES
<i>N</i>	8100	8100	8100	8100	8100	8100	8100	8100
<i>R</i> ²	0.0778	0.0900	0.0796	0.0575	0.0619	0.0540	0.0659	0.0533
PanelB	(1) <i>Inn₁</i>	(2) <i>Inn₂</i>	(3) <i>Inn₁</i>	(4) <i>Inn₂</i>	(5) <i>Inn₁</i>	(6) <i>Inn₂</i>	(7) <i>Inn₁</i>	(8) <i>Inn₂</i>
L1. <i>Digas</i>	0.0292*** (3.88)	0.0235*** (6.06)						
L2. <i>Digas</i>			0.0219*** (2.64)	0.0238*** (5.67)				
L3. <i>Digas</i>					0.0103* (2.04)			
L4. <i>Digas</i>						0.0192*** (3.89)	0.0121 (1.01)	0.0183*** (3.08)
控制变量	YES	YES	YES	YES	YES	YES	YES	YES
时间固定效应	YES	YES	YES	YES	YES	YES	YES	YES
企业固定效应	YES	YES	YES	YES	YES	YES	YES	YES
<i>N</i>	7290	7290	6480	6480	5670	5670	4860	4860
<i>R</i> ²	0.0828	0.0682	0.0594	0.0525	0.0402	0.0250	0.0164	0.0059

(二)内生性问题处理

数据资产对企业创新投入的影响可能受到内生性影响,具体而言:

一是逆向因果的影响。尽管本文从动态视角考察了数据资产对创新投入的影响,能够在一定程度上缓解逆向因果带来的内生性问题,但可能会出现创新投入相对较高的企业更倾向于拥有较大程度的数据资产,这也可以说是,企业创新投入反过来对数据资产产生影响。因此,为避免逆向因果带来的内生性问题,本文将选用工具变量法弱化潜在影响。

参考Chong等(2013)和唐松等(2020)的研究思路,本文选用企业所在行业的其他企业数据资产的平均值作为工具变量。值得注意的是,选用与该企业在同一行业内的其他企业数据资

产平均值作为工具变量,从第一阶段的检验中可发现,工具变量的Wald F统计量为227.83,这说明本文选取的工具变量不存在弱工具变量问题,且在1%显著水平上工具变量的估计系数显著为正。第二阶段的检验结果显示,在1%显著水平上核心解释变量*Digas*仍然显著为正,这表明本文核心结论仍然成立,即数据资产会提升企业创新投入(见表7)。

表7 工具变量的检验结果

	(1) <i>Inn</i> ₁		(2) <i>Inn</i> ₂	
	第一阶段	第二阶段	第一阶段	第二阶段
<i>IV</i>	0.969***(25.18)		0.969***(25.18)	
<i>Digas</i>		0.462***(19.07)		0.2126***(14.35)
常数项	-0.038***(-3.82)	0.0701***(8.49)	-0.038***(-3.82)	0.0503***(11.07)
控制变量	YES	YES	YES	YES
时间固定效应	YES	YES	YES	YES
企业固定效应	YES	YES	YES	YES
<i>F-test</i> (第一阶段)	227.83		227.83	
<i>LM</i> 统计量(<i>p</i> 值)	0.0000		0.0000	
<i>N</i>	8100	8100	8100	8100
<i>R</i> ²		0.0202		0.0125

同时,本文也采用双重差分法来解决内生性问题。2016年工信部发布了《大数据产业发展规划(2016—2020年)》。该规划鼓励各行业强化大数据产品应用和研发,建设大数据标准体系。该政策对于各行业而言是促进其数据应用和创新的一个外生冲击。为此,本文将《大数据产业发展规划(2016—2020年)》作为典型的外生冲击,并以各企业异质性构建实验组和对照组。基于此,本文构建如下模型

$$Inn_{it} = \lambda_0 + \lambda_1 Treat_{it} \times Post_{it} + \lambda Controls_{it} + \mu_t + \theta_i + \varepsilon_{it} \quad (2)$$

其中,*Treat*代表对照组和实验组的区分标记,以2016年各企业数据资产发展水平的中位数进行分组,企业所在行业2016年末的数据资产小于当年所有行业的中位数,则记为实验组,赋值为1,否则记为对照组,赋值为0。*Post*代表年份虚变量,即2016年及以后为1,否则为0。同时,为保证双重差分模型结果的稳健性,本文还加入了平行趋势检验。

表8展示了双重差分模型的估计结果。从第(1)和(2)列可看出,*Treat* × *Post*的估计系数在1%水平上显著为正,这表明在该政策出台后,企业数据资产呈现出增长的态势,且实验组中的企业数据资产增长更为明显。从第(3)和(4)列可看出,*Treat*与事前年份的虚变量均不显著,而*Treat*与当年及事后年份的交互项在1%的水平上显著为正,这表明在该政策出台前,企业在创新投入上并没有差异,模型的平行趋势假定通过检验。

尽管本文已尽可能地控制了样本偏差问题,但仍存在选择偏差问题。对此,本文考虑运用Heckman两阶段模型缓解样本选择偏差的问题。其中,Heckman模型中的第一阶段,本文将按照该企业数据资产是否超过同行业年度中位数设置哑变量,即该企业超过同行业年度中位数则赋值为1,否则为0,并以第一阶段的Probit模型估计结果计算逆米尔斯比率(*IMR*)。表9中第(1)和(2)列的结果显示,在1%显著性水平上,核心解释变量*Digas*估计系数仍然显著为正,这表明本文的研究结论是稳健的。

除企业自身特征影响创新投入外,地区的生产总值、产业结构、政府干预程度等方面会影响企业的创新投入。同时,企业所在行业随时间的变化趋势也会影响企业创新活动。因此,本文在模型(1)的控制变量中加入了地区经济发展水平(各省份生产总值的对数值)、地区产业结构(各省份第二产业产值/生产总值)、政府干预度(各省份财政收入/财政支出)以及控制了“行业

×时间”的高阶联合固定效应,结果见表10中第(3)和(4)列所示。结果显示,核心解释变量*Digas*的估计系数在1%显著水平上依然为正,这表明本文的实验结果仍然成立。

表8 大数据产业发展规划与企业创新投入

	双重差分估计		平行趋势检验	
	(1) <i>Inn</i> ₁	(2) <i>Inn</i> ₂	(3) <i>Inn</i> ₁	(4) <i>Inn</i> ₂
<i>Treat</i> × <i>Post</i>	0.0032*** (4.14)	0.0016*** (3.23)		
<i>Treat</i> × <i>year</i> 2014			0.0009 (0.57)	0.0007 (1.03)
<i>Treat</i> × <i>year</i> 2015			0.0017 (0.79)	0.0005 (0.44)
<i>Treat</i> × <i>year</i> 2016			0.0033* (1.98)	0.0009 (1.03)
<i>Treat</i> × <i>year</i> 2017			0.0041** (2.41)	0.0013*** (2.95)
<i>Treat</i> × <i>year</i> 2018			0.0053*** (3.22)	0.0021*** (3.03)
控制变量	YES	YES	YES	YES
时间固定效应	YES	YES	YES	YES
企业固定效应	YES	YES	YES	YES
<i>N</i>	8100	8100	8100	8100
<i>R</i> ²	0.0338	0.0243	0.0426	0.0379

表9 内生性检验结果

	Heckman两阶段模型		控制遗漏变量	
	(1) <i>Inn</i> ₁	(2) <i>Inn</i> ₂	(3) <i>Inn</i> ₁	(4) <i>Inn</i> ₂
<i>Digas</i>	0.0351*** (4.89)	0.0258*** (6.85)	0.1783*** (21.65)	0.0781*** (17.35)
<i>IMR</i>	-0.6296** (-2.54)	-0.1713** (-2.21)		
控制变量	YES	YES	YES	YES
时间固定效应	YES	YES	NO	NO
企业固定效应	YES	YES	NO	NO
行业×年份固定效应	NO	NO	YES	YES
<i>N</i>	8100	8100	8100	8100
<i>R</i> ²	0.0831	0.0782	0.3781	0.3284

表10 稳健性检验结果

	更换量化方式		剔除部分样本		更换估计模型	
	(1) <i>Inn</i> ₁	(2) <i>Inn</i> ₂	(3) <i>Inn</i> ₁	(4) <i>Inn</i> ₂	(5) <i>Inn</i> ₁	(6) <i>Inn</i> ₂
<i>Digas</i>	0.1432*** (4.86)	0.0837*** (5.40)	0.0197** (2.59)	0.0209** (5.12)	0.0535*** (6.74)	0.0351*** (8.19)
常数项	0.0295** (1.97)	0.0468*** (5.94)	0.0397** (2.50)	0.0644*** (7.55)	0.0715*** (4.99)	0.0741*** (9.57)
控制变量	YES	YES	YES	YES	YES	YES
时间固定效应	YES	YES	YES	YES	YES	YES
企业固定效应	YES	YES	YES	YES	YES	YES
<i>N</i>	8100	8100	6860	6860	8100	8100
<i>R</i> ²	0.0868	0.0766	0.0774	0.0776		

(三)稳健性检验

除进行内生性处理外,本文进行了一系列的稳健性检验。具体而言:

一是替换核心解释变量的测度方式。本文进一步采用文本挖掘法对图1所示关键词词频进行挖掘,并使用熵值法赋权,结果见表10中第(1)和(2)列。可以发现,更换数据资产的测度方式后,本文的核心解释变量*Digas*与被解释变量均显著为正,本文的研究结论没有发生改变。

二是剔除特定样本的影响。我国的直辖市存在较大的经济特殊性,企业数据资产拥有程度与创新投入可能存在不同。对此,本文借鉴李健等(2022)的研究,将样本中企业所属地为直辖市的进行剔除后重新回归检验,结果见表10中第(3)和(4)列。结果显示,*Digas*的系数仍然显著为正,本文的核心结论“数据资产能够提升企业创新投入”没有发生变化。

三是改变参数估计方法。根据被解释变量企业创新投入的特点(零值堆积),本文参考唐松等(2020)的研究,采用Tobit模型进行重新估计,结果见表10中第(5)和(6)列。从结果来看,核心解释变量*Digas*的显著性和符号均没有变化,这说明数据资产仍然对创新投入具有明显的提升作用。

(四)异质性检验

1.基于企业特征的异质性检验

企业数据资产对不同性质、规模的企业,其对创新投入的提升作用可能存在差异(于瑶和祁怀锦,2022)。相较于大型国有企业而言,中小民营企业的融资约束问题更严重(李旭超等,2017),这可能会严重影响企业的创新投入。为深入探究企业特征的异质性,本文按照企业所有制(*poe*)以及企业规模(*size*)进行赋值后与数据资产做交互项。

本文按照企业所有制将样本分为国有企业和非国有企业,将国有企业赋值为1,非国有企业赋值为0。同样地,按照《统计上大中小微型企业划分办法(2017)》,将研究样本按企业当年的营业收入、员工人数以及资产规模划分为大型企业和中小型企业,大型企业赋值为1,否则为0。结果见表11所示。

表 11 基于企业特征的异质性检验

	企业性质		企业规模	
	(1) <i>Inn</i> ₁	(2) <i>Inn</i> ₂	(3) <i>Inn</i> ₁	(4) <i>Inn</i> ₂
<i>Digas</i>	0.0267*** (2.69)	0.0271*** (6.40)	0.0647*** (5.72)	0.0538*** (9.07)
<i>Digas</i> × <i>poe</i>	-0.0026*** (-3.39)	-0.0021*** (-3.63)		
<i>Digas</i> × <i>size</i>			-0.0421*** (-3.34)	-0.0401*** (-6.07)
控制变量	YES	YES	YES	YES
时间固定效应	YES	YES	YES	YES
企业固定效应	YES	YES	YES	YES
<i>N</i>	8100	8100	8100	8100
<i>R</i> ²	0.0979	0.0931	0.0851	0.0850

表11的第(1)与(2)列的回归结果显示,企业数据资产与企业所有制的交互项系数(*Digas* × *poe*)₁在1%水平上显著为负,且创新投入*Inn*₁与*Inn*₂的结果均一致。从第(3)与(4)列的回归结果看,企业数据资产与企业规模的交互项系数(*Digas* × *size*)₁在1%水平上显著为负,这表明与大型国有企业相比,数据资产对中小民营企业创新投入的提升效应更大。可见,企业资产可以帮助金融机构有效评估企业风险,降低了银企间的信息不对称,进而有利于拓宽企业融资渠道。

2.基于产业性质的异质性检验

不同行业的融资需求不同,通常来说融资需求较高行业的融资约束问题会更严重(Rajan和Zingales,1998)。数据资产如果能缓解融资需求较高行业的融资约束问题,可能会促进这类行业的创新活动。本文参考Hsu等(2014)的研究,通过计算每家企业的融资需求^①,确定该企业所属行业的融资依赖度,并按照各年份各行业中位数进行分组,将行业融资依赖度高

①企业融资需求:(企业资本支出+研发支出-经营活动产生的现金流)/(企业资本支出+研发支出)。

的赋值为1,否则为0。从表12第(1)与(2)列可看出,数据资产与行业融资依赖度的交互项系数($Digas \times def$)在5%水平上显著为正,这表明相较于融资依赖度低的行业,数据资产对高融资依赖行业的创新驱动作用更显著。也就是说,在融资依赖度较高的行业中,数据资产通过缓解融资约束提升企业创新投入的边际贡献更大。

表 12 基于行业性质的异质性检验

	外部融资依赖度		行业性质	
	(1) Inn_1	(2) Inn_2	(3) Inn_1	(4) Inn_2
$Digas$	0.0452***(4.66)	0.0345***(6.77)	0.0468***(5.95)	0.0321***(7.77)
$Digas \times def$	0.0145(1.50)	0.0127**(2.49)		
$Digas \times ind$			0.0476***(3.49)	0.0259*(1.92)
控制变量	YES	YES	YES	YES
时间固定效应	YES	YES	YES	YES
企业固定效应	YES	YES	YES	YES
N	8 100	8 100	8 100	8 100
R^2	0.0978	0.0927	0.1239	0.1154

不同行业对创新的需求不同。相较于传统行业,高新技术企业更可能通过研发合作的方式来开展创新活动(Ang等,2014)。据此,本文借鉴鲁桐和党印(2014)的研究,将样本分为技术密集型行业和非技术密集型行业,其中电子行业、医药行业、信息技术以及机械、设备、仪表行业为技术密集型行业,其他为非技术密集型行业,并将技术密集型行业赋值为1,否则为0。从表12第(3)与(4)列可看出,数据资产与行业性质的交互项系数($Digas \times ind$)在10%水平上显著为正,这表明数据资产对技术密集型行业创新投入的提升作用更显著。出现这一现象的原因可能在于:技术密集型行业通常属于高新技术行业,对创新的需求更大,数据资产通过促进企业间的创新交流和共享,可以发挥不同企业的技术和人才优势,从而实现优势互补,形成专业化的分工,共同开展创新活动。

3.基于宏观环境的异质性检验

除企业特征和行业性质外,已有研究发现,企业所处的宏观环境对企业创新投入同样具有显著影响(鞠晓生等,2013)。例如,地区的金融科技发展程度越高,越能缓解中小企业的融资约束问题,从而助力中小企业进行研发创新(宋敏等,2021)。因此,本文选取地区数字普惠金融指标作为金融科技发展水平的代理变量,并参考冯俊诚(2022)的研究,以25%作为分界线,将地区金融科技发展水平高的赋值为1,否则为0。从表13第(1)与(2)列可看出,数据资产与地区金融科技的交互项系数($Digas \times fin$)在1%水平上显著为负,这表明数据资产提升企业创新的作用在金融科技发展程度较低的地区中更明显。也就是说,数据资产对融资约束的缓解,支持了金融科技不发达地区的企业创新活动。

地区的创新环境会影响企业进行研发合作创新。本文参考邱子迅和周亚虹(2021)的研究,选取各省份专利数量作为地区创新环境的代理变量,并按照25%分位数为界限划分样本,将地区创新程度高的赋值为1,否则为0。从表13第(3)与(4)列可看出,数据资产与地区创新环境的交互项系数($Digas \times rei$)在10%水平上显著为负,这说明数据资产对企业创新投入的提升作用在创新基础薄弱的地区更显著。其原因在于,一方面,对比前几次的工业革命,以数字化技术为代表的第四次工业革命对地区的创新程度不具有依赖性。创新基础薄弱的地区只要具备一定的互联网基础同样可以追赶发达地区。数据资产是第四次工业革命的核心要素,地区创新动能不足的企业可以借助数据资产,实现突破性的发展。另一方面,企业数据资产具有较强的“非

竞争性”特征(徐翔等,2021),能够满足不同地区的企业共同使用,从而打破地区间创新活动的区域壁垒,促进企业间展开创新交流活动,这在一定程度上帮助了弱势地区企业的创新活动,这也体现了数据资产具有一定的“帮扶”作用。

表 13 基于宏观环境的异质性检验

	金融科技发展程度		创新地区	
	(1) <i>Inn</i> ₁	(2) <i>Inn</i> ₂	(3) <i>Inn</i> ₁	(4) <i>Inn</i> ₂
<i>Digas</i>	0.0539*** (5.50)	0.0383*** (7.44)	0.0403*** (3.96)	0.0328*** (6.14)
<i>Digas</i> × <i>fin</i>	-0.0283*** (-2.76)	-0.0191*** (-3.53)		
<i>Digas</i> × <i>rei</i>			-0.0075** (-2.68)	-0.0105* (-1.82)
控制变量	YES	YES	YES	YES
时间固定效应	YES	YES	YES	YES
企业固定效应	YES	YES	YES	YES
<i>N</i>	8 100	8 100	8 100	8 100
<i>R</i> ²	0.0989	0.0943	0.0971	0.0926

六、拓展性分析:机制检验

本文前述研究显示,数据资产对企业创新投入具有显著的促进作用。特别是,数据价值形态转化中的每一阶段都有助于促进企业的创新投入。上述研究刻画了数据资产对企业创新投入的影响。本文前述中从创新意愿和创新资源两个视角出发,理论探究了数据资产对企业创新投入的作用机制,但需要进一步进行实证检验。因此,本文在模型(1)的基础上,构建如下模型进行验证:

$$Mediator_{it} = \gamma_0 + \gamma_1 Digas_{it} + \gamma Controls_{it} + \mu_t + \theta_i + \varepsilon_{it} \quad (3)$$

$$Inn_{it} = \phi_0 + \phi_1 Mediator_{it} + \phi_2 Digas_{it} + \phi Controls_{it} + \mu_t + \theta_i + \varepsilon_{it} \quad (4)$$

其中,*Mediator*代表中介变量。正如前文所述,数据资产降低了企业间的信息不对称,缓解了企业融资约束,从而增进企业创新意愿。因此,参考Hadlock和Pierce(2010)的研究,选取SA指数作为企业融资约束的代理变量^①。数据资产在缓解融资约束问题的同时,也能够一定程度上促进企业间展开研发合作,拓展企业创新资源,进而对企业创新活动形成助益。据此,本文参考张红娟等(2022)的做法,采用企业联合申请专利的自然对数(*CRT*)作为研发合作的代理变量。

1. 基于企业创新意愿的机制检验

表14展示了企业创新意愿视角的检验结果,即数据资产、融资约束与企业创新投入的回归结果。从表14第(1)列可看出,数据资产(*Digas*)对企业融资约束(*SA*)的回归系数在1%显著水平下为负,这表明数据资产缓解了企业融资约束问题。从第(2)与(3)列可看出,再加入融资约束(*SA*)的变量后,融资约束与企业创新投入呈现显著的负相关(*SA*的估计系数在1%显著水平上为负),这说明融资约束抑制了企业创新投入。同时,结合表5中核心解释变量*Digas*对企业创新投入的边际效应有所下降,这表明融资约束是数据资产对企业创新投入影响中的中介因子。数据资产可以有效缓解信息不对称问题,使得外部融资机构能够甄别“真创新”企业,从而提高企业融资的可获得性,避免企业因资金问题无法开展研发活动。可见,数据资产通过缓解企业融资约束提高企业创新意愿,假设H2得以验证。

2. 基于企业创新资源的机制检验

为验证假设H3,本文仍使用中介效应模型检验数据资产能否通过加强企业间研发合作,

①SA指数的计算方式: $SA = 0.043Size^2 - 0.04Age - 0.737Size$

促进企业的创新投入。表15第(1)列结果显示, *Digas* 的估计系数在1%水平上显著为正, 这表明企业数据资产能显著加强企业间研发合作。第(2)与(3)列中 *CRT* 的估计系数在1%显著水平上为正, 表明企业间研发合作能够提升企业创新投入; *Digas* 的估计系数显著为正, 且系数有所下降, 这说明数据资产通过加强企业间研发合作提升了企业创新投入。一方面, 数据资产通过移动终端、互联网服务等数字化平台, 可以增进企业间研发交流合作, 促使企业通过引进、吸收、消化新知识, 推动企业的创新投入。另一方面, 随着不同微观主体之间对创新经验的交流, 知识的外溢效应会推动企业协同创新。由此, 数据资产通过加强企业间研发合作, 扩展了企业创新资源, 进而为企业创新投入奠定资源基础, 本文假设H3得到经验支撑。

表 14 创新意愿: 数据资产、融资约束与创新投入

	(1) <i>SA</i>	(2) <i>Inn₁</i>	(3) <i>Inn₂</i>
<i>Digas</i>	-0.2092***(-6.35)	0.0334*** (4.66)	0.0256*** (6.80)
<i>SA</i>		-0.0092***(-3.62)	-0.0012***(-3.09)
常数项	-3.6693***(-53.05)	-0.0031(-0.18)	0.0451*** (4.84)
控制变量	YES	YES	YES
时间固定效应	YES	YES	YES
企业固定效应	YES	YES	YES
<i>N</i>	8 100	8 100	8 100
<i>R</i> ²	0.0946	0.1059	0.0926

表 15 创新资源: 数据资产、研发合作与创新投入

	(1) <i>CRT</i>	(2) <i>Inn₁</i>	(3) <i>Inn₂</i>
<i>Digas</i>	1.4854*** (5.07)	0.0313*** (4.39)	0.0239*** (6.39)
<i>CRT</i>		0.0027*** (9.63)	0.0013*** (8.74)
常数项	-4.4974***(-7.32)	0.0996*** (4.88)	0.0554*** (7.02)
控制变量	YES	YES	YES
时间固定效应	YES	YES	YES
企业固定效应	YES	YES	YES
<i>N</i>	8 100	8 100	8 100
<i>R</i> ²	0.2112	0.1498	0.1641

七、研究结论和政策启示

数据成为企业和国家重要的战略资产。那么, 如何发挥数据资产对企业创新的价值是当前我国面临的重要问题。本文借助2011至2020年中国A股上市公司数据, 通过量化企业数据资产考察了数据资产对创新投入的影响。研究发现, 数据资产能够显著提升企业的创新投入。数据资产对创新投入的提升作用在中小民营企业、融资依赖度较高的企业、技术密集型企业, 以及金融科技水平低和创新不足的地区的企业更为显著。此外, 数据资产不仅通过缓解企业融资约束提高企业创新意愿, 而且通过加强企业间研发合作拓展企业创新资源, 这些均有利于提升企业的创新投入。对此, 本文提出如下政策建议:

一是应加强数据资产的顶层设计。首先, 完善数据资产确权机制是确保数据资产高效运用的重要前提, 政府应联合企业与科研机构共同建构数据确权总体框架, 明确界定数据资产的权利主体、权利类型等, 从而为数据资产的使用、流通、交易等提供坚实保障。其次, 明晰数据资产的定价规则与交易机制, 将数据要素与传统生产要素的异同纳入定价考量, 建立一次性定价与长期定价相结合的数据资产定价体系, 同时搭建数据资产交易、监管、仲裁等全周期适用的数

据流通平台。最后,加快数据资产和资本市场的融合发展,政府应鼓励资本市场对企业数据资产进行合理的估值,从而能极大地缓解创新型企业的融资约束问题。二是企业在注重数据资产化过程同时,应充分挖掘数据资产的潜在价值。数据资产化过程包括了四个阶段,且每个阶段均会对企业产生价值,这意味着企业在注重建设数据基础设施的同时,应培养数字化人才,释放数据资产潜在价值,提高企业创新活力。同时,数据资产化的关键是企业数字化技术引入,无论是中小企业还是上市公司,均应积极利用数字化技术赋能企业发展,为数据资产的到来奠定技术基础。三是企业要注重数据资产与传统生产要素融合发展。数据资产价值实现的关键在于通过与传统生产要素的融合和联动,使得数据资产赋能商业模式,提升企业产品创意和服务品质。一方面,企业应通过数字化技术实现传统生产要素与数据资产的融合,进一步辅助企业进行生产决策。另一方面,企业要打通内部的数据孤岛,实现数据资产的内部流通,为激发企业创新活力奠定基础。

主要参考文献

- [1]陈国青,曾大军,卫强,等.大数据环境下的决策范式转变与使能创新[J].管理世界,2020,36(2):95-105.
- [2]冯俊诚.减税与减负——来自所得税优惠政策的经验证据[J].经济学(季刊),2022,22(1):67-86.
- [3]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.
- [4]黎文靖,彭远怀,谭有超.知识产权司法保护与企业创新——兼论中国企业创新结构的变迁[J].经济研究,2021,56(5):144-161.
- [5]彭刚,李杰,朱莉.SNA视角下数据资产及其核算问题研究[J].财贸经济,2022,43(5):145-160.
- [6]戚聿东,肖旭.数字经济时代的企业管理变革[J].管理世界,2020,36(6):135-152.
- [7]邱子迅,周亚虹.数字经济发展与地区全要素生产率——基于国家级大数据综合试验区的分析[J].财经研究,2021,47(7):4-17.
- [8]宋敏,周鹏,司海涛.金融科技与企业全要素生产率——“赋能”和信贷配给的视角[J].中国工业经济,2021,(4):138-155.
- [9]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.
- [10]徐涛,尤建新,曾彩霞,等.企业数据资产化实践探索与理论模型构建[J].外国经济与管理,2022,44(6):3-17.
- [11]许宪春,张钟文,胡亚茹.数据资产统计与核算问题研究[J].管理世界,2022,38(2):16-30.
- [12]于瑶,祁怀锦.混合所有制与民营经济健康发展——基于企业违规视角的研究[J].财经研究,2022,48(3):33-47.
- [13]张红娟,申宇,赵晓阳,等.企业外部研发合作、内部知识网络与创新绩效[J].科学学研究,2022,40(4):704-712.
- [14]Acquisti A, Taylor C, Wagman L. The economics of privacy[J]. *Journal of Economic Literature*, 2016, 54(2): 442-492.
- [15]Akcigit U, Liu Q M. The role of information in innovation and competition[J]. *Journal of the European Economic Association*, 2016, 14(4): 828-870.
- [16]Autor D, Dorn D, Katz L F, et al. The fall of the labor share and the rise of superstar firms[J]. *The Quarterly Journal of Economics*, 2020, 135(2): 645-709.
- [17]Begenau J, Farboodi M, Veldkamp L. Big data in finance and the growth of large firms[J]. *Journal of Monetary Economics*, 2018, 97: 71-87.
- [18]Farboodi M, Mihet R, Philippon T, et al. Big data and firm dynamics[J]. *AEA Papers and Proceedings*, 2019, 109: 38-42.
- [19]Farboodi M, Veldkamp L. Long-run growth of financial data technology[J]. *American Economic Review*, 2020, 110(8): 2485-2523.
- [20]Hannila H, Silvola R, Harkonen J, et al. Data-driven begins with DATA; Potential of data assets[J]. *Journal of Computer Information Systems*, 2022, 62(1): 29-38.
- [21]Jones C I, Tonetti C. Nonrivalry and the economics of data[J]. *American Economic Review*, 2020, 110(9): 2819-2858.
- [22]Leonelli S. Data — from objects to assets[J]. *Nature*, 2019, 574(7778): 317-320.

- [23]Nariri R H, Fredericks E M, Bowers K M. Uncertainty in big data analytics: Survey, opportunities, and challenges[J]. *Journal of Big Data*, 2019, 6(1): 44.
- [24]Sivarajah U, Kamal M M, Irani Z, et al. Critical analysis of Big Data challenges and analytical methods[J]. *Journal of Business Research*, 2017, 70: 263-286.

The Impact of Data Assets on Enterprise Innovation Investment

Li Jian¹, Dong Xiaofan², Zhang Jinlin², Tao Yunqing³

(1. *School of Finance, Shandong University of Finance and Economics, Jinan 250014, China*; 2. *School of Finance, Zhongnan University of Economics and Law, Wuhan 430073, China*; 3. *National Development Research Institute / Digital Finance Research Center, Peking University, Beijing 100091, China*)

Summary: In the era of the digital economy, data resources have become a key factor of production. As a new type of core productivity, the assetization process of data has become an inevitable trend, but how to exert the value of data assets has become an urgent problem for enterprises to solve, and it is also a key issue in academic research. Existing literature mainly discusses the value of data assetization from a theoretical perspective, but lacks research on the impact of data assets on enterprise innovation behavior, especially the impact path of data assets on enterprise innovation investment? In this regard, this paper uses data from Chinese A-share listed companies from 2011 to 2020 to quantify enterprise data assets and examine the impact and mechanism of data assets on innovation investment.

The conclusions are as follows: First, data assets can significantly increase enterprise innovation investment, and this improvement effect will be reflected in the process of data assetization, but the impact of data assets on enterprise innovation investment will show the dynamic decay characteristics of time series. Second, the role of data assets in improving innovation investment is more significant in small and medium-sized private enterprises, enterprises with high financing dependence, technology-intensive enterprises, and enterprises in areas with a low financial technology level and insufficient innovation. Third, data assets not only improve enterprises' willingness to innovate by alleviating corporate financing constraints, but also expand enterprise innovation resources by strengthening R&D cooperation between enterprises, which are conducive to enhancing enterprise innovation investment.

Based on the above conclusions, the following policy recommendations are put forward: First, enterprises should strengthen the top-level design of data assets. Second, while paying attention to the process of data assets, enterprises should fully tap the potential value of data assets. Third, enterprises should pay attention to the integration and development of data assets and traditional production factors.

The possible contributions of this paper are that: First, it examines the impact of data assets on enterprise innovation investment from a micro perspective, enriching the research in the micro field of data assets. Second, it quantifies data assets by discussing the process of enterprise data assetization, supplementing the quantitative research of data assets. Third, it reveals the impact path of data assets on enterprise innovation investment from the perspective of innovation willingness and innovation resources, supplementing the theoretical and empirical research of data assets.

Key words: data assets; innovation investment; financing constraints; R&D cooperation

(责任编辑:王雅丽)