

# 数据资产对制造业企业创新能力的研究

刘雪莹, 周 霞

北方工业大学经济管理学院, 北京

收稿日期: 2026年7月1日; 录用日期: 2026年8月3日; 发布日期: 2026年8月13日

## 摘 要

在数字经济与实体经济深度融合背景下, 数据资产逐渐成为制造业企业优化资源配置、提升研发效率和增强创新能力的重要资源。本文以2014~2023年A股制造业上市公司为研究样本, 基于上市公司年报中数据资产相关关键词词频构建数据资产水平代理指标, 并以发明专利申请数量衡量企业创新能力, 在剔除年度ST样本后, 检验数据资产水平与制造业上市公司创新能力之间的关系及其可能作用机制。研究发现, 数据资产水平与制造业上市公司创新能力显著正相关; 在替换变量、使用滞后一期解释变量以及控制企业固定效应后, 结论仍保持稳健。机制检验结果表明, 研发投入和人力资本水平可能是数据资产影响企业创新能力的重要作用路径。异质性分析显示, 在不同技术密集度、所有权性质和地区样本中, 数据资产变量均显著为正, 说明结论在不同分组样本中具有一定稳健性。本文丰富了数据资产与制造业企业创新关系的相关研究, 也为企业加强数据治理、提升数据资源应用能力提供了经验证据。

## 关键词

数据资产, 制造业企业, 创新能力, 研发投入, 人力资本

# Research on the Impact of Data Assets on the Innovation Capability of Manufacturing Enterprises

Xueying Liu, Xia Zhou

School of Economics and Management, North China University of Technology, Beijing

Received: July 1, 2026; accepted: August 3, 2026; published: August 13, 2026

## Abstract

Against the backdrop of the in-depth integration of the digital economy and the real economy, data

文章引用: 刘雪莹, 周霞. 数据资产对制造业企业创新能力的影响研究[J]. 国际会计前沿, 2026, 15(4): 969-978.  
DOI: 10.12677/fia.2026.154101

assets have gradually become a critical resource for manufacturing enterprises to optimize resource allocation, raise R&D efficiency and boost innovation capacity. This paper takes A-share listed manufacturing companies from 2014 to 2023 as research samples. It constructs a proxy indicator for the level of data assets based on the word-frequency of data-asset-related keywords in corporate annual reports, and measures corporate innovation capacity by the number of invention patent applications. After excluding annual ST-status samples, this paper examines the correlation between data-asset level and the innovation capacity of listed manufacturing firms, as well as its potential influencing mechanism. The empirical results demonstrate that the level of data assets is significantly and positively correlated with manufacturing listed firms' innovation capacity. The conclusion remains robust after variable replacement, adoption of the one-period-lagged explanatory variable and control for firm fixed-effects. Mechanism tests indicate that R&D investment and human capital level serve as vital transmission paths through which data assets affect corporate innovation capability. Heterogeneity analysis shows that the coefficient of the data-asset variable stays significantly positive across subgroups with varied technology intensity, ownership nature and geographic locations, which verifies the robustness of our findings among grouped samples. This study expands existing research concerning the nexus between data assets and manufacturing-firm innovation, and furnishes empirical evidence for enterprises to strengthen data governance and improve the application capability of data resources.

## Keywords

Data Assets, Manufacturing Enterprises, Innovation Capability, R&D Investment, Human Capital

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

随着数字技术的深入发展,数据已成为影响经济增长、产业升级和企业创新的重要生产要素。企业在生产经营、研发设计、供应链管理、客户服务和市场反馈过程中积累了大量数据资源。随着数据治理能力、数字技术应用水平和数据分析能力不断提升,部分数据资源逐渐具备可识别、可控制、可利用和可增值的资产属性,成为企业提升经营效率和形成竞争优势的重要基础。

制造业是实体经济的核心领域,也是数字技术与实体经济深度融合的关键场景。与服务业相比,制造业企业的数据来源更加嵌入生产和研发过程,包括生产数据、设备运行数据、工艺参数数据、质量检测数据、供应链数据、研发试验数据和客户反馈数据等。这些数据并非简单的信息记录,而是企业识别技术路线、优化产品设计、改进生产工艺和提升供应链协同能力的重要依据。

基于此,本文以2014~2023年A股制造业上市公司为研究对象,借鉴已有研究中利用上市公司年报文本刻画企业数字化转型和信息披露特征的做法,基于年报中数据资产相关关键词词频构建企业数据资产水平的代理指标,并以发明专利申请数量衡量企业创新能力,考察数据资产水平与制造业上市公司创新能力之间的关系及其可能作用机制。

## 2. 文献综述

### 2.1. 数据资产内涵及经济后果研究

随着数据要素市场化配置和企业数据资源入表制度推进,数据资产的内涵、确认、披露及其经济后

果逐渐成为会计和数字经济研究的重要议题。《企业数据资源相关会计处理暂行规定》<sup>1</sup>, 为企业数据资源在存货、无形资产等项目中的会计处理提供了制度依据; 国家数据局等部门关于印发《“数据要素×”三年行动计划(2024—2026年)》<sup>2</sup>, 强调推动数据要素在工业制造、科技创新等场景中发挥乘数效应。

围绕数据资产经济后果, 王海芳等[1]以 A 股制造业上市公司为样本, 发现数据资产与制造业企业供应链韧性存在显著关联; 王兴启等[2]从动态能力视角研究发现, 数据资产有助于制造业企业价值链升级; 吕寒和王坤[3]进一步发现, 数据资产化与企业资源配置效率提升相关。总体来看, 已有研究表明, 数据资产不仅是信息资源, 也可能成为影响企业资源配置、组织协同和价值创造的重要战略资源。

## 2.2. 数据资产与企业创新研究

企业创新是企业识别、吸收、整合并应用知识资源的过程。数据资产能够为企业 提供市场需求、技术趋势、生产反馈和供应链协同等信息基础, 因此已有文献开始从创新绩效、突破式创新、融通创新和 创新质效等角度检验数据资产的创新效应。

胡海晨和刘斌燕[4]基于资源基础理论研究发现, 数据资产与企业创新绩效显著正相关, 技术吸收能力在其中发挥传导作用; 崔晓蕾和耿倩[5]以 A 股上市公司为样本, 发现数据资产与企业创新水平提升相关, 其机制可能包括缓解融资约束和推动人工智能应用; 霍远和李文龙[6]从突破式创新视角发现, 数据资产化与企业突破式创新显著相关; 朱磊等[7]则发现数据资产能够通过打破信息壁垒和促进知识流动影响企业融通创新。上述研究为本文分析数据资产与制造业企业创新能力的关系提供了重要依据。

## 2.3. 数据资产影响企业创新的作用机制研究

数据资产影响企业创新并非孤立发生, 而是依赖研发资源投入、知识吸收能力、人力资本结构和数字化能力等条件。吕寒和王坤[3]指出, 数据资产化与资源配置效率改善相关; 霍远和李文龙[6]发现, 研发投入增加可能是数据资产化影响突破式创新的重要路径; 赵鑫宇等[8]从创新质效视角表明, 数据资产化与企业创新质量和效率提升存在显著关联。

制造业企业的数据资产价值实现需要研发人员、技术人员和管理人员共同参与, 较高的人力资本水平有助于将分散的数据资源转化为可用于研发设计、工艺优化和产品创新的知识基础。现有研究为本文从研发投入和人力资本两个方面检验数据资产影响企业创新能力的作用路径提供了理论支撑。

## 3. 研究假设

已有研究表明, 数据资产与企业创新之间存在较为稳定的正向关联。胡海晨和刘斌燕[4]发现, 数据资产有助于提升企业创新绩效, 技术吸收能力在其中发挥传导作用; 崔晓蕾和耿倩[5]指出, 数据资产可能通过缓解融资约束和推动人工智能应用改善企业创新表现; 朱磊等[7]从融通创新角度发现, 数据资产能够通过打破信息壁垒和促进知识流动影响企业创新。制造业企业在研发设计、生产控制、工艺优化和供应链协同过程中积累了大量数据资源, 这些资源经过治理和应用后, 能够为企业技术路线识别、产品设计优化和研发决策改进提供信息支持。基于此, 提出假设:

H1: 数据资产水平越高, 制造业上市公司的创新能力越强。

研发投入是企业创新产出的重要资源基础。已有研究指出, 数据资产化与企业资源配置效率提升相关[3], 并可能通过增加研发投入影响突破式创新[6]。从制造业场景看, 生产反馈数据、质量检测数据、客户需求数据和供应链数据能够帮助企业更准确识别研发方向和技术路线, 减少研发资源错配, 提高研

<sup>1</sup>[https://www.gov.cn/zhengce/zhengceku/202308/content\\_6899395.htm](https://www.gov.cn/zhengce/zhengceku/202308/content_6899395.htm)

<sup>2</sup>[https://www.nda.gov.cn/sjj/zwgk/zcfb/0830/20240830174038137859023\\_pc.html](https://www.nda.gov.cn/sjj/zwgk/zcfb/0830/20240830174038137859023_pc.html)

发投入的针对性。基于此, 提出假设:

H2: 数据资产能够通过提高研发投入促进制造业上市公司创新能力提升。

数据资产价值实现离不开高素质人才。赵鑫宇等[8]从创新质效角度指出, 人力资本结构可能是数据资产化影响创新的重要条件。制造业企业的数据采集、清洗、建模、分析和应用需要数字化人才、研发人员和技术人员共同参与, 较高的人力资本水平有助于将数据资源转化为创新产出。基于此, 提出假设:

H3: 数据资产能够通过提升人力资本水平促进制造业上市公司创新能力提升。

## 4. 研究设计

### 4.1. 样本选择与数据来源

本文选取 2014~2023 年 A 股制造业上市公司为研究样本, 企业财务、公司治理、研发创新、员工结构、行业分类及上市相关数据来源于 CSMAR 数据库, 专利数据取自 CNRDS 数据库, 核心数据资产指标通过上市公司年报文本关键词统计获得。由于 CNRDS 专利数据目前覆盖至 2023 年, 正式实证样本区间确定为 2014~2023 年。为保证研究结果的可靠性, 对初始样本进行如下筛选处理: 保留证监会行业分类中行业代码以 C 开头的制造业上市公司样本; 剔除样本期内被标记为年度 ST 的异常经营企业 - 年份观测值; 剔除主回归模型中核心变量存在数据缺失的企业 - 年份观测值; 对连续变量进行上下 1% 缩尾处理。

年度 ST 样本根据 CSMAR 个股停牌标识表(日)中的 ST 标识构建。若某公司某年度任一记录中 ST = Y, 则将该公司该年度标记为 ST 样本并予以剔除; 未匹配到 ST 标识记录的样本暂按非 ST 处理。由于不同数据库在年报文本、专利数据、员工结构和公司治理变量方面存在一定缺失, 本文在描述性统计和相关性分析中采用主回归所需变量均不缺失的样本, 以保证统计口径与基准回归保持一致。机制变量 RD 和 Human 的样本量分别在机制检验表中列示。

### 4.2. 变量定义

被解释变量为企业创新能力(Innovation), 采用企业当年发明专利申请数量加 1 后取自然对数衡量。发明专利相较于实用新型和外观设计专利更能体现企业实质性技术创新水平。稳健性检验中, 本文采用发明专利授权数量加 1 后取自然对数(Innovation\_auth)作为替代指标。

核心解释变量为数据资产水平(DA)。本文基于上市公司年报文本中数据资产相关关键词词频构建, 具体为  $\ln(1 + \text{数据资产相关关键词总词频})$ 。稳健性检验中进一步使用标准化数据资产指标(DA\_std), 即关键词总词频除以年报文本字符数, 以缓解年报篇幅差异带来的影响。该指标主要反映企业对数据资产相关活动的披露程度和重视程度。

机制变量包括研发投入(RD)和人力资本水平(Human)。研发投入采用研发投入金额除以营业收入衡量, 反映企业对研发创新活动的资源投入强度; 人力资本水平采用本科及以上学历员工人数占员工总数的比例衡量, 反映企业员工知识结构和高素质人才储备。

控制变量包括企业规模(Size)、资产负债率(Lev)、盈利能力(ROA)、企业成长性(Growth)、企业年龄(Age)、董事会规模(Board)和第一大股东持股比例(Top1)。这些变量分别反映企业资源获取能力、融资和风险承担能力、盈利能力、发展潜力、经验积累以及公司治理特征(表 1)。

### 4.3. 模型设定

为检验数据资产水平与制造业上市公司创新能力之间的关系, 构建如下基准模型:

Table 1. Variable definitions  
表 1. 变量定义表

| 变量类型  | 变量名称      | 变量符号       | 变量定义                 |
|-------|-----------|------------|----------------------|
| 被解释变量 | 企业创新能力    | Innovation | ln(1 + 发明专利申请数量)     |
| 解释变量  | 数据资产      | DA         | ln(1 + 数据资产相关关键词总词频) |
| 机制变量  | 研发投入      | RD         | 研发投入金额/营业收入          |
| 机制变量  | 人力资本水平    | Human      | 本科及以上学历员工人数/员工总数     |
| 控制变量  | 企业规模      | Size       | ln(资产总计)             |
| 控制变量  | 资产负债率     | Lev        | 负债合计/资产总计            |
| 控制变量  | 盈利能力      | ROA        | 净利润/资产总计             |
| 控制变量  | 企业成长性     | Growth     | 营业收入增长率              |
| 控制变量  | 企业年龄      | Age        | ln(1 + 年份 - 上市年份)    |
| 控制变量  | 董事会规模     | Board      | ln(董事人数)             |
| 控制变量  | 第一大股东持股比例 | Top1       | 第一大股东持股比例            |

注：Innovation 为企业创新能力，DA 为数据资产水平。连续变量均经过上下 1%缩尾处理。

$$\text{Innovation}_{i,t} = \alpha_0 + \alpha_1 \text{DA}_{i,t} + \alpha_2 \text{Controls}_{i,t} + \text{IndustryFE} + \text{YearFE} + \varepsilon_{i,t}$$

其中，Innovation\_it 表示企业 *i* 在第 *t* 年的创新能力，DA\_it 表示企业 *i* 在第 *t* 年的数据资产水平，Controls\_it 表示控制变量，IndustryFE 表示行业固定效应，YearFE 表示年份固定效应。为了缓解异方差和序列相关问题对估计结果的影响，本文在回归中采用按企业聚类的稳健标准误。

5. 实证结果与分析

5.1. 描述性统计(表 2)

Table 2. Descriptive statistics  
表 2. 描述性统计

| 变量         | 样本量    | 均值      | 标准差     | 最小值     | 中位数     | 最大值     |
|------------|--------|---------|---------|---------|---------|---------|
| Innovation | 20,441 | 3.0641  | 1.6611  | 0.0000  | 3.1355  | 7.0660  |
| DA         | 20,441 | 2.0507  | 1.2966  | 0.0000  | 2.0794  | 5.0720  |
| Size       | 20,441 | 22.1536 | 1.1657  | 19.1260 | 22.0005 | 25.6506 |
| Lev        | 20,441 | 0.3953  | 0.1916  | 0.0548  | 0.3862  | 0.9017  |
| ROA        | 20,441 | 0.0347  | 0.0707  | -0.2718 | 0.0379  | 0.2168  |
| Growth     | 20,441 | 0.1410  | 0.3421  | -0.5343 | 0.0914  | 1.8652  |
| Age        | 20,441 | 2.0974  | 0.8192  | 0.0000  | 2.1972  | 3.3673  |
| Board      | 20,441 | 2.0887  | 0.1892  | 1.6094  | 2.1972  | 2.5193  |
| Top1       | 20,441 | 32.1743 | 13.8166 | 8.7818  | 30.0000 | 72.3800 |

注：表中变量均基于主回归样本计算，连续变量均经过上下 1%缩尾处理。

5.2. 相关性分析

相关性分析显示，数据资产水平与企业创新能力之间呈正相关关系，为后续回归检验提供了初步依据(表 3)。

**Table 3.** Correlation analysis**表 3.** 相关性分析

| 变量         | Innovation | DA      | Size   | Lev     | ROA     | Growth  | Age     | Board   | Top1    |
|------------|------------|---------|--------|---------|---------|---------|---------|---------|---------|
| Innovation | 1.0000     | 0.3122  | 0.5323 | 0.1989  | 0.1070  | 0.0821  | 0.1393  | 0.1422  | 0.0221  |
| DA         | 0.3122     | 1.0000  | 0.1494 | 0.0648  | 0.0103  | 0.0076  | -0.0503 | -0.0374 | -0.0219 |
| Size       | 0.5323     | 0.1494  | 1.0000 | 0.4205  | 0.0761  | 0.0674  | 0.4316  | 0.2492  | 0.0895  |
| Lev        | 0.1989     | 0.0648  | 0.4205 | 1.0000  | -0.3960 | 0.0228  | 0.3248  | 0.1002  | -0.0516 |
| ROA        | 0.1070     | 0.0103  | 0.0761 | -0.3960 | 1.0000  | 0.2932  | -0.1784 | 0.0374  | 0.1785  |
| Growth     | 0.0821     | 0.0076  | 0.0674 | 0.0228  | 0.2932  | 1.0000  | -0.1039 | 0.0123  | 0.0105  |
| Age        | 0.1393     | -0.0503 | 0.4316 | 0.3248  | -0.1784 | -0.1039 | 1.0000  | 0.1723  | -0.1385 |
| Board      | 0.1422     | -0.0374 | 0.2492 | 0.1002  | 0.0374  | 0.0123  | 0.1723  | 1.0000  | -0.0378 |
| Top1       | 0.0221     | -0.0219 | 0.0895 | -0.0516 | 0.1785  | 0.0105  | -0.1385 | -0.0378 | 1.0000  |

注：表中为 Pearson 相关系数，变量均基于主回归样本计算。

### 5.3. 基准回归

表 4 报告了剔除年度 ST 样本后的基准回归结果。模型 1 仅控制年份固定效应，模型 2 进一步加入控制变量，模型 3 同时控制行业固定效应和年份固定效应，模型 4 控制企业固定效应和年份固定效应。结果显示，在控制年份和行业固定效应后，DA 的系数为 0.2224，且在 1%水平显著；在企业固定效应模型中，DA 的系数为 0.0735，仍在 1%水平显著。上述结果表明，数据资产水平与制造业上市公司创新能力显著正相关，假设 H1 得到支持。

**Table 4.** Benchmark regression results**表 4.** 基准回归结果

| 变量     | 模型 1                  | 模型 2                   | 模型 3                   | 模型 4                  |
|--------|-----------------------|------------------------|------------------------|-----------------------|
| DA     | 0.4356***<br>(0.0176) | 0.3363***<br>(0.0148)  | 0.2224***<br>(0.0151)  | 0.0735***<br>(0.0135) |
| Size   |                       | 0.7538***<br>(0.0218)  | 0.7796***<br>(0.0199)  | 0.5964***<br>(0.0378) |
| Lev    |                       | -0.0512<br>(0.1234)    | -0.2065*<br>(0.1134)   | -0.2901**<br>(0.1138) |
| ROA    |                       | 1.0507***<br>(0.2716)  | 1.2860***<br>(0.2385)  | 0.1375<br>(0.1722)    |
| Growth |                       | 0.0279<br>(0.0337)     | -0.0259<br>(0.0308)    | 0.0183<br>(0.0230)    |
| Age    |                       | -0.1524***<br>(0.0258) | -0.0750***<br>(0.0240) | -0.0729<br>(0.0455)   |
| Board  |                       | 0.1878*<br>(0.0988)    | 0.2470***<br>(0.0912)  | 0.1316<br>(0.0866)    |
| Top1   |                       | -0.0050***<br>(0.0015) | -0.0014<br>(0.0014)    | 0.0041*<br>(0.0025)   |
| 年份固定效应 | 是                     | 是                      | 是                      | 是                     |
| 行业固定效应 | 否                     | 否                      | 是                      | 否                     |

续表

|                |        |        |        |        |
|----------------|--------|--------|--------|--------|
| 企业固定效应         | 否      | 否      | 否      | 是      |
| 样本量            | 24,052 | 20,441 | 20,441 | 20,441 |
| R <sup>2</sup> | 0.0971 | 0.3624 | 0.3673 | 0.0591 |

注：括号内为按企业聚类稳健标准误；\*\*\*、\*\*、\*分别表示在 1%、5%、10%水平显著；回归中连续变量均采用上下 1% 缩尾后的变量。模型 1 未加入控制变量，使用 Innovation 和 DA 均不缺失的样本；模型 2~4 加入控制变量后使用主回归样本。

5.4. 稳健性检验

为了检验基准回归结果的稳健性，本文从替换被解释变量、替换核心解释变量、使用滞后一期解释变量以及控制企业固定效应四个方面进行检验。结果显示，将创新能力替换为发明专利授权数量后，核心变量系数仍显著为正；采用标准化关键词词频衡量数据资产水平后，核心变量系数仍显著为正；使用滞后一期数据资产变量后，结果仍保持一致。上述检验表明，本文结论具有一定稳健性(表 5)。

Table 5. Robustness test results  
表 5. 稳健性检验结果

| 变量             | Innovation_auth       | DA_std                     | L_DA                  | FirmFE                |
|----------------|-----------------------|----------------------------|-----------------------|-----------------------|
| 核心变量           | 0.1834***<br>(0.0145) | 1347.0750***<br>(117.2658) | 0.2201***<br>(0.0150) | 0.0735***<br>(0.0135) |
| 控制变量           | 是                     | 是                          | 是                     | 是                     |
| 年份固定效应         | 是                     | 是                          | 是                     | 是                     |
| 行业固定效应         | 是                     | 是                          | 是                     | 否                     |
| 企业固定效应         | 否                     | 否                          | 否                     | 是                     |
| 样本量            | 20,441                | 20,441                     | 20,380                | 20,441                |
| R <sup>2</sup> | 0.3323                | 0.3572                     | 0.3658                | 0.0591                |

注：括号内为按企业聚类稳健标准误；\*\*\*、\*\*、\*分别表示在 1%、5%、10%水平显著；回归中连续变量均采用上下 1% 缩尾后的变量。采用标准化关键词词频衡量数据资产水平时，因变量尺度较小，本文仅关注系数方向和显著性。

5.5. 机制检验

表 6 显示，剔除年度 ST 样本后，DA 对 RD 为正且显著，系数为 0.0050；在进一步将 RD 纳入创新能力回归后，RD 对 Innovation 为正且显著，系数为 6.7463。这说明研发投入可能是数据资产影响制造业上市公司创新能力的重要作用路径。

表 7 显示，剔除年度 ST 样本后，DA 对 Human 为正且显著，系数为 0.0328；在进一步将 Human 纳入创新能力回归后，Human 对 Innovation 为正且显著，系数为 1.1270。这说明人力资本水平可能是数据资产影响制造业上市公司创新能力的另一重要作用路径。

5.6. 异质性分析

为检验数据资产对制造业企业创新能力影响的情境稳健性，本文从行业技术密集度、产权性质和区域特征三个维度开展异质性分析。所有回归均控制企业特征、行业与年份固定效应，并采用企业聚类稳健标准误。研究发现，数据资产对创新的正向影响在不同分组中均保持显著，核心结论具有普适性。

**Table 6.** Mechanism test of R&D investment  
**表 6.** 研发投入机制检验

| 变量             | RD                    | Innovation            |
|----------------|-----------------------|-----------------------|
| DA             | 0.0050***<br>(0.0006) | 0.1847***<br>(0.0146) |
| RD             |                       | 6.7463***<br>(0.3990) |
| 控制变量           | 是                     | 是                     |
| 行业固定效应         | 是                     | 是                     |
| 年份固定效应         | 是                     | 是                     |
| 样本量            | 20,010                | 20,010                |
| R <sup>2</sup> | 0.1169                | 0.3920                |

注：括号内为按企业聚类稳健标准误；\*\*\*、\*\*、\*分别表示在 1%、5%、10% 水平显著；回归中连续变量均采用上下 1%缩尾后的变量。

**Table 7.** Mechanism test of human capital  
**表 7.** 人力资本机制检验

| 变量             | Human                 | Innovation            |
|----------------|-----------------------|-----------------------|
| DA             | 0.0328***<br>(0.0025) | 0.1814***<br>(0.0153) |
| Human          |                       | 1.1270***<br>(0.0938) |
| 控制变量           | 是                     | 是                     |
| 行业固定效应         | 是                     | 是                     |
| 年份固定效应         | 是                     | 是                     |
| 样本量            | 19,817                | 19,817                |
| R <sup>2</sup> | 0.0550                | 0.3840                |

注：括号内为按企业聚类稳健标准误；\*\*\*、\*\*、\*分别表示在 1%、5%、10% 水平显著；回归中连续变量均采用上下 1%缩尾后的变量。

从行业技术密集度来看，参考行业分类标准，本文将 C27、C35~C40 等行业界定为高技术制造业，其余为中低技术制造业。回归结果显示，高技术制造业与中低技术制造业中数据资产系数分别为 0.2187 和 0.2261，均在 1%水平上显著为正，表明数据资产对创新的促进作用不受行业技术属性影响。

从产权性质和区域特征来看，依据 CSMAR 数据划分国有企业与非国有企业、东部地区与中西部地区，并剔除信息缺失样本。结果显示，国有企业、非国有企业、东部地区及中西部地区的系数分别为 0.2184、0.2330、0.2185 和 0.2378，均通过 1%显著性检验，说明产权性质与区域差异未改变数据资产对创新的正向效应(表 8)。

6. 研究结论与建议

6.1. 研究结论

本文以 2014~2023 年 A 股制造业上市公司为研究样本，在剔除年度 ST 样本后，基于上市公司年报

文本关键词词频构建数据资产指标，并以发明专利申请数量衡量企业创新能力，考察数据资产水平与制造业上市公司创新能力之间的关系。实证结果表明，数据资产水平与制造业上市公司创新能力显著正相关。在替换被解释变量、替换核心解释变量、使用滞后一期解释变量以及企业固定效应模型后，上述结论仍然保持稳健。

Table 8. Heterogeneity analysis results  
表 8. 异质性分析结果

| 分组类型    | 分组名称    | DA 系数     | 聚类<br>稳健标准误 | 样本量    | R <sup>2</sup> | 控制变量 | 行业<br>固定效应 | 年份<br>固定效应 |
|---------|---------|-----------|-------------|--------|----------------|------|------------|------------|
| 行业技术密集度 | 高技术制造业  | 0.2187*** | 0.0193      | 11,094 | 0.4569         | 是    | 是          | 是          |
| 行业技术密集度 | 中低技术制造业 | 0.2261*** | 0.0238      | 9347   | 0.2748         | 是    | 是          | 是          |
| 所有权性质   | 国有企业    | 0.2184*** | 0.0339      | 4741   | 0.4174         | 是    | 是          | 是          |
| 所有权性质   | 非国有企业   | 0.2330*** | 0.0163      | 15,178 | 0.3021         | 是    | 是          | 是          |
| 地区      | 东部地区    | 0.2185*** | 0.0171      | 14,786 | 0.3811         | 是    | 是          | 是          |
| 地区      | 中西部地区   | 0.2378*** | 0.0314      | 5653   | 0.3451         | 是    | 是          | 是          |

注：括号内为按企业聚类稳健标准误；\*\*\*、\*\*、\*分别表示在 1%、5%、10%水平显著。

机制检验结果表明，数据资产水平与企业研发投入显著正相关，研发投入与创新能力显著正相关；同时，数据资产水平与人力资本水平显著正相关，人力资本水平与创新能力显著正相关。上述结果为研发投入和人力资本水平可能是数据资产影响制造业上市公司创新能力的重要作用路径提供了经验证据。进一步的异质性分析表明，在高技术制造业与中低技术制造业、国有企业与非国有企业、东部地区与中西部地区样本中，数据资产变量均显著为正，说明本文结论在不同分组样本中具有一定稳健性。

6.2. 实践建议

从企业层面看，制造业上市公司应重视数据资产积累和数据治理能力建设，建立覆盖数据采集、存储、清洗、分析和应用的管理体系，提高数据质量和数据可用性。同时，企业应将数据资产更好服务于研发设计、智能制造、工艺优化和产品创新等场景，提高研发资源配置效率，并加强数字化人才、研发人员和技术人员培养，提升数据资源转化为创新成果的能力。

从政策层面看，可完善数据要素市场与数据资产披露制度，健全数据确权、流通、交易及安全治理规则，为制造业企业开展数据驱动创新提供制度支撑。同时，可为制造业数字化转型、工业互联网建设及数据治理能力提升提供政策支持，推动数据要素更好服务实体经济与制造业高质量发展。

6.3. 研究不足与展望

本文仍存在以下不足。第一，数据资产指标基于年报关键词词频构建，更多反映企业披露水平和重视程度，不能完全代表真实数据资产质量、数据治理能力和数据应用深度。第二，创新能力主要以专利数量衡量，难以完全反映创新质量、产品创新、工艺创新和管理创新。第三，虽然本文进行了替换变量、滞后一期解释变量和企业固定效应等稳健性检验，但仍可能存在遗漏变量、反向因果等内生性问题，未来可采用工具变量、PSM、DID 等方法进一步识别因果关系。第四，ST 标识数据对部分公司 - 年份样本未完全匹配，本文对未匹配样本暂按非 ST 处理，未来可结合更完整的年度风险警示状态数据进行样本筛选。第五，制造业内部存在行业差异，未来可进一步区分高技术制造业、装备制造业和传统制造业开展

异质性分析。第六，本文异质性分析主要采用分组回归方式，尚未进一步进行组间系数差异检验，未来可结合更严格的交互项模型或组间差异检验进一步识别不同情境下作用强度的差异。

## 致谢

在本论文的撰写和研究过程中，得到了多方面的支持与帮助。谨此向所有关心、指导和帮助过我的老师、同学及相关单位致以衷心的感谢。

衷心感谢我的导师周霞老师。从论文选题、研究思路、框架设计到内容修改与定稿，周老师始终给予悉心指导和宝贵建议，使我得以不断完善研究，顺利完成论文。

感谢评阅专家提出的宝贵修改意见和建议，这些意见对提升论文质量帮助良多。同时，感谢课题组各位同学在资料收集、问题讨论和论文修改过程中给予的支持与帮助。

本文研究得到了教育部人文社科项目“创新券提升中小微企业创新能力的协同机制及优化路径研究”(项目编号：19YJA630127)的资助，在此表示诚挚的感谢。

最后，感谢本文所引用文献、资料及相关研究成果的作者。正是前人的研究为本文的写作提供了重要的理论基础和研究启发。

## 基金项目

教育部人文社科创新券提升中小微企业创新能力的协同机制及优化路径研究项目(19YJA630127)。

## 参考文献

- [1] 王海芳, 范文静, 邓美玲, 李梦沅. 数据资产与制造业企业供应链韧性[J]. 财会月刊, 2026, 47(8): 54-61.
- [2] 王兴启, 李吉, 刘义超, 周云波. 数据资产、动态能力与制造业企业价值链升级[J]. 证券市场导报, 2026(4): 13-23+37.
- [3] 吕寒, 王坤. 数据资产化对企业资源配置效率的影响[J]. 科技管理研究, 2026, 46(1): 231-244.
- [4] 胡海晨, 刘斌燕. 数据资产、技术吸收能力与企业创新绩效[J]. 技术经济与管理研究, 2025(12): 108-115.
- [5] 崔晓蕾, 耿倩. 数据资产对企业创新的影响研究[J/OL]. 经营与管理, 1-13.  
<https://doi.org/10.16517/j.cnki.cn12-1034/f.20260110.004>, 2026-05-07.
- [6] 霍远, 李文龙. 数据资产化能否影响企业突破式创新[J]. 财会月刊, 2025, 46(21): 30-37.
- [7] 朱磊, 李彤, 王春燕, 张茜. 数据资产对企业融通创新的影响研究[J/OL]. 科学学与科学技术管理, 1-31.  
<https://doi.org/10.20201/j.cnki.ssstm.20251202.003>, 2026-05-07.
- [8] 赵鑫宇, 肖明, 李格. 数据资产化与企业创新质效: 理论机制与实证检验[J]. 统计与决策, 2026(8): 177-182.