

DOI:10.3969/j.issn.1674-8131.2024.03.005

数据资产如何赋能企业高质量发展

——对传统生产要素的优化机制

苑泽明¹, 尹琪¹, 于翔²

(1. 天津财经大学 会计学院, 天津 300222; 2. 南开大学 商学院, 天津 300071)

摘要: 数据资产的资产属性和数据属性决定了其可以改善企业传统生产要素的配置状态, 从而赋能企业高质量发展。根据企业年报, 运用机器学习算法, 通过“种子词筛选+AutoPhrase 相似词扩充+SnowNLP 情感极性分析”对企业的数据资产水平进行评估, 以全要素生产率衡量企业高质量发展水平, 进而采用沪深 A 股上市公司 2007—2022 年的数据进行分析, 结果发现: 数据资产增加显著促进了企业高质量发展, 该作用主要由数据资产的处理和应用产生, 数据资产获取对企业高质量发展的影响不显著; 数据资产增加能够改善劳动力结构、提高技术水平、抑制非效率投资、降低管理成本, 从而通过对传统生产要素的优化机制来赋能企业高质量发展; 外部环境对数据资产赋能企业高质量发展具有调节作用, 表现为环境不确定性、信息化水平、市场化水平和数字经济政策供给水平的提高会强化数据资产对企业高质量发展的促进作用。因此, 企业应积极开发、积累和利用数据资产, 政府应持续推进地区信息化、市场化和数字经济发展, 以充分发挥数据资产的积极效应。

关键词: 数据资产; 高质量发展; 传统生产要素; 劳动力结构; 资本配置效率; 技术水平; 管理效率

中图分类号: F27; F273.4 文献标志码: A 文章编号: 1674-8131(2024)03-0054-20

引用格式: 苑泽明, 尹琪, 于翔. 数据资产如何赋能企业高质量发展——对传统生产要素的优化机制[J]. 西部论坛, 2024, 34(3): 54-73.

YUAN Ze-ming, YIN Qi, YU Xiang. Empowering high-quality development of enterprises with data assets: Optimization mechanism of traditional production factors[J]. West Forum, 2024, 34(3): 54-73.

* 收稿日期: 2024-02-28; 修回日期: 2024-04-29

基金项目: 国家社会科学基金一般项目(23BGL109)

作者简介: 苑泽明(1964), 女, 天津人; 教授, 博士生导师, 主要从事环境会计、无形资产等研究; Tel: 13820558731, E-mail: yuanyezeming@tjufe.edu.cn。尹琪(2000), 通信作者, 女, 山西太原人; 硕士研究生, 主要从事数据资产研究; Tel: 15534401011, E-mail: yinqi06025024@163.com。于翔, 女, 山东烟台人; 博士研究生, 主要从事公司治理研究; Tel: 18406557159, E-mail: 1046205890@qq.com。

一、引言

在数智时代,随着数字技术和数字经济的快速发展,数据资源呈现出指数级增长态势,推动着社会生产方式、生活方式发生深刻变革。在数字经济形态下,数据驱动的新产品、新服务、新业态和新模式不断完善和发展,数据成为一种新型生产要素,并催生出企业经济效益的新增长点,已经具备在财务报告中被确认为资产的条件,即形成了数据资产(孙颖等,2021)^[1]。截至2023年,中国信息通信研究院云计算与大数据研究所发布的《数据资产管理实践白皮书》已更新至6.0版本,其指出数据资产是指由企业拥有或控制的,能够为企业带来未来经济利益的,以物理或电子方式记录的数据资源。在此背景下,如何有效利用数据资产、激活数据资产潜能、挖掘数据资产生产力,成为国家和企业培育发展新动能、构筑竞争新优势的重要命题。与此同时,党的二十大报告指出,高质量发展是全面建设社会主义现代化国家的首要任务。企业是国民经济发展的微观基础,经济高质量发展首先需要实现企业的高质量发展。企业高质量发展意味着其生产函数整体处于最优状态,即各类生产要素在限定条件下实现利益最大化的最优配置(阳镇,2023)^[2]。数据成为新的生产要素,改变了企业的生产函数,此时,数据资产不仅本身会带来丰厚的经济利益,而且还会通过促进传统生产要素的迭代和流动来优化生产要素的整体配置,从而推动企业高质量发展。因此,深入研究数字资产对企业高质量发展的影响及其机制,有助于进一步充分发挥数据要素的积极作用,进而带动经济高质量发展。

关于企业高质量发展的影响因素,已有文献从企业特征及外部环境等方面进行了广泛研究,如营商环境(周泽将等,2022)^[3]、政府补贴(郑飞等,2022)^[4]、企业创新(陈丽珊等,2019)^[5]、高管特征(李雄飞,2022)^[6]、杠杆率(宋清华等,2021)^[7]等。随着数字经济的发展,大量文献就数字经济(刘艳霞,2022;阳镇,2023)^{[8][2]}、数字金融(张超等,2022)^[9]、数字技术(李朝鲜,2022;黄勃等,2023;栗晓云等,2023)^[10-12]、数字政府(岳宇君等,2023)^[13]、数字化转型(吴浩强等,2023;高跃等,2023;王羲等,2024)^[14-16]等企业高质量发展的影响进行了探讨,但少有研究考察数据要素或数据资产对企业高质量发展的影响。孙颖和陈思霞(2021)^[1]研究发现,数据资产促进了科技服务企业的高质量发展;杨向阳和徐从才(2024)^[17]分析表明,数据要素主要通过降本增效和提质增效两条路径促进了流通企业的高质量发展。可见,有必要进一步深入研究数字资产对企业高质量发展的影响及其机制。

近年来,数据资产受到学术界的高度关注,成为研究热点之一,然而,关于数据资产的经济效应研究相对较少。其中,在对企业发展及行为的影响方面,数据资产能够通过缓解融资约束和加强研发合作来提高企业创新意愿和拓展企业创新资源,从而增加企业创新投入(李健等,2023)^[18],并通过嵌入有价值的信息来提高企业运营效率(Hu et al.,2022)^[19],作用于企业的成本管理、组织建设和风险管控(Kiron,2017;Sivarajah et al.,2020)^[20-21],进而增加企业经济利润(Begenau et al.,2018)^[22],提升企业价值(苑泽明等,2022)^[23],促进企业高质量发展(孙颖等2021)^[1]。此外,路征等(2023)^[24]研究发现,数据资产通过降低信息不对称程度、降低交易成本和提高企业创新能力三条路径促进了企业发展;陆岷峰等(2023)^[25]分析表明,数据资产通过提高产品市场份额、抑制管理层跟风投资、提升资源配置效率、助力企业转型升级等路径提高了企业的产能利用率。但这些研究均未从数据资产对传统生产要素的影响视角来探究这一新型生产要素对企业高质量发展“赋能能”的问题。

作为数字经济时代的新型生产要素,数据资产依托于数字产业化和产业数字化,嵌入实体经济的各个领域,与劳动力、资本、自然资源、技术、管理等传统生产要素在价值链和产业链中深度融合,产生“鲶鱼效应”,促进企业的人力资源和资本积累、技术进步以及管理提效,从而显著提高企业的全要素生产

率,有效推动企业高质量发展。因此,数据资产的价值创造是建立在优化传统生产要素的基础之上的,对于数据资产赋能企业高质量发展的研究,有必要从改善生产要素配置的视角出发,探究数据资产对传统生产要素的优化机制。有鉴于此,本文在已有研究的基础上,探讨数据资产通过优化传统生产要素来促进企业高质量发展的机制,并采用沪深A股上市公司2007—2022年的数据进行实证检验。

与已有文献相比,本文的边际贡献主要在于:第一,考察了数据资产对企业高质量发展的影响,并从对劳动力、资本、技术、管理4种传统生产要素的优化角度展开机制分析(劳动力结构改善机制、资本配置效率提升机制、技术创新积累机制、管理提质增效机制),拓展和深化了数据资产的经济效应研究,并为数据资产对传统生产要素的改善效应和对企业高质量发展的促进作用提供了新的经验证据,有助于深入认识和理解数字经济形态下数据资产的积极功效。第二,基于数据资产“获取—处理—应用”的底层逻辑,采用“种子词筛选+AutoPhrase相似词扩充+SnowNLP情感极性分析”的方法对企业数据资产水平进行测算,并构建了包括“数据资产获取”“数据资产处理”“数据资产应用”3个层面的数据资产特征词谱,为客观评估企业的数据资产水平提供了一种有效方法。第三,进一步分析了环境不确定性、城市信息化水平、地区市场化水平以及数字经济政策等外部环境对数据资产影响企业高质量发展的调节作用,为充分发挥数据资产的积极效应提供了经验借鉴和策略启示。

二、理论分析与研究假说

1. 数据资产与企业高质量发展

近年来,作为数字经济时代的新型战略性资源,数据资产引起学界的高度关注。现有研究从数据资产的概念框架出发,一方面通过对数据资产的数据属性解读探求其产生经济收益的可能性及途径,另一方面基于数据资产的资产或财产属性探讨其权属问题(韩秀兰等,2023)^[26]。基于此,本文主要从数据资产的资产属性和数据属性两个方面,阐述其对企业高质量发展的影响。

从数据资产的资产属性来看:数据成为资产,首先表现为企业拥有的一种稀缺性资源。根据资源基础理论,稀有、不可替代和难以模仿的异质性资源是企业竞争优势的核心来源,对于促进企业业绩提升、推动企业高质量发展具有重要作用。企业对数据资源进行清洗和标准化处理,匹配数据源并删除冗余数据,使得数据资源成为产权明晰的资产并进入财务报表。由此,数据资产具有了可阅读、可理解和可增值的特征,企业可以利用其内含的有效信息进行商机洞察、风险管控和辅助决策等,从而降低信息搜索成本、传递成本和验证成本等;同时,企业还可利用数据资产开发新技术新产品新服务,形成新的收益增长点,促进企业价值提升。此外,资产化的数据资源具有可控制、可计量、可使用的特征,会在流通中进一步权益化,进而可以作为权益工具进行股权投资或融资,比如进入资本市场开展抵押、托管、信托及保险等金融业务,有效拓宽企业的融资渠道,降低企业的融资成本,有利于化解企业高质量发展的融资困境。因此,数据资产基于数据价值链,历经从资源化到资产化再到资本化的价值形态演进,为企业追求更高水平、高层次、高效率的价值创造提供了更好的条件和优势,进而赋能企业高质量发展。

从数据资产的数据属性来看:数据资产在本质上是一种信息知识资产,具有促进技术进步、提高生产效率、改善经营管理的能效。企业可以利用数据资产所蕴含的知识和信息开发新技术新产品、改进业务流程、提高经营管理效率,从而增强竞争优势;还可以通过数据资产来缓解企业与外部信息使用者的信息不对称,构建纵向贯通、横向互联的数据共享体系,在数据资产的流通与应用过程中与其他经济主体更好地实现优势互补。相比其他生产要素,数据资产具有虚拟替代性、无消耗性和共享性等特点,能

够源源不断地产生和循环使用,并促进传统生产要素的流动配置和优化组合,产生价值倍增效应和投入替代效应(罗玫等,2023)^[27]。一方面,数据资产与传统生产要素深度融合,比如金融科技(资本要素)、智能机器人(劳动力要素)、人工智能(技术要素)、数字孪生(自然资源)等,可以显著提升传统生产要素的生产效率,产生价值倍增效应;另一方面,数据资产可以对人力和管理等传统生产要素产生虚拟替代效应,用更少的物质资源创造出更多的物质财富。例如,利用大数据智能技术来模拟人脑思维进行分析与决策,从而实现对部分人力的替代。与此同时,数据商品也可与传统生产要素相结合,对企业的资产结构、劳动力结构、技术结构以及管理决策等进行优化,从而改善企业的生产要素配置,促进企业高质量发展。

基于上述分析,本文提出假说 H1:数据资产增加能够显著促进企业高质量发展。

2. 数据资产对传统生产要素的优化作用

不同于传统生产要素,数据资产具有无实物形态的特点,其更多的是以信息和知识来驱动企业发展。数据资产除了可以直接改善企业的生产经营状态外,还能够通过技术流、资金流、人才流等突破传统生产要素的稀缺性和排他性限制,改善企业的劳动力、资本、技术和管理等传统生产要素的配置状态,产生基础性资源优化效应,并带来全要素生产率的提升。因此,本文认为数据资产能够通过传统生产要素的优化机制来赋能企业高质量发展,并主要基于劳动力、资本、技术和管理四种传统生产要素展开分析。

(1) 劳动力结构改善机制

第一,数据资产具有劳动力筛选效应。数据本身具有复制成本较低、确权模糊和收益难以转化等特点,这使得数据资产的整合难度大,因而企业沉淀数据资产需要招揽高素质和高技能的人才来进行数据汇聚、处理、管理、应用及流通。同时,企业在利用数据资产进行生产运营模式的平台化、智能化、生态化转变过程中,高重复性、低技能的就业岗位会被知识技能型岗位替代,导致技术型员工比例提高。第二,数据资产具有劳动力素质提升效应。数据资产是企业通过数据清理、数据标准化、数据搜索匹配等形成的有价值的高质量数据,可以有效地转化为具有新的广度、深度和高度的内外部信息,有利于企业员工通过“干中学”及“再教育”等方式来提升自身的知识与技能,从而促使企业的人力资本水平整体提升。第三,数据资产具有人才吸引效应。数据资产是不可替代和难以复制的竞争性资源,企业拥有的数据资产越多,越能够向外界传递其具有独特优势和巨大发展潜力的正面信号,从而吸引具有丰富经验和专业知识的高素质人才进入。企业的高质量发展与其劳动力结构紧密相关,高层次和高质量的人才越多,学习能力、吸收能力、转化能力和扩散能力越强,越能够实现技术突破、转型升级和新旧动能转化。因此,数据资产可以通过劳动力结构改善机制提升企业的人力资本水平,进而促进企业高质量发展。

基于上述分析,本文提出假说 H2a:数据资产增加可以通过改善劳动力结构的路径来赋能企业高质量发展。

(2) 资本配置效率提升机制

第一,数据资产具有投资决策优化效应。数据资产的增加有助于企业投资决策的可视化,可以通过直观的方式呈现相关信息,便于管理者从海量的数据中快速获取有效的信息和知识,为投资决策提供更全面准确的数据支持,还能够通过“数据+算法”的科学决策模式对资本配置组合进行智能评估、模拟择优,从而提高资本配置的有效性。第二,数据资产具有投资效率提升效应。企业可以利用数据资产进行动态持续的数据收集、反馈和分析,实时监控资本配置过程中的资本流向、使用与回收等情况,并通过风

险和收益评估制定最优处置方案,及时调整资金流向,以最大化投资收益。第三,数据资产具有后期投资优化效应。企业能够利用数据资产所包含的本期资本配置信息进行事后评价,将资本配置实际表现与原有预期相比较,深入分析投资经验及问题所在,并据此调整后期的投资方向和结构,以进一步提高资本配置效率。资本错配会导致企业的投资和生产行为偏离最优决策,是造成企业全要素生产率损失的主要原因之一(Song et al., 2015;陈诗一等, 2017;才国伟等, 2019)^[28-30],而数据资产对资本配置的优化作用可以显著改善企业的投资结构、提高企业的投资效率,从而有效缓解资本错配对企业全要素生产率提升的制约作用。因此,数据资产能够通过资本配置效率提升机制提高企业的投资效率,进而促进企业高质量发展。

基于上述分析,本文提出假说 H2b:数据资产增加可以通过提升资本配置效率的路径来赋能企业高质量发展。

(3) 技术创新积累机制

第一,数据资产具有知识溢出和积累效应。数据资产将复杂多样的信息和知识进行标准化的数据诠释,有助于各种显性和隐性知识资源在企业各部门、各环节共享和传输,催生出网络化的协同创新,这不仅使得企业能够通过学习效应和积累效应获得更多异质性技术资源和技术创新思维,提升技术创新认知,还有助于企业与相关知识主体之间的技术交流合作,从而打破原有的技术组合,重塑并赋能企业的技术资源生成和积累。第二,数据资产具有技术创新促进效应。基于与产品研发、技术革新相关的数据资产,企业可以通过数据分析和挖掘对目标市场进行画像,准确掌握市场需求,从而有针对性地开展技术研发活动,提高技术创新的产出和绩效。第三,数据资产具有创新成本节约效应。在技术创新过程中,企业可以利用数据资产进行创新模拟,以数据试错替代实物试错,不但节约了技术创新成本,也提高了技术创新质量。技术进步是经济发展的内生动力,新技术的使用会推动企业转型升级,加快生产经营方式转变,建立市场竞争优势,进而促进企业高质量发展。因此,数据资产能够通过技术创新积累机制提高企业的技术水平,进而促进企业高质量发展。

基于上述分析,本文提出假说 H2c:数据资产增加可以通过促进技术创新和积累的路径来赋能企业高质量发展。

(4) 管理提质增效机制

第一,数据资产具有管理改善效应。资产化的数据提高了信息的有序性、确定性和清晰度,有助于企业利用数据信息平台实现数据信息的实时共享,使得生产经营的各环节由串联转向并联,推动信息结构从层次式转向网络式,进而打破各环节、各部门之间的信息屏障(陈冬梅等, 2020)^[31],减少不同部门和团队之间的信息失真和沟通阻滞,不仅可以提高生产效率,也能够优化管理流程和机制,提高企业管理的质量。第二,数据资产具有监管强化效应。数据资产以各类信息系统为载体,留存了企业日常生产经营的全过程信息,实现了事前、事中和事后监督全链条的信息穿透,并依托数据中台等放大和扩散监管范围,使得各行为主体的履责情况能够得到及时客观的反映,这将增大管理层藏匿信息的难度,从而有效缓解委托代理冲突,抑制高管自利行为,提高企业监管的有效性和内部管理的效率。企业管理质量和效率的提升,为企业更好地配置各种生产要素和改善生产经营提供了必要的内部治理保障,有利于企业的高质量发展。因此,数据资产能够通过管理提质增效机制提高企业的经营管理水平和效率,进而促进企业高质量发展。

基于上述分析,本文提出假说 H2d:数据资产增加可以通过提高经营管理效率的路径来赋能企业高质量发展。

三、企业数据资产评估与分析

1. 企业数据资产评估方法

已有文献对企业数据资产的评估,通常采用“企业当年是否具备共享平台、数据中台或信息系统”或“数据资产关键词是否出现在年度报告中”来体现企业数据资产的“有或无”(孙颖等,2021;Hu et al., 2022)^{[1][19]}。近期,部分学者基于企业财务报表指标对数据资产的“多与少”进行了测度(李健等, 2023;路征等,2023)^{[18][24]}。但由于数据资产目前还未被纳入资产负债表中进行披露,基于财务报表对企业数据资产的评估难免存在遗漏。数据资产是企业利用大数据、人工智能等信息技术进行数智化转型形成的重要经济资源,此类特征信息会体现在具有总结和指引性质的企业年报中,因此,本文基于机器学习算法,采用“种子词筛选+AutoPhrase 相似词扩充+SnowNLP 情感极性分析”的文本分析方法对企业的数字资产水平进行评估。具体步骤如下:

第一,构建底层逻辑并选取种子词。数据从非结构化的低价值零散形态向结构化的可被应用的高价值资产形态转化,形成“数据价值链”(许宪春等,2022)^[32]。基于此,本文以“数据资产获取—数据资产处理—数据资产应用”价值链为底层逻辑,结合数据资产相关文件和研究报告,从企业年报中手工筛选出分别代表“数据资产获取”“数据资产处理”“数据资产应用”三个层面的种子词,然后通过访谈及调查问卷的形式向学术界、大数据行业协会以及相关企业人员征询意见,对种子词进行调整和完善。

第二,扩展种子词形成数据资产词典。刘景江等(2023)^[33]指出,机器学习算法拥有强大的数据处理和挖掘能力,能够从复杂高维的非结构化数据环境中提取有价值的信息。因此,本文借鉴相关文献对机器学习算法的应用(胡楠等,2021;孙昌玲等,2021)^[34-35],采用 AutoPhrase 方法对种子词进行扩展,以挖掘企业年报中能够体现数据资产的高质量特征词。同时,进一步对特征词谱进行检验与补充:首先,2023年中国信息通信研究院发布的《数据资产管理实践白皮书(6.0版本)》指出,数据资产是以电子或其他方式(如文本、图像、语音、视频、网页、数据库、传感信号等)记录的结构化或非结构化数据。基于此,对与各数据记录方式相关的关键词是否纳入本文特征词谱进行检验与补充。其次,根据国务院印发的《“十四五”数字经济发展规划》、工业和信息化部与国家互联网信息办公室等十六部门联合印发的《关于促进数据安全产业发展的指导意见》、工业和信息化部发布的《“十四五”大数据产业发展规划》等官方文件,将其中提到的数据资产关键词与本文的特征词谱进行对比和扩展,以确保词谱的全面性。最终,本文构建了如图1所示的数据资产特征词谱。

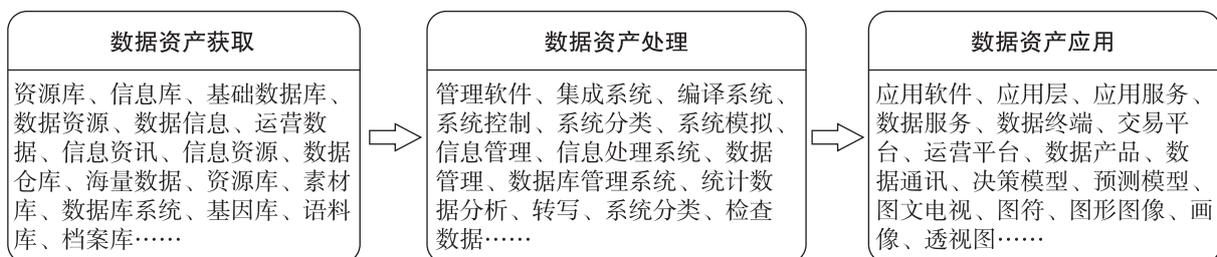


图1 数据资产特征词谱

第三,计算词频和情感极性分析。考虑到中文语言在财经语料下的表达较为复杂,根据数据资产词典,采用 Python 技术计算企业年报中各特征词的词频,并且利用 SnowNLP 技术对每个出现的词汇进行

情感极性分析,最终得到经情感极性(正向、中立、负向)调整后的各特征词词频。

第四,估算企业数据资产水平。将经情感极性调整后的正向及中立的各特征词词频进行加总,除以年报总词频再乘以 100,得到反映企业数据资产水平的指标。计算公式为: $DA_{i,t} = \sum Dwords_{i,t,n} / Tword_{i,t} \times 100$ 。其中, $DA_{i,t}$ 为 i 企业 t 年的数据资产水平, $Dwords_{i,t,n}$ 为词典中第 n 个特征词在 i 企业 t 年年度报告中的词频, $Tword_{i,t}$ 为 i 企业 t 年年度报告的总词频(排除英文和数字)。 DA 值越大,则企业的数据资产水平越高。

2. 数据资产评估指标的有效性检验

本文采用上述方法对样本企业(沪深 A 股的上市公司,详见后文)2007—2022 年的数据资产水平进行了评估,为了确保评估结果能够客观反映企业的数据资产水平,进行如下指标有效性检验。

(1) 指标对比检验

本文基于企业、行业和地区层面的相关资料和数据对企业数据资产水平评估结果进行对比分析。

一是企业层面的对比。由于数据资产具有非实体性和依附性(需依附于实物载体或介质而存在),是否拥有数据中台和信息管理系统等对于企业的数据资产化进程具有重要意义(孙颖等,2021)^[1],因此可认为,企业的数据中台发展状况能够在一定程度上体现其数据资产水平。借鉴胡楠等(2021)^[34]的检验思路,将《互联网周刊》、德本咨询、eNet 研究院联合发布《2022 数据中台 TOP50》榜单企业与本文的样本企业进行匹配,得到 5 家企业(用友软件、浪潮软件、星环科技、恒生电子、普元信息);然后根据各样本企业 2007—2022 年数据资产水平的均值将其分为最高、较高、较低和最低四组,结果发现这 5 家企业中有 4 家位于最高组,另 1 家位于较高组,表明本文测算的数据资产水平是有效的。此外,抽取数据资产特征词词频占比前五的企业(高鸿股份、威星智能、机器人、科大智能、东方国信),仔细研读其年报和投资者关系活动记录表等相关资料,发现相较于大部分企业,其确实拥有较多的各类(运营类、市场和客户类、工业类及知识产权类等)数据资产,这也在一定程度上表明本文评估得到的数据资产水平较高的企业确实拥有较多的数据资产。

二是行业层面的对比。计算出 2007—2022 年各行业企业数据资产水平的均值(见表 1),结果显示“信息传输、软件和信息技术服务业”“文化、体育和娱乐业”“综合类”等行业的企业数据资产水平较高,而“农、林、牧、渔业”和“采矿业”等行业的企业数据资产水平较低,这一数据资产的行业分布状态与已有相关研究结果一致(李健等,2023)^[18]。

三是地区层面的对比。计算出 2007—2022 年 31 个省份(不包括港澳台地区)企业数据资产水平的均值(见表 2)。贵州省的企业数据资产水平最高,这得益于 2015 年贵州大数据交易所的建立,贵州省的企业率先探索数据流动体系,并通过数据交易沉淀了较多数据资产。企业数据资产水平较高的省份还有北京、广东、上海和湖北,这也与《中国数字经济发展白皮书》中数字经济规模突破万亿元的省份相契合。可见,本文测算的企业数据资产水平在区域分布上也与实际情况基本一致。

(2) 指标相关性检验

根据财政部发布的《企业数据资源相关会计处理暂行规定》,企业使用的数据资源,符合规定的应确认为无形资产。在此之前,已有研究也认为数据资产应纳入企业无形资产的范畴。基于此,本文根据样本企业的财务报表,将附注中含有与数据资产化过程相关词汇的无形资产占总资产的比例,作为衡量企业“数据类无形资产”的代理变量。例如无形资产附注中出现大数据、数字平台、数据中心和数据共享等词汇,即将其视为数据类无形资产。借鉴胡楠等(2021)^[34]的验证思路,以样本企业的“数据类无形资

产”($DA_IA_{i,t}$) 为被解释变量,以“数据资产水平”($DA_{i,t}$) 为核心解释变量,并纳入一系列控制变量($Controls$,控制变量与后文的基准模型一致)以及个体($Company$)、行业(Ind)、年份($Year$)固定效应,构建如下相关性检验模型: $DA_IA_{i,t} = \alpha_0 + \alpha_1 DA_{i,t} + \sum \alpha_k Controls_{i,t} + \sum Company + \sum Ind + \sum Year + \varepsilon_{i,t}$ 。回归结果显示,“数据资产水平”的系数在 1%的水平上显著为正(0.078 1),表明本文测算的企业数据资产水平与企业数据类无形资产水平显著正相关,因此可以认为该指标是有效的。

表 1 2007—2022 年各行业企业数据资产水平均值

行业	数据资产水平	行业	数据资产水平
信息传输、软件和信息技术服务业	0.524 4	住宿和餐饮业	0.077 1
文化、体育和娱乐业	0.204 0	交通运输仓储和邮政业	0.073 8
综合类	0.168 1	建筑业	0.062 7
租赁和商务服务业	0.127 8	水利、环境和公共设施管理业	0.057 1
制造业	0.113 0	电力、热力、燃气及水的生产和供应业	0.044 0
卫生和社会工作	0.105 6	房地产业	0.046 5
教育	0.107 5	农、林、牧、渔业	0.042 0
批发和零售业	0.091 9	采矿业	0.037 9
居民服务、修理和其他服务业	0.089 3		

表 2 2007—2022 年各省份企业数据资产水平均值

省份	数据资产均值	省份	数据资产均值	省份	数据资产均值	省份	数据资产均值
贵州	0.207 7	江苏	0.120 4	吉林	0.085 0	云南	0.058 1
北京	0.202 7	湖南	0.112 3	陕西	0.083 4	新疆	0.056 1
广东	0.157 4	四川	0.106 6	河北	0.073 6	青海	0.053 1
上海	0.137 7	河南	0.105 0	重庆	0.071 4	宁夏	0.051 0
湖北	0.137 6	山东	0.092 4	广西	0.069 7	甘肃	0.048 6
浙江	0.136 6	江西	0.091 6	黑龙江	0.068 9	内蒙古	0.043 3
福建	0.125 7	安徽	0.090 8	海南	0.063 8	西藏	0.042 1
辽宁	0.124 4	天津	0.090 7	山西	0.062 8		

四、实证研究设计

1. 基准模型构建与变量选择

为检验数据资产对企业高质量发展的影响,构建如下基准模型:

$$TFP_LP_{i,t} = \alpha_0 + \alpha_1 DA_{i,t} + \sum \alpha_k Controls_{i,t} + \sum Company + \sum Year + \varepsilon_{i,t}$$

其中,下标 i 和 t 为分别代表企业和年份,被解释变量($TFP_LP_{i,t}$)“高质量发展”为企业 i 在 t 年的高质量发展水平,核心解释变量($DA_{i,t}$)“数据资产”为企业 i 在 t 年的数据资产水平; $Controls_{i,t}$ 表示一系列控制变量, $Year$ 和 $Company$ 分别表示年份和个体固定效应, $\varepsilon_{i,t}$ 为随机误差项。

被解释变量“高质量发展”采用基于 LP 法测算的全要素生产率来衡量。全要素生产率能体现企业各生产要素的使用状况和配置效率,是企业发展质量和发展数量的综合反映,具有信息丰富和综合性强等优点,能够反映企业高质量发展水平,且与本文探究数据资产驱动传统生产要素改善的研究重点相契合。因此,本文选用全要素生产率作为企业高质量发展的代理变量,并借鉴鲁晓东和连玉君(2012)^[36]的做法,采用 LP 法测算样本企业的全要素生产率。核心解释变量“数据资产”采用前文构建的企业数据资产评估方法进行测算。

参考黄勃等(2023)^[11]和黄先海(2023)^[37]的研究,从企业层面选取以下控制变量:一是“资产规模”,采用总资产的自然对数值来衡量;二是“企业年龄”,采用当年年份减企业成立年份加 1 的自然对数值来衡量;三是“成长性”,采用营业收入的增长率来衡量;四是“资产负债率”,采用期末总负债与期末总资产之比来衡量;五是“现金流比率”,采用经营活动产生的现金流量净额与营业收入之比来衡量;六是“市场价值”,采用总市值与总资产之比来衡量;七是“研发能力”,采用专利申请数量加 1 的自然对数值来衡量;八是“第一大股东持股比例”,采用年末第一大股东持股比例来衡量;九是“董事会规模”,采用董事会人数的自然对数值来衡量;十是“产权性质”,为是否国有企业的虚拟变量,国有企业赋值为 1,否则赋值为 0。

2. 样本选择与数据处理

本文以沪深 A 股的上市公司为研究样本,考虑到在 2006 年以后数字化工具开始广泛使用(李琦等,2021)^[38],样本期间选定为 2007—2022 年,剔除金融类行业样本、ST 和 *ST 类样本、数据缺失样本,最终获得 36 589 个观测值,并对连续型变量进行前后 1%的缩尾处理。企业数据资产的相关数据来自深圳证券交易所和上海证券交易所官方网站披露的上市公司年报,其他数据源自 CSMAR 数据库。表 3 为主要变量的描述性统计结果。从整体样本来看,“高质量发展”的标准差为 1.043,最大值为 11.11,最小值为 6.032,表明不同企业间的发展质量存在较大差异;“数据资产”均值为 0.134,最大值为 1.227,最小值为 0,说明样本企业的数据资产水平总体较低,有待提高。

表 3 主要变量的描述性统计结果

变量	样本量	最小值	最大值	均值	中位数	标准差
高质量发展	36 589	6.032	11.11	8.279	8.179	1.043
数据资产	36 589	0	1.227	0.134	0.0574	0.209
资产规模	36 589	19.32	26.45	22.19	22.00	1.286
企业年龄	36 589	0.693	3.611	2.878	2.944	0.356
成长性	36 589	-0.658	4.024	0.171	0.108	0.413
资产负债率	36 589	0.0274	0.908	0.434	0.429	0.203
现金流比率	36 589	-0.223	0.283	0.048	0.047	0.070
市场价值	36 589	0.064	1.246	0.618	0.614	0.248
研发能力	36 589	0	9.571	1.721	1.609	1.686
第一大股东持股比例	36 589	8.020	75.84	34.48	32.27	14.72
董事会规模	36 589	1.609	2.708	2.130	2.197	0.200
产权性质	36 589	0	1	0.389	0	0.488

五、实证检验结果分析

1. 基准模型回归

基准模型检验结果见表4,“数据资产”的回归系数在1%的水平上显著为正,表明企业数据资产的增加对其高质量发展具有显著的正向影响,假说H1得到验证。考虑到本文从数据资产获取、数据资产处理和数据资产应用3个层面确定企业的数据资产的特征词词谱,进一步分别统计该3个层面的特征词词频数,得到相应的指标(“数据资产获取”“数据资产处理”“数据资产应用”),用以反映不同层面的数据资产水平,并分别以其为核心解释变量进行模型检验,回归结果见表5。“数据资产获取”的回归系数为正但不显著,“数据资产处理”和“数据资产应用”的回归系数在1%的水平上显著为正,表明在企业的数字价值链中,数据资产的处理和应用显著促进了企业高质量发展,而数据资产的获取对企业高质量发展的影响不显著。其原因在于,处于获取阶段的数据往往是单一、零散的,这种尚未聚合的数据价值有限,往往需要经过处理后才具有应用价值;而在进入数据价值链的处理阶段后,大量无序的原始数据经过汇集、规整和分析,呈现出可阅读、可理解和可增值的状态,可以从中提炼出有效的信息与知识,并应用于企业的生产经营全过程中,进而为企业带来多元化的效益,助力企业高质量发展。

表4 基准模型回归结果

变 量	高质量发展	高质量发展	高质量发展	高质量发展
数据资产	1.237 9*** (20.202 8)	0.477 8*** (8.810 7)	0.197 4*** (4.891 5)	0.149 2*** (3.563 2)
资产规模			0.531 3*** (41.774 7)	0.564 6*** (43.199 5)
企业年龄			0.148 7*** (5.219 4)	0.141 3** (2.370 6)
成长性			0.224 1*** (30.370 2)	0.219 7*** (30.253 4)
资产负债率			0.135 0*** (3.051 8)	0.100 3** (2.285 9)
现金流比率			0.787 8*** (15.973 1)	0.745 5*** (15.176 5)
市场价值			-0.048 0*** (-2.643 6)	-0.282 8*** (-11.015 0)
研发能力			-0.004 7 (-1.244 9)	-0.003 2 (-0.847 3)
第一大股东持股比例			0.000 6 (0.743 7)	0.001 0 (1.326 3)
董事会规模			0.039 0 (1.120 9)	0.020 7 (0.597 9)

续表 4

变 量	高质量发展	高质量发展	高质量发展	高质量发展
产权性质			-0.004 9 (-0.186 3)	-0.013 3 (-0.508 9)
常数项	8.113 1*** (988.179 7)	7.599 6*** (377.778 3)	-4.163 5*** (-17.893 1)	-4.704 2*** (-16.406 9)
年份固定效应	未控制	未控制	控制	控制
个体固定效应	控制	控制	控制	控制
观测值数量	36 589	36 589	36 589	36 589
调整 R ²	0.0607	0.3049	0.5780	0.5916

注: *、**、***分别表示在 10%、5%、1%的水平上显著,括号内为 t 值(企业层面聚类处理),下表同。

表 5 不同层面数据资产水平对企业高质量发展的影响

变 量	高质量发展	高质量发展	高质量发展
数据资产获取	0.078 4 (1.091 8)		
数据资产处理		0.367 7*** (4.752 2)	
数据资产应用			0.198 6*** (4.254 0)
控制变量	控制	控制	控制
个体和年份固定效应	控制	控制	控制
观测值数量	36 589	36 589	36 589
调整 R ²	0.591 0	0.591 3	0.591 5

2. 内生性处理

为了缓解模型可能存在的样本自选择偏误、逆向因果关系以及遗漏变量等内生性问题对回归结果的干扰,本文进行如下内生性处理:

(1) 倾向得分匹配法

数据资产水平较高的企业可能本身的发展质量就较高,并非一定是由于数据资产的增加促进了其高质量发展。因此,根据企业数据资产水平的中位数将样本划分为“高水平”(实验组)与“低水平”(对照组)两组,选取“资产规模”“企业年龄”“成长性”“资产负债率”作为协变量,进行 logit 回归得到倾向得分值,并借助协变量进行 1:1 最近邻匹配。样本匹配后,变量的标准化偏差都小于 2%,且实验组与对照组无系统差异,通过了平衡性检验(见表 6),ATT 值为 9.2,与前文分析结果一致。删除没有匹配上的样本后重新进行模型检验,回归结果见表 7 的 Panel A,“数据资产”的回归系数依然在 1%的水平上显著为正。

表6 匹配变量的平衡性检验

变量	未匹配/匹配	处理组	控制组	%偏差	T值	P值
公司规模	匹配前	22.180	22.206	-2.0	-1.94	0.052
	匹配后	22.179	22.167	0.9	-0.91	0.363
成立年限	匹配前	2.8937	2.8632	8.6	8.20	0.000
	匹配后	2.8936	2.8896	1.1	1.08	0.279
成长性	匹配前	0.1772	0.1643	3.1	3.00	0.003
	匹配后	0.1761	0.1771	-0.4	-0.42	0.672
资产负债率	匹配前	0.4159	0.4519	-17.8	-17.03	0.000
	匹配后	0.4160	0.4146	0.7	0.68	0.498

(2) 工具变量法

企业的发展质量也可能会影响其数据资产水平,产生反向因果关系。对此,本文将剔除样本企业自身的分年度、行业的数据资产水平均值作为“数据资产”的工具变量,进行2SLS检验。由于数据资产在企业间存在同群效应,样本企业的数据资产水平与同行业其他企业的数据资产水平具有相关性,同时行业内其他企业的平均数据资产水平不会对样本企业的高质量发展产生直接作用,因而该工具变量满足工具变量有效的相关性和外生性要求。工具变量法检验结果见表7的Panel B,工具变量通过了不可识别检验(Kleibergen-Paap rk LM统计量显著)与弱工具变量检验(Kleibergen-Paap Wald rk F统计量显著)。第一阶段回归结果显示,工具变量与“数据资产”显著正相关;第二阶段回归结果显示,工具变量拟合的“数据资产”的回归系数仍然在1%的水平上显著为正。

(3) 更换固定效应

尽管基准模型已经控制了个体固定效应和年份固定效应,但仍可能遗漏不同年份发生在特定地区或特定行业的冲击,因此进一步增加省份固定效应和行业固定效应重新进行模型检验。此外,为了缓解大量干扰性因素的影响,将基准模型的固定效应更换为“省份—行业”“年份—行业”“年份—省份”三种高维交互固定效应中的两种,分别进行模型检验。更换固定效应的回归结果见表7的Panel C,“数据资产”的回归系数均在1%的水平上显著为正。

上述内生性处理结果表明,在缓解模型内生性问题后,本文的基本结论依然成立,即企业数据资产的增加促进了企业高质量发展。

表7 内生性处理结果

变量	Panel A:	Panel B:工具变量法		Panel C:更换固定效应			
	PSM 检验	第一阶段	第二阶段	高质量发展	高质量发展	高质量发展	高质量发展
	高质量发展	数据资产	高质量发展	高质量发展	高质量发展	高质量发展	高质量发展
数据资产	0.1319*** (2.7936)		1.1123*** (12.6349)	0.1105*** (2.7990)	0.1151*** (2.7848)	0.1196*** (3.1900)	0.1232*** (3.1978)
工具变量		0.5163*** (11.4386)					
控制变量	控制	控制	控制	控制	控制	控制	控制
年份固定效应	控制	控制	控制	控制	未控制	未控制	未控制

续表 7

变 量	Panel A:	Panel B:工具变量法		Panel C:更换固定效应			
	PSM 检验	第一阶段	第二阶段	高质量发展	高质量发展	高质量发展	高质量发展
	高质量发展	数据资产	高质量发展	高质量发展	高质量发展	高质量发展	高质量发展
个体固定效应	控制	控制	控制	控制	未控制	未控制	未控制
行业固定效应				控制	未控制	未控制	未控制
省份固定效应				控制	未控制	未控制	未控制
省份—行业固定效应				未控制	未控制	控制	控制
年份—行业固定效应				未控制	控制	未控制	控制
年份—省份固定效应				未控制	控制	控制	未控制
观测值数量	19 824	36 578	36 578	35 508	35 496	35 491	35 480
调整 R ²	0.593 6	0.213 6	0.521 4	0.611 0	0.710 5	0.732 3	0.736 9
K-P rk LM 统计量		305.208***					
K-P Wald rk F 统计量		594.250***					

3. 稳健性检验

为进一步验证基准模型分析结果的可靠性,进行以下稳健性检验:

(1) 变量替换

一是替换核心解释变量。在数据资产词典中,与中立情感词汇相比,正向情感词汇更能体现企业对数据资产的认可及取得的成就,因此,仅加总经情感分析调整后的正向数据资产特征词词频,从而得到“数据资产 1”,将其作为核心解释变量重新进行模型检验。此外,将数据类无形资产占比(“数据资产 2”)作为核心解释变量,重新进行模型检验。二是替换被解释变量。分别采用 OP 法、OLS 法和 FE 法测算样本企业的全要素生产率,将其作为被解释变量重新进行模型检验。上述检验结果见表 8,核心解释变量的回归系数均在 1%的水平上显著为正,表明本文的分析结果是稳健的。

表 8 替换变量的稳健性检验结果

变 量	替换核心解释变量		替换被解释变量		
	高质量发展	高质量发展	OP 法	OLS 法	FE 法
数据资产			0.083 7** (1.979 7)	0.097 4*** (2.785 2)	0.093 2*** (2.713 4)
数据资产 1	0.117 4*** (3.181 8)				
数据资产 2		0.162 8*** (4.057 8)			
控制变量	控制	控制	控制	控制	控制
个体和年份固定效应	控制	控制	控制	控制	控制
观测值数量	36 589	28 083	36 589	36 589	36 589
调整 R ²	0.591 4	0.596 6	0.498 0	0.725 3	0.744 8

(2) 样本优化

一是剔除直辖市样本。考虑到直辖市的特殊性,将企业所在城市为直辖市的样本剔除后,重新进行检验。二是剔除部分省份样本。考虑到企业的信息披露可能存在跟风现象(比如政府倡导时可能披露更多相关的内容),根据卜令通和嘉伟(2023)^[39]的研究结论,将浙江、上海、江苏、北京、甘肃和宁夏的样本剔除后,重新进行检验。三是剔除2013年之前样本。根据人民网的报道,2013年被称为我国的“大数据元年”,据此,采用2013—2022年的样本重新进行检验。四是剔除数据资产为0的样本。由于存在数据资产为0的样本,将其剔除后重新进行检验。五是剔除信息披露质量较差样本。考虑到企业的信息披露可能存在投机性和策略性动机,导致信息披露存在偏误,借鉴袁淳等(2021)、戴翔和马皓巍(2023)^[40-41]的方法,根据证券交易所对企业信息披露质量的评级,只选取评级结果为优秀和良好的企业样本进行检验。上述检验结果见表9,“数据资产”的回归系数依然均在1%的水平上显著为正,进一步表明本文基准模型的分析结果具有较好的稳健性。

表9 优化样本的稳健性检验结果

变 量	剔除直辖市样本	剔除部分省份样本	2013—2022年样本	剔除数据资产为0样本	剔除信息披露质量较差样本
数据资产	0.175 1*** (3.463 0)	0.185 6*** (2.849 0)	0.115 8*** (3.199 7)	0.148 4*** (3.531 8)	0.130 2*** (3.039 8)
控制变量	控制	控制	控制	控制	控制
个体和年份固定效应	控制	控制	控制	控制	控制
观测值数量	29 009	21 295	28 288	36 078	27 387
调整 R ²	0.611 8	0.611 7	0.529 1	0.593 4	0.600 9

4. 机制检验

为检验数据资产能否通过优化传统生产要素的机制来促进企业高质量发展,依照江艇(2022)^[42]提出的中介效应检验方法,构建如下计量模型:

$$M_{i,t} = \beta_0 + \beta_1 DA_{i,t} + Controls_{i,t} + \gamma_i + \lambda_t + \varepsilon_{i,t}$$

其中, M 为中介变量。根据前文理论分析,选取以下4个中介变量:一是“劳动力结构”,采用本科及以上学历员工人数与企业员工总数之比来衡量,用以检验数据资产的劳动力结构改善机制;二是“非效率投资”,参考刘姝雯等(2023)^[43]的方法,采用样本企业的非效率投资作为资本配置效率的代理变量(其值越小,企业的资本配置效率越高),用以检验数据资产的资本配置效率提升机制;三是“技术水平”,借鉴黄先海和高亚兴(2023)^[37]的做法,采用样本企业专利被引数量加1的自然对数值来衡量(其值越大,企业的技术水平越高),用以检验数据资产的技术创新积累机制;四是“管理费用率”,采用管理费用与营业收入之比作为经营管理效率的代理变量(其值越小,企业的经营管理效率越高),用以检验数据资产的管理提质增效机制。

机制检验结果见表10。“数据资产”对“劳动力结构”和“技术水平”的回归系数显著为正,表明数据资产的增加改善了企业的劳动力结构、提高了企业的技术水平;“数据资产”对“非效率投资”和“管理费用率”的回归系数显著为负,表明数据资产的增加抑制了企业的非效率投资、降低了企业的管理成本。

可见,企业的数据资产增加,能够通过劳动力结构改善机制提高企业的人力资本水平、通过资本配置效率提升机制提高企业的投资效率、通过技术创新积累机制提高企业的技术水平、通过管理提质增效机制提高企业的管理效率,从而促进企业高质量发展。由此,本文提出的假说 H2a、H2b、H2c、H2d 均得到验证。

表 10 机制检验结果

变 量	劳动力结构	非效率投资	技术水平	管理费用率
数据资产	0.104 2** (2.483 1)	-0.021 4** (-2.105 4)	0.508 3*** (5.708 7)	-0.017 0** (-2.110 8)
控制变量	控制	控制	控制	控制
年份固定效应	控制	控制	控制	控制
个体固定效应	控制	控制	控制	控制
观测值数量	29 319	31 835	26 472	36 589
调整 R ²	0.187 4	0.032 1	0.547 6	0.064 3

六、进一步分析:外部环境的调节作用

企业的发展以及数据资产的积累会受到外部环境的影响,在不同的发展环境中,数据资产对企业高质量发展的影响程度可能存在明显差异。比如,良好的市场和政策环境可以引导生产要素有序流动和合理配置,从而使数据资产对企业高质量发展的赋能作用得以更充分地发挥。有鉴于此,本文进一步从环境不确定性、城市信息化水平、地区市场化水平以及数字经济政策等方面来考察外部环境对数据资产影响企业高质量发展的调节作用。在基准模型中加入调节变量(Z)及其与“数据资产”的交互项($DA \times Z$),构建如下计量模型:

$$TFP_LP_{i,t} = \delta_0 + \delta_1 DA_{i,t} + \delta_2 DA_{i,t} \times Z_{i,t} + Z_{i,t} + Controls_{i,t} + \gamma_i + \lambda_t + \varepsilon_{i,t}$$

第一,环境不确定性的调节效应。在现实经济中,企业的发展会面临众多的不确定性因素。当环境不确定性较强时,市场机会和风险并存,且复杂多变的各类信息导致难以预测未来,企业科学决策的难度加大;此时,数据资产将有助于企业聚合和分析庞杂的数据信息,并挖掘和利用隐性知识,从而提高企业对机会和风险的识别能力以及决策的有效性和行动的及时性,助力企业高质量发展。而当环境不确定性较弱时,市场较为稳定,预测发展趋势相对容易,此时数据资产对企业决策的改进作用相对较小。因此,数据资产对企业高质量发展的促进作用会随着环境不确定性的提高而增强。参考 Ghosh 和 Olsen (2009)^[44]、申慧慧等(2012)^[45]的研究,采用经过行业调整的过去五年销售收入变异系数来衡量企业的“环境不确定性”,以其为调节变量的检验结果见表 11 的(1)列。“数据资产×环境不确定性”的回归系数显著为正,表明环境不确定性的提高会强化数据资产对企业高质量发展的促进作用。

第二,城市信息化水平的调节效应。当企业所在城市的信息化水平较高时,信息和知识传播的阻碍较小,有利于数据资产的交流、扩散和共享,此时,数据资产的共享性、无消耗性和报酬递增性等优势可以得到较为充分的发挥,能够更有效地赋能企业的生产运营各环节,进而推动企业全要素生产率增长。借鉴张桅和胡艳(2020)^[46]的研究,采用城市固定电话用户数、移动电话用户数、互联网用户数之和与城市土地面积的比值来衡量“城市信息化水平”,以其为调节变量的检验结果见表 11 的(2)列。“数据资产×城市信息化水平”的回归系数显著为正,表明城市信息化水平的提高能够强化数据资产对企业高质量发展的促进作用。

第三,区域市场化水平的调节效应。当企业所在区域的市场化水平较高时,一方面,资源配置的市场化程度较高,生产要素的流动性较强,而生产要素的自由流动则会增强数据资产的要素优化和驱动作用;另一方面,市场体系和制度也较为健全,可以为数据资产的流通和交易提供良好的外部环境 with 制度保障。因此,区域市场化水平的提高可以通过市场发育和制度健全的双重作用,促使数据资产更有效地赋能生产要素配置改善,从而增强数据资产对企业高质量发展的促进作用。借鉴王小鲁等(2019)^[47]的研究,使用市场化指数来衡量各省市的“区域市场化水平”,以其为调节变量的检验结果见表 11 的(3)列。“数据资产×区域市场化水平”的回归系数显著为正,表明区域市场化水平的提高可以强化数据资产对企业高质量发展的促进作用。

第四,数字经济政策的调节效应。当前,各地政府均十分重视数字经济的发展,并采用各种政策工具激励和帮助企业推进数字化转型。促进数字经济发展的相关政策实施,不仅为企业的数字化转型和数据资产积累提供了直接的支持,还为企业的数字化发展以及数据资产的流通和应用提供了更好的外部环境,从而促使数据资产对企业高质量发展的促进作用得以更充分地发挥。考虑到政府工作报告能够反映政府对各领域的关注度及政策支持力度,借鉴金灿阳等(2022)^[48]、陶长琪和丁煜(2022)^[49]的做法,采用省级政府工作报告中数字经济特征词词频作为数字经济政策供给水平的代理变量“数字经济政策”,以其为调节变量的检验结果见表 11 的(4)列。“数据资产×数字经济政策”的回归系数显著为正,表明政府提高数字经济政策供给水平能够强化数据资产对企业高质量发展的促进作用。

表 11 调节效应检验结果

变 量	(1)	(2)	(3)	(4)
调节变量	环境不确定性	城市信息化水平	地区市场化水平	数字经济政策
数据资产	0.106 9** (2.398 6)	0.087 3** (1.974 9)	0.109 2** (2.489 5)	0.153 4** (2.450 6)
数据资产×调节变量	0.035 5** (2.210 5)	0.468 1** (2.296 2)	0.518 4** (2.493 8)	0.003 9* (1.939 3)
调节变量	-0.033 9*** (-6.748 2)	0.032 3 (0.479 3)	0.039 2 (0.587 5)	0.000 4 (0.685 4)
控制变量	控制	控制	控制	控制
个体和年份固定效应	控制	控制	控制	控制
观测值数量	35 151	34 066	34 066	34 066
调整 R ²	0.604 4	0.582 5	0.541 6	0.597 5

七、结论与启示

在数字经济时代,数据成为一种新型生产要素,数据资产则成为企业的新型战略性基础资源。数据资产的资产属性和数据属性决定了其不仅可以直接改善企业的生产经营状态,而且能够改善企业的劳动力、资本、技术和管理等传统生产要素的配置状态,产生基础性资源(生产要素)优化效应,进而推动企业高质量发展。本文以沪深 A 股上市公司为研究样本,根据企业年报,运用机器学习算法,通过“种子词筛选+AutoPhrase 相似词扩充+SnowNLP 情感极性分析”对样本企业的数据资产水平进行评估,并以全要素生产率衡量企业高质量发展水平,进而采用 2007—2022 年的数据进行实证检验,结果表明:(1)数据

资产增加显著促进了企业高质量发展,该结论在经过一系列内生性处理和稳健性检验后依然成立;此外,从数据价值链来看,数据资产对企业高质量发展的赋能主要发生于数据资产的处理和应用层面,数据资产的获取对企业高质量发展的影响不显著。(2)数据资产增加改善了企业的劳动力结构、抑制了企业的非效率投资、提高了企业的技术水平、降低了企业的管理成本,表明数据资产能够通过劳动力结构改善机制提高企业的人力资本水平、通过资本配置效率提升机制提高企业的投资效率、通过技术创新积累机制提高企业的技术水平、通过管理提质增效机制提高企业的管理效率,从而进一步赋能企业高质量发展。(3)环境不确定性、城市信息化水平、区域市场化水平和数字经济政策对数据资产赋能企业高质量发展具有正向调节作用,即环境不确定性、信息化水平、市场化水平和数字经济政策供给水平的提高会强化数据资产对企业高质量发展的促进作用。

基于以上结论,得出以下启示:第一,企业应积极开发、积累和利用数据资产,并充分发挥数据资产对传统生产要素的优化作用。一方面,各企业要调整改进数据资产开发使用的路径和模式,通过业务数据化和数据业务化赋能人才升级、技术变革、资本优化和管理提效;另一方面,各企业要不断完善与其他经济主体的数据资产交易、流通、共享机制,提升数据资产的要素优化和驱动作用。第二,政府和相关部门应针对数据资产在确认、核算、披露和交易等环节存在的问题,完善数据资产的统计与信息发布工作,积极培育数据要素市场,并加强引导和监管,为充分激发企业数据资产的潜在价值提供良好的市场环境和制度支持。第三,各地区应持续提高信息化和市场化水平,并大力支持数字经济发展。

此外,针对本文研究存在的不足,提出进一步研究的两点建议:(1)本文采用机器学习和文本分析方法对企业的数据资产水平进行度量,只能大致反映数据资产的相对水平。目前,对企业数据资产的核算和评估仍处于探索阶段,数据资产入表还存在收入成本匹配难、摊销期确定难等障碍。在企业的数据资产完成入表后,可利用企业的财务报表对数据资产进行精确测算,以提高实证分析结果的准确性。(2)受分析方法的限制,本文只验证了数据资产对于4种传统生产要素的优化作用,未能从数据资产与其他生产要素的协同作用角度展开分析。今后可采用NCA与fsQCA结合等研究方法,寻求数据资产与其他要素资源协同驱动企业高质量发展的有效路径,以进一步丰富企业高质量发展研究的经验分析工具。

参考文献:

- [1] 孙颖,陈思霞.数据资产与科技服务企业高质量发展——基于“宽带中国”准自然实验的研究[J].武汉大学学报(哲学社会科学版),2021,74(5):132-147.
- [2] 阳镇.数字经济如何驱动企业高质量发展?——核心机制、模式选择与推进路径[J].上海财经大学学报,2023,25(3):92-107.
- [3] 周泽将,雷玲,伞子瑶.营商环境与企业高质量发展——基于公司治理视角的机制分析[J].财政研究,2022(5):111-129.
- [4] 郑飞,李腾,刘晗.政府补贴对企业高质量发展的影响研究[J].经济经纬,2022,39(5):140-150.
- [5] 陈丽珊,傅元海.融资约束条件下技术创新影响企业高质量发展的动态特征[J].中国软科学,2019,(12):108-128.
- [6] 李雄飞.董事会多元化对国有上市企业高质量发展的影响研究[J].经济问题,2022(6):85-93.
- [7] 宋清华,林永康.杠杆率会影响全要素生产率吗——基于企业和地区异质性的视角[J].山西财经大学学报,2021,43(3):112-126.
- [8] 刘艳霞.数字经济赋能企业高质量发展——基于企业全要素生产率的经验证据[J].改革,2022(9):35-53.
- [9] 张超,钟昌标,杨佳妮.数字金融对实体企业高质量发展的影响研究——基于浙江的实证[J].华东经济管理,2022,36(3):63-71.

- [10] 李朝鲜.“双循环”背景下数字技术如何赋能商贸流通企业高质量发展[J].北京工商大学学报(社会科学版),2022,37(5):59-70.
- [11] 黄勃,李海彤,刘俊岐,等.数字技术创新与中国企业高质量发展——来自企业数字专利的证据[J].经济研究,2023,58(3):97-115.
- [12] 栗晓云,夏传信,施建军.数字技术驱动制造企业高质量发展战略研究——基于三一重工、特斯拉和酷特智能的多个案例研究[J].技术经济,2023,42(5):149-161.
- [13] 岳宇君,洪长威.数字政府建设与企业高质量发展:内在机理与实证分析[J].学术交流,2023(7):88-103.
- [14] 吴浩强,胡苏敏.数字化转型、技术创新与企业高质量发展[J].中南财经政法大学学报,2023(1):136-145.
- [15] 高跃,王跃堂,冀云阳.数字化转型赋能企业高质量发展——基于城市异质性视角的分析[J].城市问题,2023(6):95-103.
- [16] 王羲,陈雪姣,曹晶,等.数字化转型对企业高质量发展的影响——企业创新与风险承担视角[J].科技进步与对策,2024,41(7):1-10.
- [17] 杨向阳,徐从才.数据要素与流通企业高质量发展[J].商业经济与管理,2024(3):5-17.
- [18] 李健,董小凡,张金林,等.数据资产对企业创新投入的影响研究[J].外国经济与管理,2023,45(12):18-33.
- [19] HU C, LI Y, ZHENG X. Data assets, information uses, and operational efficiency[J]. Applied Economics, 2022, 54(60): 6887-6900.
- [20] KIRON D. Lessons from becoming a data-driven organization[J]. MIT Sloan Management Review, 2017, 58(2).
- [21] SIVARAJAH U, IRANI Z, GUPTA S, et al. Role of big data and social media analytics for business to business sustainability: a participatory web context[J]. Industrial Marketing Management, 2020, 86: 163-179.
- [22] 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.
- [23] 苑泽明,于翔,李萌.数据资产信息披露、机构投资者异质性与企业价值[J].现代财经(天津财经大学学报),2022,42(11):32-47.
- [24] 路征,周婷,王理,等.数据资产与企业发展——来自中国上市公司的经验证据[J].产业经济研究,2023(4):128-142.
- [25] 陆岷峰,王稳华,朱震.数据资产如何赋能企业高质量发展——基于产能利用率视角的经验证据[J].上海商学院学报,2023,24(4):22-41.
- [26] 韩秀兰,王思贤.数据资产的属性、识别和估价方法[J].统计与信息论坛,2023,38(8):3-13.
- [27] 罗玫,李金璞,汤珂.企业数据资产化:会计确认与价值评估[J].清华大学学报(哲学社会科学版),2023,38(5):195-209+226.
- [28] SONG Z, WU G L. Identifying capital misallocation[R]. Chicago: Univ. Chicago, Work Pap, 2015.
- [29] 陈诗一,陈登科.中国资源配置效率动态演化——纳入能源要素的新视角[J].中国社会科学,2017(4):67-83+206-207.
- [30] 才国伟,杨豪.外商直接投资能否改善中国要素市场扭曲[J].中国工业经济,2019(10):42-60.
- [31] 陈冬梅,王俐珍,陈安霓.数字化与战略管理理论——回顾、挑战与展望[J].管理世界,2020,36(5):220-236+20.
- [32] 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+2.
- [33] 刘景江,郑畅然,洪永森.机器学习如何赋能管理学研究?——国内外前沿综述和未来展望[J].管理世界,2023,39(9):191-216.
- [34] 胡楠,薛付婧,王昊楠.管理者短视主义影响企业长期投资吗?——基于文本分析和机器学习[J].管理世界,2021,37(5):139-156+11+19-21.
- [35] 孙昌玲,王化成,王芃芃.企业核心竞争力对供应链融资的影响:资金支持还是占用?[J].中国软科学,2021(6):

- 120-134.
- [36] 鲁晓东,连玉君. 中国工业企业全要素生产率估计:1999—2007[J]. 经济学(季刊),2012,11(2):541-558.
- [37] 黄先海,高亚兴. 数实产业技术融合与企业全要素生产率——基于中国企业专利信息的研究[J]. 中国工业经济,2023(11):118-136.
- [38] 李琦,刘力钢,邵剑兵. 数字化转型、供应链集成与企业绩效——企业家精神的调节效应[J]. 经济管理,2021,43(10):5-23.
- [39] 卜令通,张嘉伟. 基于PMC指数模型的数字经济政策量化评价[J]. 统计与决策,2023,39(7):22-27.
- [40] 袁淳,肖土盛,耿春晓,等. 数字化转型与企业分工:专业化还是纵向一体化[J]. 中国工业经济,2021(9):137-155.
- [41] 戴翔,马皓巍. 数字化转型、出口增长与低加成率陷阱[J]. 中国工业经济,2023(5):61-79.
- [42] 江艇. 因果推断经验研究中的中介效应与调节效应[J]. 中国工业经济,2022(5):100-120.
- [43] 刘姝雯,刘建秋,阳旸. 企业金融化与生产效率:“催化剂”还是“绊脚石”[J]. 南开管理评论,2023,26(1):55-68.
- [44] GHOSH D, OLSEN L. Environmental uncertainty and managers' use of discretionary accruals [J]. Accounting Organizations and Society,2009,34(2):188-205.
- [45] 申慧慧,于鹏,吴联生. 国有股权、环境不确定性与投资效率[J]. 经济研究,2012,47(7):113-126.
- [46] 张桅,胡艳. 长三角地区创新型人力资本对绿色全要素生产率的影响——基于空间杜宾模型的实证分析[J]. 中国人口·资源与环境,2020,30(9):106-120.
- [47] 王小鲁,樊纲,胡李鹏. 中国分省份市场化指数报告(2018)[M]. 北京:社会科学文献出版社,2019.
- [48] 金灿阳,徐蔼婷,邱可阳. 中国省域数字经济发展水平测度及其空间关联研究[J]. 统计与信息论坛,2022,37(6):11-21.
- [49] 陶长琪,丁煜. 数字经济政策如何影响制造业企业创新——基于适宜性供给的视角[J]. 当代财经,2022(3):16-27.

Empowering High-quality Development of Enterprises with Data Assets: Optimization Mechanism of Traditional Production Factors

YUAN Ze-ming¹, YIN Qi¹, YU Xiang²

(1. School of Accounting, Tianjin University of Finance and Economics, Tianjin 300222, China;

2. School of Business, Nankai University, Tianjin 300071, China)

Abstract: In the era of the digital economy, accelerating the construction of Digital China and promoting high-quality development of the real economy are important engines for advancing Chinese modernization. Against this background, data is experiencing exponential growth, driving profound changes in social production and lifestyles. Enterprise data assets have also become a focal topic in academia. As a new factor of production in the digital economy era, data assets are embedded in various fields of the real economy, relying on data industrialization and industrial data-oriented businesses, serving as a key factor in talent upgrading, technological innovation, optimal capital allocation, and management efficiency improvement, thereby promoting high-quality development of enterprises.

Based on the data of A-share listed companies in Shanghai and Shenzhen Stock Markets from 2007 to 2022, this paper describes the data assets using text analysis, tests its influence on the high-quality development of enterprises and its mechanism of action, and identifies the regulatory effects in different external

environments. It is found that data assets are the effective driving force of high-quality development. This conclusion is still valid after the robustness test and the endogeneity test. The mechanism test shows that the application of data assets can improve the development quality of enterprises through four channels: human resource optimization mechanism, capital allocation improvement mechanism, technology resource accumulation mechanism, and management resource efficiency mechanism. Further analysis shows that, for the areas with high environmental uncertainty, information level, marketization degree, and government's concern about the digital economy, data assets have a more significant role in promoting the high-quality development of enterprises. The conclusion of this paper provides a micro-theoretical basis and policy enlightenment for enterprises to use data assets to promote the optimization and upgrading of traditional production factors and achieve high-quality development.

The marginal contributions of this paper are as follows: firstly, in terms of research content, this paper examines the impact of data assets on the high-quality development of enterprises, not only providing a "new perspective" for high-quality development but also expanding the economic consequences of data assets from the dual attributes of "data" and "assets". Secondly, in terms of variable measurement, this paper selects seed words based on the underlying logic of data asset "acquisition-processing-application", uses the Autophrase method to construct a "data asset" text dictionary, and then mines textual information from annual reports through sentiment analysis to measure the level of enterprise data assets. Furthermore, the validity of this indicator is verified through feature word spectrum tests and comparative tests, providing useful references for innovating the measurement of enterprise data asset levels. Thirdly, in terms of analysis perspective and mechanism testing, this paper takes data assets as a new factor of digital economic development and the basic fact that they can substitute inputs and optimize resources for traditional production factors as a research starting point, providing evidence for the impact of enterprise data asset mechanisms on high-quality development through four channels: human resources optimization mechanism, capital allocation improvement mechanism, technological resource accumulation mechanism, and management resource efficiency enhancement mechanism, and further discusses the regulatory effects of regional and industry environments, providing theoretical support and empirical evidence for differentiated strategies for high-quality development of enterprises.

Key words: data assets; high-quality development; traditional production factors; labor structure; capital allocation efficiency; technical level; management efficiency

CLC number: F27; F273.4

Document code: A

Article ID: 1674-8131(2024)03-0054-20

(编辑:刘仁芳)