

# 数据交易平台建设如何影响企业全要素生产率<sup>\*</sup>

戴魁早 王思曼 黄姿

**摘要:**本文以政府参与型数据交易平台建设反映数据要素市场发展,在理论分析数据交易平台建设影响企业全要素生产率的基础上,结合数据交易平台建设实践与2011—2020年A股上市公司数据,运用渐进DID方法对理论预期进行了验证。研究发现:数据交易平台建设显著地促进了制造业企业全要素生产率增长,这一结论在安慰剂检验、PSM-DID等一系列稳健性检验后仍然成立;数据交易平台建设的促进作用是通过降低交易成本、提高创新效率与生产效率等机制实现的。数字基础设施、数据应用场景等数据产业链关键环节强化了数据交易平台建设的促进作用;而且数据法治建设、数据安全保障等数据治理水平提高,增强了数据交易平台建设的影响。此外,对于规模较大、智能化水平较高、技术密集度较高以及政府数据开放度较高地区的企业来说,数据交易平台建设的促进作用更强。本文的研究验证了数据交易平台建设在制造业全要素生产率增长中的重要作用,为数据要素市场发展与制造业高质量发展提供了有益的参考。

**关键词:**要素市场 数据交易平台 生产率 数据产业链 数据治理

## 一、引言

数据要素作为一种新的生产要素,突破了传统要素的竞争性和稀缺性限制,具有非竞争性和可再生性两个显著特征,是国家基础性战略资源。理论上,数据要素被引入知识积累和生产过程以后,可以使得企业创新和生产活动更有效率,能够更好地促进长期经济增长(Jones & Tonetti, 2020; Cong et al, 2021, 2022);而要素市场发展(或数据要素市场化配置)<sup>①</sup>可以激发数据要素潜能,能够充分发挥数据要素在推动经济发展质量变革、效率变革、动力变革中的价值。实际上,我国非常重视数据要素市场发展,2020年3月,中共中央、国务院出台《关于构建更加完善的要素市场化配置体制机制的意见》,<sup>②</sup>提出“要建立健全统一开放的要素市场”,明确要求“要加快培育和发展数据要素市场”,这标志着数据要素市场化配置已经上升为国家战略,将对我国经济高质量发展产生深远影响。2022年12月,中共中央、国务院印发《关于构建数据基础制度更好发挥数据要素作用的意见》,助力数据要素市场体系建设,标志着数据要素市场培育进入到全面推进的新阶段。鉴于制造业是国民经

<sup>\*</sup> 戴魁早,湖南科技大学商学院、区域经济高质量发展中心,邮政编码:411201,电子邮箱:daikz2007@sina.com;王思曼、黄姿,湖南科技大学商学院,邮政编码:411201,电子邮箱:20301502001@hnust.edu.cn, huangzii2022@163.com。基金项目:国家自然科学基金项目“技术要素市场发展对中国制造业生产率增长的影响机制及调控政策研究”(72173042);湖南省教育厅重点项目“劳动力要素市场化改革推动湖南省高质量发展的内在机理与政策协同研究”(23A0383)。感谢匿名审稿专家的建设性意见,文责自负。

<sup>①</sup> 国家工业信息安全发展研究中心《中国数据要素市场发展报告(2020—2021)》明确指出,数据要素市场化配置就是将尚未完全由市场配置的数据要素转向由市场配置的动态过程,其目的是形成以市场为根本调配机制,实现数据流动的价值或者数据在流动中产生价值。

<sup>②</sup> 2021年3月,《中华人民共和国国民经济和社会发展第十四个五年规划和2035年远景目标纲要》进一步强调,“要推动大数据采集等技术创新,培育数据采集、标注、存储、传输、管理、应用等全生命周期产业体系,完善大数据标准体系”。

济的支柱产业,而制造业高质量发展本质上是质量和效益替代规模和增速,就是需要持续提高制造业全要素生产率。因而,在新发展阶段,探究数据要素市场发展对制造业企业生产率增长的影响,是值得研究的重大课题。

实际上,已有大量文献关注了要素市场发展对全要素生产率的影响。相关文献集中于要素市场扭曲带来的资源效率损失。研究发现,贫困国家与发达国家的资源错置的不同造成了国家间全要素生产率的差异(Hsieh & Klenow, 2009),中国要素市场发展相对滞后形成的要素价格扭曲引致的资源错置效应,显著地抑制了企业或产业生产率的提高(Hsieh & Klenow, 2009; 罗德明等, 2012; 曹亚军, 2019)。尽管上述文献深入解释了要素市场发展如何影响企业生产率,但尚未关注到数据要素市场——这一新型要素市场对企业生产率的可能影响。事实上,数据要素突破了传统要素的竞争性和稀缺性限制,对生产率的影响明显区别于传统要素(Jones & Tonetti, 2020; Cong et al, 2021, 2022);因而,数据要素市场发展对生产率增长的影响可能有自身特殊规律。这为本文的研究提供了可能的突破空间。

数据要素市场是数据要素的交易场所及其交易活动的总和。由于数据交易平台为社会化的、自由有序的数据要素交易与流动,提供了规范有序的交易场所,因而,数据交易平台对数据要素市场发展至关重要。<sup>①</sup>由此可以认为,数据交易平台建设的效果能够在很大程度上反映数据要素市场发展的影响。理论上,数据交易平台带来的海量数据交易或服务,可以改善产品和要素供需之间的信息不对称问题,不仅可以降低企业的交易成本,还可以提高企业内部的创新效率和生产效率(Cong et al, 2021, 2022),因而很可能对制造业企业生产率增长产生重要影响。实际上,我国各地区从2014年起就陆续开始进行数据交易平台建设的有益探索和实践,截至2021年,全国共建立了28家政府参与型数据交易平台。数据交易平台建设作为一个有效的准自然实验,为我们研究数据要素市场发展对制造业企业生产率增长的影响提供了一个绝佳样本。

与现有要素市场发展领域的文献不同,本文基于数据交易平台建设是我国数据要素市场发展的关键举措这一事实,以数据交易平台建设反映数据要素市场发展,采用DID方法评估数据要素市场发展是否促进了制造业企业生产率增长。具体地,在理论分析数据交易平台建设影响企业生产率机理的基础上,结合数据交易平台建设实践与A股上市企业数据,运用DID方法对理论预期进行验证;进一步地,探讨了数据产业链关键环节和数据治理在数据交易平台建设促进企业生产率增长中的作用。

与现有文献相比,本文的可能贡献主要体现在:第一,以数据交易平台建设反映数据要素市场发展,创新性地分析了数据交易平台建设对制造业企业全要素生产率增长的影响。这拓展了数据要素、要素市场发展与企业生产率的研究领域,极大地补充了现有研究。第二,考察了数据产业链关键环节是否会影响数据交易平台建设的作用效果,并探讨了数据治理在数据交易平台建设促进企业生产率增长中的作用。这不仅使得数据交易平台建设的影响研究更接近于我国数据要素市场发展的客观实际,还加深了数据交易平台建设影响企业生产率内在规律的认识和理解,为通过发展数据要素市场推动制造业生产率增长提供了参考依据。第三,探究了数据交易平台建设促进企业生产率增长的内在机制,并考察了数据交易平台建设对规模、智能化水平、技术密集度等特征不同企业的异质性影响。这丰富了要素市场与生产率增长的研究内容,在政策层面更具明确的启示意义。

## 二、文献回顾与数据交易平台建设实践

### (一)文献简单回顾

1. 要素市场发展与企业生产率增长。这类研究大多集中在要素市场扭曲的经济影响,研究发

<sup>①</sup>数据交易平台是数据要素流通、数据要素交易等活动的场所,业务活动涵盖了数据要素市场的六大模块(数据采集、数据存储、数据加工、数据流通、数据分析、生态保障),因此,数据交易平台建设对数据要素市场发展至关重要。

现,要素市场发展相对滞后形成的要素市场扭曲显著地抑制了企业或产业生产率的提高(Hsieh & Klenow,2009;罗德明等,2012;曹亚军,2019)。与这类研究有着紧密联系的文献还有关于经济体制改革(或市场化改革)对生产率的影响,马光荣(2014)着重关注了市场化转型带来的企业内部激励机制与企业间资源配置改善对生产率产生的促进作用,毛其淋和许家云(2015)分析了市场化转型如何通过就业再配置这一途径提高企业生产率,孙晓华等(2020)则认为市场化改革通过同质性集聚和异质性集聚作用促进了制造业企业生产率增长。

2. 数据要素与内生经济增长。数据要素对经济增长的影响受到了关注,代表性文献仅有 Jones & Tonetti(2020)、Cong et al(2021,2022)等。Cong et al(2021)构建一个新的内生经济增长模型,将消费者所产生的数据作为一种关键要素引入知识积累的过程当中;研究发现,数据要素使得企业创新更有效率,产生出更多的新知识、新技术、新专利,进而提高长期经济增长率。Jones & Tonetti(2020)以数据要素“水平非竞争性”为假设前提,理论模型的结果显示,将数据要素产权授予消费者可以提高生产效率,而将数据要素产权授予生产者则会导致数据要素的低效率使用。而 Cong et al(2022)则进一步假设数据要素具有“垂直非竞争性”,理论模型的研究发现,数据要素既可以提高创新过程的效率,又可以提高生产过程的效率,进而提高了长期经济增长率。

3. 数据交易平台的经济效应。这类研究的代表性文献有 Akcigit & Liu(2016)、Bergemann & Bonatti(2019)和陈舟等(2022)等。Akcigit & Liu(2016)构建了一个两企业参与、成功不确定的探索创新模型研究发现,数据交易平台作为第三方数据机构可以帮助数据买方补偿创新失败的损失,弥补企业激励不足带来的社会福利损失。Bergemann & Bonatti(2019)则认为,企业从数据交易平台购买个人消费者信息,能帮助企业通过实行三级价格歧视获得超额利润,但是这种行为会造成消费者损失,进而带来社会总福利损失。而陈舟等(2022)的定性分析认为,数据交易平台提炼的数据信息、知识和智慧,可以通过乘数效应作用于全要素生产率。

综上所述,就研究对象而言,上述文献并未从制度层面探讨数据要素市场发展对企业生产率增长的影响,无法真实反映我国各地区正在建设的数据交易平台对制造业企业生产率增长的影响。<sup>①</sup>就研究内容而言,上述文献并未探讨数据产业链的关键环节和数据治理在数据交易平台建设促进企业生产率增长中的可能影响。就关于数据交易平台如何影响企业生产率增长而言,上述文献尚未从交易成本、创新效率和生产效率视角进行解释。

由于数据交易平台建设的效果能够在很大程度上反映数据要素市场发展的影响,而且数据交易平台建设是一个有效的准自然实验,这为探究数据要素市场发展如何影响企业生产率增长提供了绝佳的机会。此外,鉴于数据交易(或数据服务)是数据产业链的一个环节,其活动不仅会受到上游关键环节——数字基础设施、下游关键环节——数据应用场景的影响,还会受到的数据法治建设、数据安全保障等方面数据治理水平的影响,因此,我们可以将数据产业链、数据治理与数据交易平台结合起来展开研究,这使得关于数据交易平台建设影响的研究更接近于我国数据要素市场的客观实际。

## (二)数据交易平台建设实践

从数据要素的产生到发生要素作用的过程来界定,数据要素市场主要包括数据采集、数据存储、数据加工、数据流通、数据分析、生态保障六大模块。鉴于数据要素自主有序流动是数据要素市场化配置的前提,而数据交易平台是数据要素流通、数据要素交易等数据服务活动的场所,且其业务活动涵盖了数据要素市场的六大模块,因此,数据交易平台建设对数据要素市场发展乃至数据要素市场化配置至关重要。

实际上,为了推动数据要素市场发展(或数据要素交易),我国地方政府从2014年起就陆续开始

<sup>①</sup>实际上,第一类和第二类文献并未从制度层面探讨数据要素市场的影响。第三类文献是通过构建理论模型、定性分析数据交易平台的经济效应,但并没有专题研究数据交易平台建设对企业生产率增长的影响,因而,无法反映我国各地区数据要素市场发展对制造业生产率增长的真实影响。

了政府参与型数据交易平台建设的有益探索和实践。这类数据交易平台依托隐私计算、区块链等数字技术,采取的是国资控股、管理层持股和主要数据提供方参股的混合所有制模式。从发展历程来看,我国政府参与型数据交易平台的建设可以划分为三个阶段:

第一阶段为初创期。2014年,北京成立了中关村大数据交易平台和北京大数据交易服务平台,标志着我国正式进入数据交易平台建设的探索阶段。第二阶段为快速发展期。这一阶段数据交易平台发展迅猛,数量规模迅速扩大、市场占有率持续提高、业务范围不断拓宽。2015年4月,贵阳大数据交易所在贵阳市国资委的支持下挂牌运营;8月,华中地区第一家数据交易所—长江大数据交易中心落户武汉;此后,西咸新区大数据交易所、武汉东湖大数据交易所、河北京津冀数据交易中心等交易平台也相继成立。截至2017年,全国共成立了22家由地方政府发起、指导或批准的数据交易平台。<sup>①</sup> 第三阶段为成长期。由于前期数据交易的市场活跃程度并没有达到预期,中央政府采取了一系列政策加快了数据交易平台与流通体系的建设,越来越多的地方政府把建设数据交易平台写入大数据发展规划中。2019年起,新一批数据交易平台陆续在山东、山西、广西北部湾和北京成立。截至2021年,全国共建立了28家政府参与型数据交易平台。

随着政府参与型数据交易平台的建立和运营,我国数据要素的交易规模也在迅速上升。据国家工信安全中心测算数据,我国数据要素市场规模从2016年的62亿元迅速上升到2020年的545亿元,四年增长了779%,年均增长率高达72%左右。<sup>②</sup>

### 三、理论分析

从产业链视角,数据产业可以划分为基础支撑、数据服务和融合应用三层业态,数据交易平台的业务活动对应于数据服务这层业态,而数据基础设施、数据应用场景则对应于基础支撑、融合应用。依据产业链理论,产业链协同状况决定了整个产业的价值创造和价值实现,因而,数据交易平台的价值创造和价值实现,会受到数据基础设施、数据应用场景的影响。与此同时,数据治理作为实现提升数据质量与应用价值、促进数据资源整合与流通、保障数据安全等目标的制度安排,<sup>③</sup>对数据交易平台流通的数据质量和数据应用价值有着重要影响,因此,数据交易平台的价值创造和价值实现,还会受到数据治理水平的影响。

结合在数据交易平台建设过程中,数据产业链关键环节不断发展与数据治理水平不断提高的客观事实,以及我国经济进入高质量发展阶段以后迫切需要通过实现制造业企业生产率持续提升的现实需要,通过梳理相关领域研究成果,本文构建如下理论分析框架探讨数据交易平台建设与企业全要素生产率增长的关系(如图1所示)。

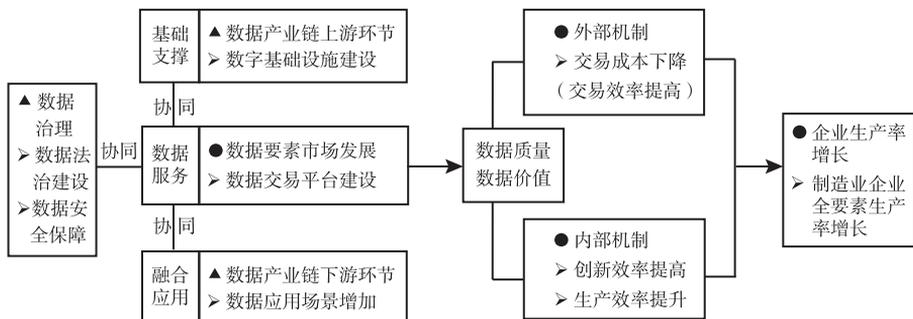


图1 理论分析框架

①数据来源于2021年12月中国信息通信研究院发布的《大数据白皮书(2021)》。

②数据来源于2022年11月国家工业信息安全发展研究中心发布的《中国数据要素市场发展报告(2020~2021)》。

③这里的数据治理定义源于2021年12月中国信息通信研究院发布的《大数据白皮书(2021)》。

### (一)数据交易平台与企业全要素生产率增长

依据内生增长理论,全要素生产率增长本质上是资源配置效率提高,而优化资源配置能够提高全要素生产率。理论上,数据交易平台建设带来的数据质量和数据价值提高,能够更好地发挥数据要素的应用价值,提高资源配置效率。从产业层面来看,能够大大降低市场的不确定,使得要素资源以更快的速度流向更高效率的地区和行业(即效率更高的制造业企业),从总体上提高了制造业的资源配置效率。从企业层面来看,既可以降低制造业企业外部交易成本,提高企业间的资源配置效率;又可以缓解企业创新环节和生产环节的信息不对称问题,提高企业的创新效率和生产效率。

1. 外部机制:交易成本降低机制。理论上,数据交易平台可以更好地采集和整合碎片化、分割分散等数据资源,形成海量的高质量和高价值的规范化、体系化数据要素,可以从两个方面降低制造业企业外部的交易成本。一方面,降低了现有交易的成本及交易匹配成本,促进了现有交易发生,这提高了企业现有交易的资源利用效率(李文莲、夏健明,2013;韩先锋等,2014;许宪春等,2019)。另一方面,降低了产品和要素资源(人才、资金和技术)等交易发生的信息门槛,降低了制造业企业的搜寻成本,匹配了大量产品和要素资源的新交易,创造了大量原本没有发生的新交易,这也提高了企业资源配置效率。

2. 内部机制:创新效率提高机制。理论上,数据交易平台形成海量的高质量和高价值的要素,可以帮助制造业企业在研发前期获取更具前瞻性的市场需求信息,以满足头部市场、头部产品、头部用户的标准化、大众化需求;而企业更有针对性的新产品研发活动,可以大大降低研发失败的风险,提高企业的创新效率(Cockburn et al,2019)。实际上,Cong et al(2021)构建一个新的内生经济增长模型发现,将消费者所产生的数据作为一种关键要素引入到知识积累过程以后,数据要素使得企业创新更有效率,产生出更多的新知识、新技术、新专利。

3. 内部机制:生产效率提升机制。理论上,数据交易平台形成海量的高质量和高价值的要素,可以从两个方面提高制造业企业生产效率。一方面,既降低企业生产过程中人才、资金和设备等要素资源的信息不对称,优化企业现有的要素资源配置(陈中飞、江康奇,2021);又可以与企业现有要素结合,被应用到企业生产流程再造中(Baesens et al,2016),改进企业的生产工艺,提高企业生产效率。另一方面,制造业企业将针对市场需求而研发的新知识、新技术、新专利,应用到企业生产过程中,也可以提升企业的生产效率(Cong et al,2022)。

综上所述,由此有如下待检验的两个假说:

假说1:数据交易平台建设促进了制造业企业全要素生产率增长。

假说2:数据交易平台建设可以通过降低交易成本、提高创新效率与生产效率等机制,促进制造业企业全要素生产率增长。

### (二)数据交易平台、数据产业链与企业全要素生产率增长

1. 数据产业链上游关键环节:数字基础设施的影响。数字基础设施是数据产业链上游关键环节,为数据交易平台的数据服务提供的技术支撑。实际上,随着我国工业互联网基础设施建设逐渐完善,5G、物联网和区块链等技术支撑起工业数据采集、汇聚、流通、分析、应用的价值闭环,这一过程给数据交易平台提供海量的数据和更智能的数据收集和分析工具(李海舰、赵丽,2021;冀雁龙等,2022)。理论上,数据交易平台得到更为先进数字基础设施(包括网络设备、计算机设备、存储设备等硬件供应和软件分析工具)的支撑,能够更为迅速地采集、分析和挖掘更大规模、更高维度的数据资源,形成更高质量和更高价值的体系化数据要素,因而可以更好地为区域内制造业企业解决信息不对称问题,进而提高制造业企业要素资源配置效率。

2. 数据产业链下游关键环节:数据应用场景的影响。数据应用场景是数据产业链下游关键环节,是数据服务(或数据交易)的应用环节,决定了数据要素的需求,为数据交易平台的业务活动提供需求激励。事实上,我国轨道交通、装备制造、航空航天等制造业企业已经具备了较好的数据积累和处理能力,且应用场景也从最初的生产监控到降本增效,逐步向支撑服务化转型。理论上,更宽广的数据应用场景增加了对数据服务的需要,为数据交易平台的业务活动提供了更好的激励,有助于激

发数据交易平台创新数据挖掘技术,形成适应不同应用场景的分类体系化数据要素,进而可以更好地缓解企业的信息不对称问题,提高制造业企业资源配置效率。

综上所述,有如下待检验的假说:

假说3:数据交易平台建设对制造业企业生产率增长的促进作用,会受到数字基础设施建设、数据应用场景等数据产业链关键环节的影响。

### (三)数据交易平台、数据治理与企业全要素生产率增长

1. 数据法治建设的影响。数据成为一种生产要素,必然具有产权等资产的诸多属性(产权是其他所有资产属性的前提),需要被相关法律充分认同、明确界定和完全保护;同时,数据交易平台如何对个体信息及个体信息集合进行收集、加工、交换等,也需要法律体系支撑。理论上,健全的数据法律制度缓解了条块分割、不完全契约和违法违规交易等,促进了数据开放与共享(赵云辉等,2019;尹华容、王惠民,2022)。因而,在更优的数据法治环境下,数据交易平台可以更方便获取高质量和高价值的数据库资源,并以更低的数据要素交易成本,为制造业企业提供数据库服务,进而有助于企业更好地应用数据库要素,优化配置资源。实际上,我国一直在不断探索和完善数据库要素产权和数据库交易等方面的法律法规体系,数据库法治相关制度体系建设不断取得新成效。由此可以推测,数据库法治建设可以增强数据交易平台对制造业企业全要素生产率增长的促进作用。

2. 数据库安全保障的影响。在数据库处理过程中,存在个人隐私、个人信息泄露与数据库攻击等数据库安全风险。理论上,在数据库安全制度体系不完善、数据库安全技术落后的情况下,数据库安全风险程度较高;社会公众对数据库交易的不信任程度较高,参与数据库交易的主动性、积极性较低;数据库交易平台为制造业企业提供高质量和高价值的数据库要素难度较大,成本更高。这种情形下,不利于充分发挥数据库要素在优化配置企业要素资源中的应用价值。实际上,近年来我国一直在逐步完善数据库安全制度建设、大力发展数据库安全技术来保障数据库安全,数据库安全保障不断取得新成效。由此可以推测,数据库安全保障可以增强数据交易平台建设的促进作用。

综上所述,由此有如下待检验的假说:

假说4:数据交易平台建设对制造业企业生产率增长的促进作用,会受到数据库法治建设、数据库安全保障等数据库治理状况的影响。

## 四、研究设计

### (一)计量模型

为了检验数据库要素市场发展对制造业企业全要素生产率的影响,本文采用渐进 DID 方法进行经验研究(Beck et al,2010),具体模型设定如下:

$$Y_{it} = \alpha_0 + \alpha_1 D_{it} + \alpha_2 [X_p \times f(t)] + \tau_t + \mu_i + \varepsilon_{it} \quad (1)$$

式(1)中, $i$ 代表企业, $t$ 代表年份。 $Y_{it}$ 为 $i$ 制造业企业在 $t$ 年的全要素生产率。 $D_{it}$ 作为双重差分变量,是按照《大数据白皮书(2021)》中公布的数据交易平台建立时间作为划分标准。当 $i$ 企业所在城市 $t$ 年建立了数据交易平台,则在当年及该年之后, $D_{it}$ 赋值为1,否则 $D_{it}$ 赋值为0。如果 $D_{it}$ 系数 $\alpha_1$ 显著为正,说明假设1提出的数据交易平台建设对当地制造业企业全要素生产率增长有正向影响得到了验证。 $X_p$ 为企业层面和地区层面的前定变量; $\tau_t$ 为年份固定效应, $\mu_i$ 为企业固定效应, $\varepsilon_{it}$ 为随机扰动项。考虑到企业层面聚类效应可能对结果造成的偏误,本文对标准误差进行了企业层面的聚类调整(cluster at firm level)。

在本文样本期间内(2011—2020年),国内先后在16个省,20个城市成立了26家由地方政府发起、指导或批准成立的数据交易平台。<sup>①</sup>考虑到研究需要和数据的可获得性,本文的研究对象为

<sup>①</sup>城市包括北京、贵阳、西安、重庆、盐城、武汉、承德、哈尔滨、上海、广州、杭州、深圳、郑州、青岛、潍坊、新乡、长春、济南、太原、南宁。

2011—2020年间所有A股上市制造业企业。样本处理时,剔除了非正常交易的上市制造业企业(ST、\*ST和PT)和数据大量缺失的企业,最终得到9300个面板观测值。

(二)变量说明

1. 全要素生产率。本文借鉴企业层面全要素生产率测算的通常做法(杨汝岱,2015;黄群慧等,2019;王浩等,2022),利用LP方法(Levinsohn & Petrin,2003)来测算制造业企业全要素生产率。具体来说,关于企业产出和中间投入衡量,借鉴刘莉亚等(2018)的做法,采用销售收入进行衡量制造业企业产出;用销售额减去增加值来度量制造业企业的中间投入;而增加值为折旧、劳动者报酬、生产税净额和营业盈余四项之和。关于劳动力投入和资本投入的衡量,本文参考何光辉和杨咸月(2012)的做法,用现金流量表中“支付给职工以及为职工支付的现金”(即企业实际劳动成本)反映劳动力投入,用企业固定资产净值来衡量资本投入变量。

2. 控制变量。本文参考该领域文献的通常做法(范子英等,2016;金刚、沈坤荣,2018;贾智杰等,2023),从外商直接投资、人力资本水平、交通基础设施、政府参与度、第三产业比重、研发投入强度等地区层面选取控制变量。从企业资本密度、企业管理水平、存货周转率、企业财务状况等企业层面选取控制变量。同时,为了避免控制事后变量导致估计结果存在的可能偏误,本文借鉴该领域文献的通常做法(陈诗一等,2021),控制数据交易平台建设以前某一年(2013年)前定变量与时间趋势 $f(t)$ 的乘积项,以解决可能存在的内生性问题和估计结果偏误。

(三)数据说明

考虑到数据的可获得性,本文的样本区间为2011—2020年。数据来源包括几个部分:一是制造业企业全要素生产率测算数据。相关数据来源于国泰安数据库和Wind数据库。二是数据交易平台数据。本文手工整理《大数据白皮书(2021)》中公布的数据交易平台数据,并将其与Wind数据库中A股上市制造业企业进行匹配,进而识别出所在城市建立了数据交易平台的上市制造业企业(实验组)和所在城市未建立数据交易平台的上市制造业企业(控制组)。三是控制变量数据。地区层面控制变量的测算数据主要源于《中国统计年鉴》《中国工业统计年鉴》《中国劳动统计年鉴》和《中国科技统计年鉴》相关年度,而企业层面控制变量的测算数据主要来源于国泰安数据库和Wind数据库。为了避免极端值对估计结果的干扰,本文对连续变量进行1%分位的缩尾处理。主要变量描述性统计结果见表1。

表1 主要变量描述性统计

变量名称	变量符号	变量指标或测度方法	均值	标准差	最小值	中位数	最大值
全要素生产率	$Lp$	LP方法测算企业层面生产率	8.1457	0.8998	6.7217	8.0544	10.0660
平台建立虚拟变量	$D$	企业所在城市建立数据交易平台当年及之后的年份, $D$ 取1,否则,取值为0。	0.3303	0.4703	0.0000	0.0000	1.0000
企业资本密集度	$Kl$	企业固定资产/企业从业人员(%)	39.3815	44.9016	0.0000	26.5394	825.6050
企业应收账款周转率	$Retur$	营业收入/平均应收账款(%)	10.6792	19.7740	0.8187	4.9713	48.9224
企业存货周转率	$Intur$	销货成本/存货平均余额(%)	4.5880	4.3151	0.0590	3.4933	58.9130
企业财务状况	$Finance$	负债总额/资产总额(%)	40.6203	20.4596	0.7080	39.6271	78.5560
外商直接投资	$Fdi$	地级市外商直接投资金额/GDP(%)	3.2190	5.7117	0.0000	2.7071	10.4128
人力资本水平	$Hcl$	地级市平均受教育年限(年/人)	8.8007	1.0287	7.1142	8.5087	13.4879
交通基础设施	$Ie$	地级市人均道路面积( $m^2$ )	15.3748	7.9043	0.4056	13.6800	59.3000
政府参与程度	$Govern$	地级市一般预算支出/GDP(%)	16.2589	8.4012	4.3967	13.9098	91.6367
第三产业比重	$Tgdp$	地级市第三产业产值/GDP(%)	39.5613	12.7681	15.3700	35.5690	80.9800
研发投入强度	$Tech$	地级市R&D经费内部支出/GDP(%)	2.3943	1.9823	0.5234	1.6878	10.0123

注:为了直观地反映各个变量的信息,表中是未经过缩尾处理、标准化处理、前定变量处理的描述性统计结果。

## 五、实证结果

### (一) 基准回归结果分析

在估计之前,运用方差膨胀因子方法对解释变量之间的多重共线性进行检验,结果显示,主要模型的方差膨胀因子(VIF)总值和单个解释变量的VIF值都小于10,这表明主要模型的解释变量之间不存在多重共线性。此外,估计时还控制了年度固定效应。

从表2列(1)(2)的估计结果可以看出,无论是否加入控制变量,解释变量( $D_{it}$ )的系数 $\alpha_1$ 在1%水平上显著为正(系数值为0.3114和0.1118),表明数据交易平台建设对制造业企业生产率产生了正向影响,即数据交易平台建设显著地促进了制造业企业生产率增长,验证了假说1。此外,为了避免由于行业特性带来的估计偏误,列(3)(4)报告了加入时间趋势与行业固定效应乘积项的估计结果,可以看出,系数 $\alpha_1$ 在影响方向和显著性方面没有明显的差异(系数值为0.0578和0.0899),这验证了列(1)(2)的结论具有较好的稳健性,即数据交易平台建设对制造业企业生产率增长的促进作用显著且稳健。表2的结果还显示,控制变量也对全要素生产率增长产生了重要影响。上述结论大都符合理论预期。

表2 数据交易平台建设的总体影响

变量	(1)	(2)	(3)	(4)
<i>D</i>	0.3114*** (16.7866)	0.1118*** (5.3537)	0.0578** (2.3768)	0.0899*** (4.5167)
<i>Kl</i>		0.0134 (1.6037)		0.0606*** (3.9262)
<i>Retur</i>		0.1068*** (12.4639)		0.0299* (1.8750)
<i>Intur</i>		0.3100*** (35.8154)		0.2362*** (8.8932)
<i>Finance</i>		0.4001*** (46.6830)		0.1703*** (9.1522)
<i>Fdi</i>		0.0026 (0.2832)		-0.0080 (-0.8317)
<i>Hcl</i>		0.0288*** (2.6179)		0.1101*** (3.0237)
<i>Ie</i>		0.0176* (2.0118)		0.0543*** (4.3701)
<i>Govern</i>		0.0161* (1.6697)		0.0131 (1.3461)
<i>Tgdp</i>		0.0813*** (5.6380)		0.2005*** (7.0078)
<i>Tech</i>		0.0361** (2.4122)		0.1568*** (8.5888)
观测值	9300	9270	9300	9270
时间/企业固定效应	是	是	是	是
行业效应-时间趋势项	否	否	是	是
聚类到企业层面	是	是	是	是
R <sup>2</sup>	0.1084	0.4039	0.2104	0.3151

### (二) 平行趋势检验及动态效应

本文在式(1)基础上,利用事件分析法,构建如下模型进行检验(Beck et al,2010;吕越等,2019;

祝树金、汤超,2020):

$$Y_{it} = \theta_0 + \theta_1 D_{it}^{-4} + \theta_2 D_{it}^{-3} + \theta_3 D_{it}^{-2} + \theta_4 D_{it}^{-1} + \theta_5 D_{it}^0 + \theta_6 D_{it}^1 + \theta_7 D_{it}^2 + \theta_8 D_{it}^3 + \theta_9 D_{it}^4 + \theta_{10} (X_{it} \times f(t)) + \tau_t + \mu_i + \epsilon_{it} \quad (2)$$

式(2)中, $D_{it}^{\pm n}$ 表示数据交易平台建设这一事件,是一个虚拟变量。如果上市企业所在城市建立数据交易平台的前 $n$ 年,则设 $D_{it}^{-n}=1$ ,否则为0;如果上市企业所在城市建立数据交易平台的后 $n$ 年,则设 $D_{it}^{+n}=1$ ,否则为0;如果上市公司所在城市建立数据交易平台的当年, $D_{it}^0=1$ ,否则为0。其他符号的含义同式(1)。回归时,将早于前4年的观测值归并于 $n=-4$ ,晚于后4年的观测值归并于 $n=4$ 。

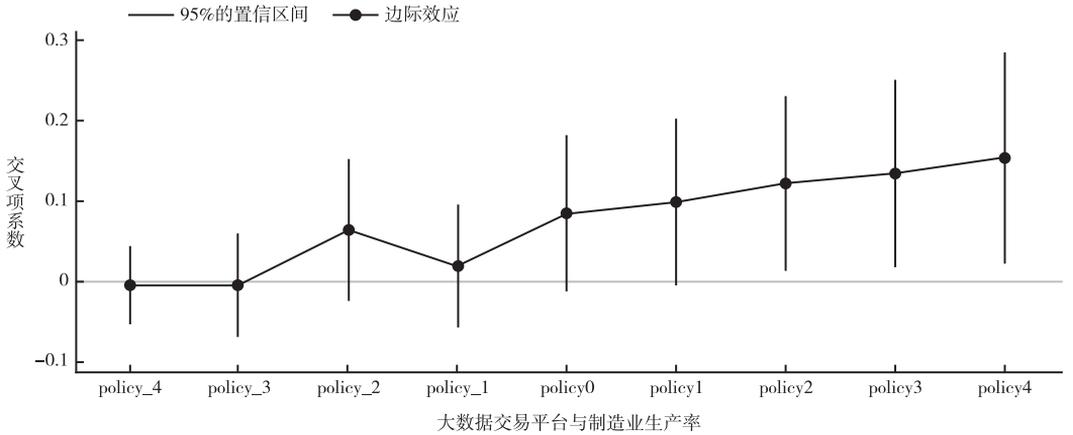


图2 平行趋势检验和动态效应检验

图2绘制的是平行趋势检验结果,可以看出, $D_{it}^0$ 以前的年份,回归结果均不显著,而 $D_{it}^0$ 以后的年份,实验组生产率相比控制组显著上升。因此,样本通过了双重差分法估计所需的平行趋势检验。此外, $D_{it}^1$ 、 $D_{it}^2$ 、 $D_{it}^3$ 和 $D_{it}^4$ 的系数值都显著为正,这说明数据交易平台建设具有动态效应。

### (三) 偏误诊断与稳健估计

1. 偏误诊断。为了验证本文双向固定效应是否存在异质性问题,本文采用 Goodman-Bacon (2021)提出的 Bacon 分解方法来进行诊断。可以看出,坏控制组的估计量所占权重较小(仅5.4%),并且其平均处理效应与第一组符号都是正向的。由此可见,本文双向固定效应估计量由负权重引起的偏误较小,再次验证了表2估计结果的有效性。

2. 异质性稳健估计量。为了进一步修正 TWFE 估计量的潜在偏误,本文首先借鉴 Cengiz et al (2019)的思路,采用事件研究法来进行验证;接着采用 Gardner(2022)提出的两阶段 DID 方法评估数据交易平台建设对企业生产率增长的影响。

### (四) 工具变量结果

为了进一步缓解内生性问题,本文尝试用工具变量的方式来克服内生性问题可能导致的偏误。本文选取夏季平均气温和发生地震等地质灾害次数的倒数作为数据交易平台建设的工具变量(分别用符号 TBD 和 EBD 表示)。由表3列(1)(2)结果可知,固定效应第一阶段的回归结果满足工具变量的相关性假设;Cragg-Donald Wald F 检验(简称 CDF 检验)统计量分别为 19.326 和 35.187,大于 Stock & Yogo(2005)审定的 F 值在 10%偏误水平上的 16.38 的临界值,说明不存在弱工具变量问题。同时,第二阶段回归结果显示,加入协变量后内生变量数据交易平台建设的系数没有发生明显变化,从而间接说明工具变量满足排他性约束(Burchardi & Hassan, 2013),即工具变量是外生的。表3列(3)(4)是 IV Tobit 估计结果显示,第一步回归的工具变量的系数都显著且整个模型的 F 值为 156.60 和 404.56,说明不存在弱工具变量问题,且工具变量满足相关性要求。可见,这里选取的工具变量有效。工具变量第二阶段的估计结果显示,数据交易平台建设的系数值都显著为正,而且,协变量的系数值显著性和影响方向大多相一致,说明前文相应的结论具有较好的稳健性。

表3 工具变量估计结果

	(1)	(2)	(3)	(4)
被解释变量	制造业全要素生产率			
	工具变量 1(TBD)	工具变量 2(EDC)	工具变量 1(TBD)	工具变量 2(EDC)
估计方法	2sls	2sls	IV Tobit	IV Tobit
$D$	0.4976*** (4.7370)	0.3364*** (3.8416)	0.3835*** (4.2566)	0.9509*** (5.3565)
控制变量	是	是	是	是
$R^2$	0.2255	0.5394		
Cragg-Donald Wald F statistic	19.326	35.187		
Kleibergen-Paap rk LM statistic	34.1362***	53.7351***		
F 值			156.60	404.56
第一阶段估计结果				
工具变量	0.0448*** (3.9267)	0.0365*** (2.9117)	0.3425*** (14.8629)	0.3168*** (13.2169)
控制变量	否	是	否	是
观测值	9300	9270	9300	9270
时间/企业固定效应	是	是	是	是
聚类到企业	是	是	是	是
$R^2$	0.0024	0.0773		

注:限于篇幅,资本密集度、应收账款周转率、存货周转率、财务状况等企业层面前定变量以及外商直接投资、人力资本水平、交通基础设施、政府参与程度、第三产业比重、研发投入强度等地区层面前定变量(控制变量)的估计结果未报告。

### (五) 稳健性检验

1. 安慰剂检验。时间安慰剂检验:正如上述分析, DID 使用的前提条件是事件冲击发生之前实验组和控制组生产率增长不存在较大的差异。因此,这里假设将各城市数据交易平台建立的年份提前一年和两年,预期核心解释变量( $D_{it}$ )的估计系数不会显著。企业安慰剂检验:本文借鉴 Chetty et al(2009)、Cantoni et al(2017)的做法,让数据交易平台建设对企业生产率的冲击变得随机。500 个系数估计值的核密度分布图可以发现(限于篇幅,分布图未呈现),数据交易平台建设对制造业企业生产率增长的估计系数均值接近于零且服从正态分布,与基础回归  $D_{it}$  的系数值存在明显差异,这说明基准估计的结果具有稳健性,并不是由某些偶然因素引起的。

2. 更换制造业全要素生产率测算指标。为了保证估计结果的可靠性,本文采用了 OP 方法重新测算了企业全要素生产率;还借鉴了 Gandhi et al(2020)提出的 GNR 方法进行估算,这一方法首先是识别弹性  $e_{ik}$  和随机误差项  $\epsilon_i$ ,再根据弹性识别与中间投入有关的弹性向量  $\theta_1$ ,最后利用所有与劳动、资本投入产出弹性有关的向量,识别出在柔性产出条件下的企业全要素生产率。

3. 基于 PSM-DID 方法的检验。为克服系统性差异,本文利用 PSM-DID 方法进行稳健性检验。本文运用了 1:2 邻近匹配、半径匹配和核匹配等三种不同的匹配方法进行逐年匹配,以检验结果的稳健性。

表4 稳健性检验估计结果

变量	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	更换因变量				PSM-DID		
	OP	OP	GNR	GNR	近邻匹配	半径匹配	核匹配
$D$	0.2946*** (16.6949)	0.0797*** (4.1677)	0.1725*** (12.2092)	0.0738*** (5.4038)	0.0988*** (2.8130)	0.0979*** (4.2984)	0.0979*** (4.2961)
控制变量	否	是	否	是	是	是	是

变量	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	更换因变量				PSM-DID		
	OP	OP	GNR	GNR	近邻匹配	半径匹配	核匹配
观测值	9300	9270	9300	9270	4100	7500	7500
R <sup>2</sup>	0.1020	0.3053	0.0157	0.3821	0.2347	0.2835	0.2835
时间/企业固定效应	是	是	是	是	是	是	是
聚类到企业层面	是	是	是	是	是	是	是

4. 纳入多维固定效应。在前文双重差分双重固定效应模型基础上,加入行业年份固定效应(行业固定效应与时间固定效应交互项),以控制行业层面随时间变化的不可观测变量。

5. 排除其他政策。考虑到 2015 年出台的《促进大数据发展行动纲要》,2020 年出台的《中共中央国务院关于构建更加完善的要素市场化配置体制机制的意见》,其中政策与本文密切相关。因此,本文在基准回归模型中,依次加入政策实施的年份虚拟变量,以控制其对估计结果的影响。

## 六、影响机制与异质性分析

### (一)影响机制分析

这里对理论分析部分的影响机制进行检验,以更好地认识和理解数据交易平台建设影响制造业生产率增长的基本规律。借鉴学术界通常的做法(黄群慧等,2019;毛其淋,2019),采用以下模型进行验证:

$$M_{it} = \rho_1 D_{it} + \rho_2 (X_{it} \times f(t)) + \tau_i + \mu_t + \varepsilon_{it} \quad (3)$$

式(3)中, $M_{it}$ 代表影响机制变量,包括交易成本下降机制、创新效率和生产效率提高机制的代理变量。

1. 交易成本降低机制。本文选用企业销售费用( $Tref$ )作为衡量企业交易成本下降机制的代理变量。理由是,销售费用中很大一部分是企业的广告推广费用,而企业发布广告的最终目标是为了降低信息不对称进而减少交易成本,因而,采用企业销售费用占主营业务收入占比可以较好地反映企业交易成本(刘凤委等,2009;黄群慧等,2019),预期数据交易平台建设对交易成本的影响系数为负。以上数据均来自于 Wind 数据库 A 股上市制造业企业财务报表。表 5 的列(1)(2)汇报了相应的回归结果。可以看出,数据交易平台建设对交易成本的影响系数在 1%的水平上显著为负,值为-0.6927和-0.2633,表明数据交易平台建设显著地降低了制造业企业的交易成本(或提高了交易效率)。这验证了,数据交易平台建设通过交易成本下降这一机制促进了制造业企业生产率增长。

2. 创新效率提高机制。本文选用研发投入与创新产出比值这一变量( $Inno$ )作为创新效率提升机制的代理变量。理由是,创新效率反映了一定创新投入下的科技成果产出水平,又因为投入与产出会出现一定的时间差,因而采用专利申请数与前一年研发投入的比值能在一定程度上反映创新效率的高低(Hirshleifer et al,2013;姚立杰、周颖,2018)。预期数据交易平台建设对创新效率的影响系数为正。数据来自于 CNRDS 数据库 A 股上市制造业企业创新数据。表 5 的列(3)(4)汇报了相应的回归结果,可以看出,数据交易平台建设对创新效率的影响系数在 1%的水平上显著为正,值为 0.4287 和 0.1405,表明数据交易平台建设显著地提升制造业企业的创新效率。这说明,数据交易平台建设通过提升创新效率这一机制促进了制造业生产率增长。

3. 生产效率提升机制。本文选用营业收入与在职员工人数的比值( $Labour$ )作为生产效率提升机制的代理变量。劳动效率反映的是一定量劳动投入所得的有效成果,即单位劳动投入的产出水平,因而,采用营业收入与在职员工人数比值能较好地反映制造业企业劳动者的生产效率(戴魁早等,2020)。预期数据交易平台建设对劳动效率的影响系数为正。数据来自于 Wind 数据库 A 股上

市制造业企业数据。表5列(5)(6)汇报了相应的回归结果,可以看出,数据交易平台建设对劳动效率的影响系数在1%的水平上显著为正,值为0.885和0.8341,表明数据交易平台建设显著地提高了制造业企业劳动者的生产效率。这验证了,数据交易平台建设通过提升生产效率这一途径和机制促进了制造业生产率增长。

表5 数据交易平台建设的影响机制估计结果

变量	(1)	(2)	(3)	(4)	(5)	(6)
	交易成本下降机制		创新效率提高机制		生产效率提升机制	
<i>D</i>	-0.6927*** (-44.2722)	-0.2633*** (-11.0805)	0.4287*** (18.3434)	0.1405*** (4.3787)	0.8850*** (32.5060)	0.8341*** (30.6263)
控制变量	否	是	否	是	否	是
观测值	9300	9272	9300	9270	9300	9270
时间/企业固定效应	是	是	是	是	是	是
聚类到企业层面	是	是	是	是	是	是
<i>R</i> <sup>2</sup>	0.1811	0.2832	0.1285	0.2250	0.1350	0.2065

## (二)异质性分析

理论上,规模不同和行业不同的企业对数据的需求和依赖程度存在差异,不同地区数据交易平台的数据来源也会存在差异。因而,对于不同规模、不同行业、不同地区的企业来说,数据交易平台建设的影响程度可能不同。这里采用如下三重差分法进行验证:

$$Y_{it} = \sigma_0 + \sigma_1 D_{it} \times Size_i + \sigma_2 D_{it} + \sigma_3 DSize_i + \sigma_4 (X_p \times f(t)) + \tau_t + \mu_i + \epsilon_{it} \quad (4)$$

$$Y_{it} = \sigma_0 + \sigma_1 D_{it} \times H_r + \sigma_2 D_{it} + \sigma_3 H_r + \sigma_4 (X_p \times f(t)) + \tau_t + \mu_i + \epsilon_{it} \quad (5)$$

$$Y_{it} = \sigma_0 + \sigma_1 D_{it} \times G_j + \sigma_2 D_{it} + \sigma_3 G_j + \sigma_4 (X_p \times f(t)) + \tau_t + \mu_i + \epsilon_{it} \quad (6)$$

式(4)中, $D_{it} \times Size_i$ 为企业规模乘积项, $Size_i$ 为企业规模虚拟变量。类似的,式(5)中, $D_{it} \times H_r$ 为特定行业乘积项, $H_r$ 为行业特征虚拟变量。式(6)中, $D_{it} \times G_j$ 为地方政府数据开放程度乘积项, $G_j$ 为地区特征虚拟变量。

1. 企业规模。理论上,规模较大的制造业企业在人才与技术等方面具有资源优势(孙晓华、王昀,2014),创新动力更足,更倾向于把数据交易获得的数据用于企业创新中(Cong et al,2022),进而生产率相对较高。因此,对于规模较大的制造业企业来说,数据交易平台建设对生产率增长的促进作用可能更大。为了验证,式(4)中制造业企业总资产值大于中位数时, $Size_i$ 取1;小于中位数时, $Size_i$ 取0。从表6列(1)(2)可以看出,乘积项( $D_{it} \times G_j$ )的系数在1%水平上显著为正,值为0.1867和0.1543。这说明,对于规模较大的制造业企业来说,数据交易平台建设对其生产率增长的促进作用更大。这与理论预期相符。

2. 行业智能化水平。理论上,智能化水平不同的行业,企业生产率会存在差异(王林辉等,2022)。智能化水平较高的行业不仅能把新兴技术融入生产环节,而且还具备了一定的数据分析能力,能把数据交易平台提供的的数据资源转化为生产要素,因此,生产率增长会更快。实际上,国家针对智能化水平相对较高的行业(如汽车行业、通信业)出台了有关数据安全管理的政策,这可能提高了智能化水平较高行业的数据交易效率。由此可以推测,对于智能化水平较高的行业来说,数据交易平台建设对生产率增长的促进作用可能更显著。这里参考芦婷婷和祝志勇(2021)的做法,使用机器人的安装密度进行表征。当行业智能化水平数值大于其中位数时,对式(5)中 $H_1$ 取1;小于其中位数时, $H_1$ 取0。从表6列(3)(4)的估计结果可以看出, $D_{it} \times H_1$ 系数显著为正,值为0.3043和0.0881,这说明,对于智能化水平较高的工业行业来说,数据交易平台建设对其生产率增长的促进作用更大。

表6 异质性估计结果

变量	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	企业规模虚拟变量		行业智能化水平虚拟变量		行业技术密集度虚拟变量		地方政府数据开放程度虚拟变量	
$D \times Size$	0.1867*** (5.4535)	0.1543*** (5.1150)						
$D \times H1$			0.3043*** (12.8440)	0.0881*** (3.7261)				
$D \times H2$					0.2714*** (10.0934)	0.0692** (2.3193)		
$D \times G$							0.3151*** (16.2583)	0.0919*** (4.5554)
$D$	0.2014*** (6.9116)	0.0545* (1.8785)	0.3061*** (12.3943)	0.0675*** (2.6645)	0.3123*** (15.2356)	0.0010 (0.0360)	0.2755*** (4.3459)	0.0342 (0.5375)
异质性虚拟变量	是	是	是	是	是	是	是	是
控制变量	否	是	否	是	否	是	否	是
观测值	9300	9270	9290	9260	9240	9220	9300	9270
时间/企业固定效应	是	是	是	是	是	是	是	是
行业年份固定效应	是	是	是	是	是	是	是	是
聚类到企业层面	是	是	是	是	是	是	是	是
$R^2$	0.1189	0.3561	0.0431	0.3035	0.0428	0.3625	0.1005	0.3051

3. 行业技术密集度。不同技术密集度的行业对创新的依赖程度存在明显差异(戴魁早、刘友金, 2020);技术密集度较低的行业、非熟练劳动力或者资本的投入占主导的行业,对创新的依赖程度较低;而技术密集度更高的行业更依赖于创新,会把数据用于创新过程中,提高了创新过程的效率(Cong et al, 2022),进而提高生产率。实际上,数据交易平台会率先在技术密集度较高的行业构建应用场景(如医疗、电子通信等行业)。由此可以推测,对于技术密集度较高的工业行业来说,数据交易平台建设对其生产率增长的促进作用更大。为了验证,式(5)中行业虚拟变量 $H_i$ 设为 $H2$ 。本文用工业细分行业研发经费支出占总产值的比重衡量行业技术密集度(戴魁早、刘友金, 2020)。当行业技术密集度数值大于其中位数时,对式(7)中 $H2$ 取1,而小于其中位数时, $H2$ 取0。表6列(5)(6)的结果显示, $D_i \times H2$ 的系数为正,值为0.2714和0.0692。以上结果说明,对于技术密集度较高的工业行业来说,数据交易平台建设对其生产率增长的促进作用更大。

4. 地方政府数据开放程度。理论上,政府数据开放程度更高的地区,能给数据交易平台提供种类更多、潜在价值更大的数据。更高质量的数据在不同制造业企业之间非竞争性共享使用(Jones & Tonetti, 2020),能够更好地促进了先进技术的扩散和溢出,因而,能够更好地增强数据交易平台建设对制造业企业生产率增长的促进作用。由此可以推测,政府数据开放程度更高的地区,数据交易平台建设的促进作用可能更大。关于地方政府数据开放度的衡量,鉴于开放数林综合指数累计分值<sup>①</sup>反映了地方政府数据开放的持续水平,因此,本文采用中国开放数林综合指数累计分值来衡量。当该地区开放数林综合指数累计分值高于样本中位数时,对式(6)中 $G_i$ 取1,而小于其中位数时, $G_i$ 取0。表6列(7)(8)的结果显示,乘积项系数都在1%水平上显著为正(值为0.3151和0.0919),以上结果说明,政府数据开放程度更高的地区(或数据更多来源于政府部门的地区),数据交易平台建设的促进作用更大。

<sup>①</sup>数据来源于历年复旦大学联合国家信息中心数字中国研究院发布的《中国地方政府数据开放报告》;中国开放数林指数自2017年首次发布以来,每半年发布一次;开放数林综合指数累计分值由2017—2020年这4年的年终开放数林综合指数分值累加而成。

## 七、调节机制分析

前文验证了数据交易平台建设对制造业企业生产率增长的促进作用,本部分检验数据产业链关键环节和数据治理的影响,即验证了假说3。为了验证,这里借鉴钱雪松(2019)的做法,构建如下三重差分模型进行检验:

$$Y_{it} = \omega_0 + \omega_1 D_{it} \times X_{it} + \omega_2 D_{it} + \omega_3 X_{it} + \omega_4 (X_{it} \times f(t)) + \tau_i + \mu_i + \varepsilon_{it} \quad (7)$$

式(7)中, $D_{it} \times X_{it}$ 为调节机制变量的乘积项。 $X_{it}$ 为调节机制变量,包括数据产业关键环节的基础设施建设和数据应用场景代理变量,以及数据法治的数据法治建设和数据安全保障代理变量。

### (一)数据产业链关键环节的调节机制分析

1. 数字基础设施的影响。鉴于数字基础设施主要体现在移动终端、网络发展基础设施、5G 基站建设等方面,为了更为准确的衡量这一变量,本文采用城市的移动终端普及率、网络发展基础设施指数、5G 基站覆盖率等三个指标的算术平均值来测算;其中,网络发展基础设施指数的数据来源于《大数据蓝皮书:中国大数据发展报告 No. 5》,其他数据来源于《中国通信年鉴》相关年度。在此基础上,借鉴袁淳等(2021)的做法,将测算的综合指数高于中位数的样本( $U_i$ )设为1,否则设为0。

表7列(1)(2)报告了数字基础设施建设调节机制的估计结果,可以看出,乘积项系数都在1%水平上显著为正(值为0.03161和0.0874),说明数字基础设施建设增强了数据交易平台建设对制造业企业生产率增长的作用效果。这个结论表明,对于北京、上海、杭州等数字基础设施建设较完善的城市来说,数据交易平台建设的作用效果更显著,而对于哈尔滨、长春、太原等数字基础设施建设较落后的城市来说,需要加快以5G网络为基础的通信网络基础设施建设,以增强数据交易平台建设在制造业企业生产率提升中的作用。

2. 数据应用场景的影响。本文用《大数据蓝皮书:中国大数据发展报告 No. 5》中场景应用指数作为衡量大数据应用场景变量<sup>①</sup>。理由是,场景应用指数反映了各地区数据技术的创新成效和场景应用能力,而技术和能力是大数据应用场景范围的重要影响因素,因而能较好地反映大数据应用场景范围大小。若该指数的值高于样本值中位数,则虚拟变量  $Du$  取值为1,否则取0。

表7列(3)(4)的结果显示, $D \times Du$ 的系数都在1%水平上显著为正(值为0.3152和0.0821),说明数据应用场景范围越广,越有助于驱动制造业企业对数据要素的需求,进而增强数据交易平台对制造业企业生产率增长的作用。这个结论表明,对于北京、杭州、广州等大数据应用场景范围较广的城市来说,数据交易平台建设的作用效果更显著;而对于哈尔滨、长春、太原等大数据应用场景范围较窄的城市来说,需要深度挖掘更多行业的数据应用需求,以增强数据交易平台建设在制造业企业生产率提升中的作用。

表7 调节机制估计结果

变量	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	数字基础设施		数据应用场景		数据法治建设		数据安全保障	
$D \times U_i$	0.3199*** (15.5366)	0.0945*** (4.5301)						
$D \times Du$			0.3152*** (16.3848)	0.0821*** (4.0131)				
$D \times Larw$					0.3161*** (16.3237)	0.0874*** (4.3121)		
$D \times Safe$							0.3118*** (16.0991)	0.0806*** (3.9542)

①《大数据蓝皮书:中国大数据发展报告》系列报告是我国首套大数据蓝皮书,为大数据各领域研究发展提供了重要理论借鉴和路径支撑,已成为研究大数据各领域发展水平的重要窗口,因而,该系列报告的大数据发展指数及其相关分项指数都具有权威性。

续表 7

变量	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	数字基础设施		数据应用场景		数据法治建设		数据安全保障	
<i>D</i>	0.2710*** (6.4073)	0.0365 (0.8627)	0.2749*** (4.0957)	0.0890 (1.2909)	0.2609*** (4.0439)	0.0596 (0.8543)	0.3065*** (4.7684)	0.0980 (1.5295)
调节变量	是	是	是	是	是	是	是	是
控制变量	否	是	否	是	否	是	否	是
观测值	9300	9270	9300	9270	9300	9270	9300	9270
<i>R</i> <sup>2</sup>	0.0944	0.3053	0.1004	0.3042	0.1022	0.3047	0.0991	0.3041
时间/企业固定效应	是	是	是	是	是	是	是	是
聚类到企业层面	是	是	是	是	是	是	是	是

## (二)数据治理的调节机制分析

1. 数据法治建设的影响。本文用《大数据蓝皮书:中国大数据发展报告 No. 5》中大数据法治指数来衡量数据法治变量。数据法治指数是对某一地区数据法治状况在不同阶段总体发展水平和发展特点进行的综合描述和价值判断,包括了数据立法、数字司法和数权保护三个方面,较全面地反映了各地区数据法治建设状况。若该指数的数值高于样本中位数,数据法治建设调节变量 *Law* 取值 1;否则,取值 0。

从表 7 列(5)(6)可以看出,*D*×*Law* 的系数都在 1%水平上显著为正(值为 0.3161 和 0.0874),说明数据法治建设越完善越能为数据要素交易营造有序的市场环境,进而可以更好地增强数据交易平台建设对制造业企业生产率增长的作用效果。这个结论表明,对于杭州、广州、贵阳等数据法治建设越完善的城市来说,数据交易平台建设的作用效果更显著,而对于太原、哈尔滨、南宁等数据法治建设相对不足的城市来说,应加快推进数据立法进程,这能够更好地发挥数据交易平台建设在制造业企业生产率增长中的促进作用。

2. 数据安全保障的影响。鉴于《大数据蓝皮书:中国大数据发展报告 No. 5》中大数据安全指数包括了安全制度、安全产业、安全能力、安全生态等方面,能较全面地反映各地区数据安全保障状况。因此,本文采用大数据安全指数来衡量城市数据安全保障状况。如果该指数的数值高于样本中位数,数据安全保障调节机制变量 *Safe* 取值为 1;否则,取值 0。

表 7 列(7)(8)的结果显示,*D*×*Safe* 的系数值都在 1%水平上显著为正(值为 0.3118 和 0.0806),说明数据安全保障水平越高,越能为数据要素交易提供更为安全的市场环境,进而更好地强化了数据交易平台建设对制造业企业生产率增长的作用效果。这个结论表明,对于北京、广州、杭州数据安全保障水平较高的城市来说,数据交易平台建设的作用效果更显著,而对于郑州、哈尔滨、长春等数据安全保障水平较低的城市来说,应加大对数据安全的资金、人员、技术等投入以及安全制度建设,全面提升数据安全防护能力,这有助于增强数据交易平台建设对制造业企业生产率增长的促进作用。

## 八、研究结论和政策启示

本文以数据交易平台建设这一重要举措反映数据要素市场发展,在理论分析数据交易平台建设影响企业全要素生产率的基础上,运用渐进双重差分法,结合中国 2011—2020 年上市公司数据,对理论预期进行了验证。进一步,探讨了数字基础设施、数据应用场景等数据产业链关键环节,以及数据法治建设、数据安全保障等方面数据治理水平,在数据交易平台建设促进制造业全要素生产率增长中的作用。

经验证据支持理论分析的结论,即数据交易平台建设显著地促进了制造业企业全要素生产率增长,这一结论在安慰剂检验、PSM-DID 等一系列稳健性检验后仍然成立。数据交易平台建设的促进作用是通过降低交易成本、提高创新效率与生产效率等机制实现的。数据交易平台建设的促进作用受到数据产业链上下游关键环节的影响,即数字基础设施建设、数据应用场景增加强化了数据交易

平台建设的促进作用;而且,数据法治建设、数据安全保障等方面数据治理水平提高,增强了数据交易平台建设对企业全要素生产率增长的影响。此外,对于规模较大、智能化水平较高、技术密集度较高及政府数据开放度较高地区的企业来说,数据交易平台建设的促进作用更强。

本文的研究结论,从政策层面来说,具有如下重要启示意义:

第一,政府的制造业生产率提升政策措施需要与正在重点推进的数据交易平台建设结合起来,需要将提升制造业全要素生产率作为数据要素市场培育与发展的重要方向,规模较大的企业、工业行业智能化水平较高与技术密集度较高的行业所在地区要加快数据交易平台建设、加快政府数据开放与共享,以提升数据交易平台在提升企业制造业全要素生产率过程中的作用。具体来说,数据交易平台需要整合前期业务积累的海量数据,如工业企业产生的生产材料数据、采购数据、生产流程数据、物流数据以及销售数据等,形成全链条可信数据,加快数据价值的变现。大型国有企业要重视在数据资产化、数据托管与开发利用模式上的探索。工业行业智能化水平较高与技术密集度较高的行业要建立数据产品交易机制,并且探索数据资产化的实现路径,以更好地促进数据交易平台的互联互通性和数据流通共享效率。各地区要建立国家政府数据开放平台,明确对外开放标准和规范,有序推动政府数据开放、增值化开发和利用。

第二,为了更有效地提升制造业全要素生产率,需要推进数据交易平台建设与降低交易成本、提高创新效率、生产效率等政策措施相互协调与配合。具体政策措施包括:首先,建立有利于数据资源要素自由流通的市场分配机制,有效消除当前存在的数据垄断、数据滥用、非法交易等现象。搭建包括数据交易撮合、交易监管、资产定价、争议仲裁在内的全流程数据要素流动平台;明确数据登记、评估、定价、交易跟踪和安全审计机制,建立全社会数据资源质量评估和信用评级体系;整合区块链等新技术,搭建全社会数据授权存证、数据溯源和数据完整性检测平台。其次,提高数据交易平台的技术水平,使其更好的帮助企业利用数据信息来提高产品的创新和设计能力,实现产品的个性化和多样化。最后,建立企业与高校之间的人才交流机制,定期开展数据人才培养,以更好地提高企业数据资源利用能力,实施流程再造来改进企业生产工艺。

第三,数据交易平台建设需要与数据产业链发展、数据治理的政策措施协同与配合,尤其是数据产业链发展和数据治理状况较差的地区,要高度重视推进数据交易平台与数据产业链关键环节、数据治理协同发展的重要意义。具体来说:提高数字基础设施服务的质量和效率,主要是提升网络带宽和网络性能;优化数据中心布局,引导算力升级;协调推进新基建与各个智能应用结合,为智慧场景创新提供应用场所和市场支撑,推动以5G为基础的各种智能场景落地;建立健全数据产权制度,要重点解决对各种类型数据的合理确权,同时,逐步探索建立国家层面数据交易的法律法规和行业标准,支持数据“安全、可信”流通;加大对数字安全的资金、人员、技术等投入;加强数据安全关键技术研究 and 试点应用,以更好地保护数据安全及隐私,营造安全规范的交易市场环境。

值得指出的是,由于资料和数据的局限,本文并未考察数据交易平台建立后的不同类型政策(尤其是数据要素定价政策)对企业生产率增长的影响,因而,探讨数据交易平台建立后具体政策如何影响企业生产率,可能是该领域进一步研究的重要方向。

#### 参考文献:

- 曹亚军,2019:《要素市场扭曲如何影响了资源配置效率:企业加成率分布的视角》,《南开经济研究》第6期。
- 陈诗一 张建鹏 刘朝良,2021:《环境规制、融资约束与企业污染减排——来自排污费标准调整的证据》,《金融研究》第9期。
- 陈舟 郑强 吴智崧,2022:《我国数据交易平台建设的现实困境与破解之道》,《改革》第2期。
- 陈中飞 江康奇,2021:《数字金融发展与企业全要素生产率》,《经济学动态》第10期。
- 戴魁早 刘友金,2020:《市场化改革能推进产业技术进步吗?——中国高技术产业的经验证据》,《金融研究》第2期。
- 范子英 彭飞 刘冲,2016:《政治关联与经济增长——基于卫星灯光数据的研究》,《经济研究》第1期。
- 韩先锋 惠宁 宋文飞,2014:《信息化能提高中国工业部门技术创新效率吗》,《中国工业经济》第12期。
- 何光辉 杨咸月,2012:《融资约束对企业生产率的影响——基于系统GMM方法的国企与民企差异检验》,《数量经济技术经济研究》第5期。

- 黄群慧 余泳泽 张松林,2019:《互联网发展与制造业生产率提升:内在机制与中国经验》,《中国工业经济》第8期。
- 贾智杰 林伯强 温师燕,2023:《碳排放权交易试点与全要素生产率——兼论波特假说、技术溢出与污染天堂》,《经济学动态》第3期。
- 冀雁龙 李金叶 赵华,2022:《数字化基础设施建设与旅游经济增长——基于中介效应与调节效应的机制检验》,《经济问题》第7期。
- 金刚 沈坤荣,2018:《以邻为壑还是以邻为伴?——环境规制执行互动与城市生产率增长》,《管理世界》第12期。
- 李海舰 赵丽,2021:《数据成为生产要素:特征、机制与价值形态演进》,《上海经济研究》第8期。
- 李文莲 夏健明,2013:《基于“大数据”的商业模式创新》,《中国工业经济》第5期。
- 连玉明 张涛 龙荣远 宋希贤,2021:《大数据蓝皮书:中国大数据发展报告 NO.5》,社会科学文献出版社。
- 芦婷婷 祝志勇,2021:《人工智能是否会降低劳动收入份额——基于固定效应模型和面板分位数模型的检验》,《山西财经大学学报》第11期。
- 吕越 陆毅 吴嵩博 王勇,2019:《“一带一路”倡议的对外投资促进效应——基于2005—2016年中国企业绿地投资的双重差分检验》,《经济研究》第9期。
- 刘凤委 李琳 薛云奎,2009:《信任、交易成本与商业信用模式》,《经济研究》第8期。
- 刘莉亚 金正轩 何彦林 朱小能 李明辉,2018:《生产效率驱动的并购——基于中国上市公司微观层面数据的实证研究》,《经济学(季刊)》第4期。
- 罗德明 李晔 史晋川,2012:《要素市场扭曲、资源错置与生产率》,《经济研究》第3期。
- 马光荣,2014:《制度、企业生产率与资源配置效率——基于中国市场化转型的研究》,《财贸经济》第8期。
- 毛其淋 许家云,2015:《市场化转型、就业动态与中国地区生产率增长》,《管理世界》第10期。
- 毛其淋,2019:《人力资本推动中国加工贸易升级了吗?》,《经济研究》第1期。
- 钱雪松 唐英伦 方胜,2019:《担保物权制度改革降低了企业债务融资成本吗?——来自中国〈物权法〉自然实验的经验证据》第7期。
- 孙晓华 王昀,2014:《R&D投资与企业生产率——基于中国工业企业微观数据的PSM分析》,《科研管理》第11期。
- 孙晓华 袁方 郭旭,2020:《市场化改革的选择效应与集聚效应——来自工业企业的经验证据》,《经济科学》第6期。
- 王浩 尚泉西 李金达,2022:《股权激励行权业绩考核如何影响企业全要素生产率》,《湖南科技大学学报(社会科学版)》第5期。
- 王林辉 姜浩 董直庆,2022:《工业智能化会重塑企业地理格局吗》,《中国工业经济》第2期。
- 许宪春 任雪 常子豪,2019:《大数据与绿色发展》,《中国工业经济》第4期。
- 杨汝岱,2015:《中国制造业企业全要素生产率研究》,《经济研究》第2期。
- 姚立杰 周颖,2018:《管理层能力、创新水平与创新效率》,《会计研究》第6期。
- 尹华容 王惠民,2022:《隐私计算的行政法规制》,《湖南科技大学学报(社会科学版)》第6期。
- 袁漳 肖土盛 耿春晓 盛誉,2021:《数字化转型与企业分工:专业化还是纵向一体化》,《中国工业经济》第9期。
- 赵云辉 张哲 冯泰文 陶克涛,2019:《大数据发展、制度环境与政府治理效率》,《管理世界》第11期。
- 祝树金 汤超,2020:《企业上市对出口产品质量升级的影响——基于中国制造业企业的实证研究》,《中国工业经济》第2期。
- Akcigit, U. & Q. Liu(2016), “The role of information in innovation and competition”, *Journal of the European Economic Association* 14(4):828—870.
- Autor, D. H. (2013), “Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing”, *Journal of Labor Economics* 21(1):1—42.
- Baesens, B. et al(2016), “Transformational issues of big data and analytics in networked business”, *MIS Quarterly* 40(4):807—818.
- Beck, T. et al(2010), “Big bad banks? The winners and losers from bank deregulation in the United States”, *Journal of Finance* 65(6):1637—1667.
- Bergemann, D. & A. Bonatti(2019), “The economics of social data: An introduction”, Cowles Foundation Discussion Paper, No. 2171.
- Burchardi, K. B. & T. A. Hassan(2013), “The economic impact of social ties: Evidence from German reunification”, *Quarterly Journal of Economics* 128(3):1219—1271.
- Cantoni, D. et al(2017), “Curriculum and ideology”, *Journal of Political Economy* 125(2):338—392.

- Cengiz, D. et al(2019), “The effect of minimum wages on low-wage jobs”, *Quarterly Journal of Economics* 134(3): 1405—1454.
- Chetty, R. et al(2019), “Salience and taxation: Theory and evidence”, *American Economic Review* 99(4):1145—1177.
- Cong, L. W. et al(2021), “Knowledge accumulation, privacy, and growth in a data economy”, *Management Science* 67(10):6480—6492.
- Cong, L. W. et al(2022), “Endogenous growth under multiple uses of data”, *Journal of Economic Dynamics and Control* 141, 104395.
- Gandhi, A. et al(2020), “On the identification of gross output production functions”, *Journal of Political Economy* 128(8):2973—3016.
- Gardner, J. (2022), “Two-stage differences in differences”, Department of Economics, University of Mississippi, University Working Paper.
- Goodman-Bacon, A. (2021), “Difference-in-differences with variation in treatment timing”, *Journal of Econometrics* 225(2):254—277.
- Hirshleifer, D. et al(2013), “Innovative efficiency and stock returns”, *Journal of Financial Economics* 107(3):632—654.
- Hsieh, C. T. & P. J. Klenow(2009), “Misallocation and manufacturing TFP in China and India”, *Quarterly Journal of Economics* 124(4):1403—1448.
- Cockburn, I. M. et al(2019), “The impact of artificial intelligence on innovation: An exploratory analysis”, in: A. Agrawal et al(eds), *Economics of Artificial Intelligence*, University of Chicago Press.
- Jones, C. I. & C. Tonetti(2020), “Nonrivalry and the economics of data”, *American Economic Review* 110(9):2819—2858.
- Levinsohn, J. & A. Petrin(2003), “Estimating production functions using inputs to control for unobservables”, *Review of Economic Studies* 70(2):317—341.
- Stock, J. H. & M. Yogo(2005), “Testing for weak instruments in linear IV regression”, NBER Technical Working Papers, No. 284.

## How the Construction of Data Trading Platform Affects Firms' TFP?

DAI Kuizao WANG Siman HUANG Zi

(Hunan University of Science and Technology, Xiangtan, China)

**Abstract:** This paper measures the development of the data factor market with the construction of the government participation data trading platform(CDTP). Based on the impact of the construction of the data trading platform on firms' total factor productivity (TFP), this paper makes an empirical verification of the theoretical expectations by using the time-varying DID approach, which combines the construction practice of the data trading platform and the data of A-share listed company in Shanghai and Shenzhen Stock exchanges from 2011 to 2020. It is found that the data factor market construction has significantly promoted the TFP of the manufacturing industry. The results remain robust after dealing with the endogeneity problem, conducting a series of robustness tests including a placebo test, and utilizing difference in differences—propensity score matching approach (such as PSM-DID). The promotion effect of CDTP is mainly through the mechanisms that reduce transaction costs and increase innovation efficiency and productivity. The key links of the data industry chain such as digital infrastructure and data application scenarios have strengthened the promotion of the CDTP. Meanwhile, the impact of the CDTP has been enhanced by the increased level of data governance such as developing data laws and ensuring data security. Furthermore, the promotion of the CDTP is more effective for firms that have larger scale, higher levels of intelligence, higher technology intensity, and higher openness of government data. Moreover, this paper verifies the effect of the CDTP on TFP and provides a beneficial reference for the development of the data factor market and the high-quality development of China's manufacturing industry.

**Keywords:** Factor Market; Data Trading Platform; Total Factor Productivity; Data Industry Chain; Data Governance

(责任编辑:木丰)

(校对:金禾)