

# 政府数据开放的价值创造作用： 企业全要素生产率视角

彭远怀<sup>\*</sup>

**摘要：**作为推进数字经济战略和数字政府建设的重大举措，政府数据开放被广泛认为将创造巨大的经济社会价值，但政府数据开放能否以及如何创造价值仍有待深入论证。基于省级政府开通政府数据开放平台的准自然实验，本文利用双重差分模型检验政府数据开放对企业全要素生产率的影响，以揭示政府数据开放在微观企业层面的价值创造作用。经验证据表明，政府数据开放能显著提高全要素生产率，其主要原因是政府数据开放有助于优化营商环境、增强企业能力、促进高质量创新，且政府数据开放对全要素生产率的提高作用主要存在于年轻企业和非国有企业。基于开放数林指数的检验结果表明，政府数据开放水平越高则全要素生产率越高，而开放数据质量、平台建设质量、政策保障力度是决定全要素生产率提高幅度的重要维度。本文从企业全要素生产率视角揭示了政府数据开放的价值创造作用，为推进政府数据高质量开放提供了重要决策参考。

**关键词：**政府数据开放 全要素生产率 价值创造 数字政府

**中图分类号：**F270 **文献标识码：**A **文章编号：**1000-3894(2023)09-0050-21

**DOI:**10.13653/j.cnki.jqte.20230725.001

## 一、引言

随着数字技术的飞速发展，数据资源在推动经济社会高质量发展中的重要性也日益凸显。特别地，政府履行公共职责时依法制作、采集或获取的资源储量、工商行政处罚、企业信用状况等政府数据，不仅种类丰富和量级庞大，更是领域内的关键性、基础性、权威性数据(郑磊,2018)，被视为国家基础战略资源(朱峥,2020)，对推进数字经济发展和数字政府建设具有重要作用。<sup>①</sup>然而，数据只有被充分利用才能发挥最大效用，但现实情况是，大量政府数据因主管单位不愿或难以向社会公众开放而无法被充分利用(陈尚龙,2016)。为了促进社会公众利用政府数据，世界各国政府近年来积极推动政府数据开放，即免费且无差别地向社会公众主动提供原始政府数据。自国务院2015年发布的《促进大数据发展行动纲要》提出“加快政府数据开放共享”以来，政府数据开放作为推进数字经济战略和数字政府建设的重大举措已上升至国家战略规划。如《中共中央国务院关于构建更

---

<sup>\*</sup> 彭远怀，博士后，暨南大学经济学院，电子邮箱：pengyh@jnu.edu.cn。本文获得国家自然科学基金重点项目(72132002)、中国博士后科学基金(2023M731313)、教育部哲学社会科学研究重大课题攻关项目(22JZD007)的资助。感谢匿名审稿专家的宝贵意见，文责自负。

<sup>①</sup> 政府数据有狭义和广义之分，狭义政府数据指政务数据，而广义政府数据指包含政务数据的所有公共数据。从政策实践和文献研究来看，政府数据开放的对象为广义政府数据，因而本文也是基于广义政府数据展开分析的。

加完善的要素市场化配置体制机制的意见》将“推进政府数据开放共享”作为“加快培育要素市场”的首要措施。此外,《“十四五”规划和2035年远景目标纲要》《“十四五”数字经济发展规划》等文件也特别强调了政府数据开放在推进数字经济发展和数字政府建设中的重要作用。

与实践者积极推动政府数据开放相呼应的是,学术界广泛认为政府数据开放将创造巨大经济社会价值(Zhao和Fan,2021a、2021b)。但现有文献多是沿用公共管理学的研究范式围绕政府数据开放的制度基础、开放程度、决定因素等展开分析(Zuiderwijk和Janssen,2014;段盛华等,2020;韩啸和吴金鹏,2020;朱峥,2020;Zhao等,2022),而政府数据开放如何甚至能否创造价值仍模糊不清(Jetzek等,2019;Magalhaes和Roseira,2020)。特别地,无论是推进数字经济发展还是数字政府建设,政府数据开放的深层次目标均是推动经济高质量发展,而在经济增速放缓和要素成本上升的背景下,提高全要素生产率已成为经济高质量发展的主要动力源泉,但却鲜有文献从全要素生产率视角讨论政府数据开放的价值创造作用。此外,作为市场经济活动的微观主体,企业既是推动经济高质量发展的主要力量,也是被鼓励利用开放政府数据创造价值的主要对象(Wang和Lo,2020),那么政府数据开放又是否以及如何影响企业全要素生产率呢?

政府数据开放的实践形式是职能部门将需要开放的数据以多类常见格式放置于在线政府数据开放平台,以供社会公众自由查询、预览和下载,或是通过数据接口调用。自美国在2009年搭建全球首个国家级政府数据开放平台以来,英国、新加坡、俄罗斯等国也陆续搭建了各自的政府数据开放平台。上海是国内最早探索政府数据开放的地区,其在2012年便已试运行上海政府数据服务网。同年,北京开始测试北京市政府数据资源网。此后,天津市、浙江省、广东省等地也相继开通地方政府数据开放平台<sup>①</sup>,这为利用双重差分模型探究政府数据开放与企业全要素生产率的因果关系提供了较好的应用前提。此外,《“十四五”规划和2035年远景目标纲要》等文件提出“构建统一的国家公共数据开放平台和开发利用端口”。由此可见,利用地方政府数据开放平台考察政府数据开放在微观企业层面的价值创造作用也有助于为国家公共数据开放平台的构建提供经验启示。

基于此,本文以2007~2021年A股非金融企业为样本,利用省级政府开通政府数据开放平台的准自然实验,通过双重差分模型考察政府数据开放对企业全要素生产率的影响及机制,以揭示宏观政府数据开放在微观企业层面的价值创造作用。本文的边际贡献主要体现在以下三方面:

第一,补充了政府数据开放价值创造作用的研究。政府数据开放作为数字时代政府提供的新型公共服务,学术界虽广泛预计将创造巨大社会经济价值,但现存研究多是沿用公共管理学的研究范式就开放论开放,如分析权利基础、评估开放程度、探究决定因素,而政府数据开放如何甚至是否创造价值仍模糊不清(Jetzek等,2019;Magalhaes和Roseira,2020)。本文较早地聚焦于政府数据开放对微观企业主体的影响,补充了政府数据开放价值创造作用的理论和实证研究。

第二,丰富了企业全要素生产率决定因素的研究。数字经济的本质特征是数据驱动,现存研究侧重讨论数字管理能力、数字化转型等衍生决定因素的影响(Müller等,2018;李唐等,2020),但却较少关注数据作为新型生产要素本身的重要性,尤其是缺乏实证分析。政府数据开放的直接作用在于企业可获取数据的范围和量级增加且成本降低,因而考察政府数据开放对企业全要素生产率的影响有助于丰富数字经济下企业全要素生产率决定因素的研究文献,加深对数据经济效益的理解。

第三,深化了异质性数据要素价值创造的研究。政府数据、企业数据、个人数据因法律地位、制

<sup>①</sup> 此处为统称。当前各地主要以“政府数据开放平台”“公共数据开放平台”作为名称后缀,但部分地区也较为特别,如“天津市信息资源统一开放平台”“湖南政务大数据公众门户”“开放广东”等。

度基础、信息含量等的差异而迥然不同,但现有文献分析数据要素的价值创造时多为笼统的价值创造,或多偏向于企业数据的价值创造,如 Ghasemaghaei 和 Calic(2019、2020)。相反,政府数据开放聚焦于开放政务服务、经济建设等领域内的关键性、基础性、权威性数据,分析其在企业层面的经济后果及作用机制是对现存数据要素价值创造研究文献的有益补充。

余文结构安排如下:第二部分阐述制度背景、文献回顾与研究假说;第三部分介绍研究设计;第四部分报告基准回归的检验结果;第五部分报告作用机制的检验结果;第六部分报告基于企业特征异质性、政府数据开放质量及维度的检验结果;第七部分归纳研究结论并提出政策建议。

## 二、制度背景、文献回顾与研究假说

### (一)制度背景

数据在现代经济中的重要作用虽已获得普遍认可,但开放的数据才能被充分利用,从而发挥推动经济高质量发展的关键驱动力作用。其中,政府部门因履行公共职责依法制作、收集或获取的海量政府数据,对加快培育数据要素市场、促进经济转型具有极大价值。正如时任总理李克强在2016年全国推进简政放权放管结合优化服务改革会议所强调的“目前中国信息数据资源80%以上掌握在各级政府部门手里,‘深藏闺中’是极大浪费”。但政府数据开放并不是以保障公民知情权为目标的政府信息公开制度的扩展和延伸,而是政府在数字时代治理改革的重要制度工具,其强调社会公众对政府数据的利用权,因而与政府信息公开存在具有本质差异的制度框架(朱峥,2020;宋烁,2021)。有鉴于此,政府数据开放亦可理解为数字时代政府提供的新型公共服务(朱峥,2020),此间政府扮演数据供应商角色,社会主体扮演数据使用者角色。

《2012年上海市人民政府工作报告》提出,推进政府部门可以公开的经济社会发展数据、各类基础政务数据对社会公众开放。随后,上海开通国内最早的省级政府数据开放平台——上海政府数据服务网,旨在向社会公众免费提供城市建设、经济建设、教育科技等原始数据。特别地,为全面推进政府数据资源开放工作,上海市经济和信息化委员会自2014年起印发《上海市政府数据资源向社会开放年度工作计划》。尽管上海早在2012年便已尝试开放政府数据,但国内真正开始大规模推进政府数据开放是在2015年国务院印发《促进大数据发展行动纲要》之后,这也是政府数据开放首次出现在中央政策文件。该纲要提出“通过政务数据公开共享,引导企业、行业协会、科研机构、社会组织等主动采集并开放数据”。此后,《关于全面推进政务公开工作的意见》等也提出要支持鼓励社会力量充分开发利用政府数据资源,并明确了政府数据开放在促进信息资源规模化创新应用、培育新的经济增长点、推动国家治理体系和治理能力现代化中的重要作用。

整体来看,各地政府数据开放平台的数据种类基本类似,大致可分为生产资料、产品市场、政务服务类数据。其中,生产资料数据包括地理、能源、资源勘探等;产品市场数据包括人口、消费、行业等;政务服务数据包括财政、司法、信用服务等。以上海公共数据开放平台为例,该平台目前已开放5360个数据集,合计20亿余条数据。数据供给方面:从数据领域来看,城市建设有698个数据集,民生服务有633个数据集,经济建设有603个数据集;从资源类型来看,接口型数据集有2259个,产品型数据集有2230个;从数据提供单位来看,市场监管局提供了226个数据集,市交通委提供了208个数据集,市商务委提供了147个数据集。数据利用方面:该平台已被访问557万余人次,数据下载153万余次,需求请求2万余次,其中科研和商业是最多的需求请求来源。此外,该平台还与银行机构合作开放了普惠金融行业应用——中小微企业信息,包括基本信息、处罚信息、经营状况等数据,这些数据已被调用2609万余次。



## (二)文献回顾

关于政府数据开放的研究主要是讨论政府数据开放的制度基础、开放程度、决定因素等。首先,制度基础方面,政府数据开放的外在表现形式也是将掌握的数据资源公之于众,故而早期文献沿用此逻辑认为政府数据开放与政府信息公开均以知情权为权利基础(陈尚龙,2016)。但无论是制度目标,还是作用对象,抑或是实施主体和方式,政府数据开放与政府信息公开均存在显著差异,因而政府数据开放非政府信息公开的扩展和延伸(朱峥,2020;宋烁,2021)。其次,开放程度评估方面,现有文献从不同角度开发了诸多评估模型,但各模型均是围绕政府、数据、公众维度的因素构建。如Zuiderwijk和Janssen(2014)基于环境情景、政策内容、公共价值等构建开放程度的评估模型;Donker和Loenen(2017)基于开放数据的供应质量、治理及用户特征等构建开放程度的评估模型。最后,决定因素方面,现有文献分析的技术因素包括数据存储获取与使用方式、通信技术基础设施、技术管理能力等(韩啸和吴金鹏,2020;Zhao和Fan,2021a、2021b);组织决定因素包括领导者参与、文化特征、金融资源等(Jetzek等,2019;Zhao和Fan,2021a、2021b);环境因素包括相关法律法规政策、政府竞争、媒体压力等(Wang和Lo,2020;Zhao等,2022)。

关于数据要素价值创造作用的研究主要围绕数字化转型、数据要素利用等展开。首先,数字化转型方面,现有文献指出数字化转型既会强化创新能力、优化人力资本结构、促进专业化分工等以提高企业全要素生产率(袁淳等,2021;赵宸宇等,2021),也会通过改善信息不对称并强化市场正面预期、促进创新等改善企业的股票市场绩效(吴非等,2021;董松柯等,2023)。部分文献还发现数字化转型对降低交易费用、缓解融资约束等有积极作用,以促进企业拓宽供应链地理分布、设立异地子公司、提升市场势力等(唐浩丹等,2022;李万利等,2023;王康等,2023)。其次,数据要素利用方面,数据速度、多样性、准确性等能通过增强洞察力促进企业创新,但仅是增强描述性和预测性洞察力,且数据量不会产生显著影响(Ghasemaghaei和Calic,2019、2020);而利用大数据技术分析则有助于促进价值创造以提高经营绩效(Chen等,2015;Wamba等,2017)。蔡继明等(2022、2023)将数据要素纳入广义价值论框架的分析指出,数据初始存量规模越大越有助于形成较高的绝对生产力,而数据要素还有助于提高其他产品的生产率。类似地,Cong等(2021)指出,消费数据在知识积累中不仅有助于生产新产品和服务,还有助于研发和知识创造。进一步,徐翔和赵墨非(2020)分析认为,数据资本与信息通信技术结合后,既能直接作为生产要素促进经济增长,也能通过改善生产要素配置间接提高生产率。

归纳现有文献可知其主要存在如下不足:首先,现有文献虽围绕政府数据开放的制度基础、开放程度、决定因素等展开了较为丰富的讨论,但政府数据开放是否成功的衡量标准并非是否开放,而是社会主体是否利用开放数据创造价值以及价值创造量(Nikiforova和McBride,2021),因而关注政府数据开放的价值创造作用可能更为重要(Conradie和Choenni,2014;Jetzek等,2019)。其次,数字化转型固然是数字经济下的重要企业行为特征,但其终归需要依赖数据实现,而现有文献虽实证分析了数字化转型对企业全要素生产率的影响,但却难以直接揭示数据要素本身的作用。最后,现有文献讨论数据要素的价值创造作用时多隐性假定数据要素同质,但不同类别的数据要素可能存在极大差异,如Cong等(2021)所提及的消费数据属于企业数据,是企业与客户交易产生的数据,其与政府数据的差异可能影响数据要素的价值创造作用。基于此,本文将考察政府数据开放对企业全要素生产率的影响,以揭示其在微观企业层面的价值创造作用。

## (三)研究假说

数字经济已然成为落实国家重大战略的关键力量,但碎片化存储、共享机制不畅通等却阻碍着

数据要素的自由流动。政府数据开放旨在向社会主体主动、免费且无差别地提供政府行使公共权力时获得的原始数据,以激励其利用开放政府数据创造更大的社会经济价值(韩啸和吴金鹏,2020;胡业飞和刘梦露,2021),以及通过数据分析和公私部门信息共享促进政策创新(Zuiderwijk 和 Janssen,2014;Park 和 Gil-Garcia,2022)。而相关方组成的整体则被视为以社会技术现象为特征的生态系统(Wang 和 Lo,2020),企业在生态中可以获取无版权的标准数据,或辅助管理层决策,或用于新产品和服务的创造。以著名房地产估价平台 Zillow 为例,其收集了包括财政税收、土地价格、地理信息、待售房屋等在内的诸多开放政府数据,并利用基于机器学习技术的预测为房屋买卖双方提供精准建议(Jetzek 等,2019)。此外,开放数据研究院 2015 年发布的《开放数据意味着商业:英国跨部门和地区的创新》报告也显示,不同规模、年龄、行业、地区的企业都在广泛利用政府数据开放,以期创造更多的社会、环境、经济效益。从这些案例不难看出,政府数据开放的确可能对企业产生显著影响,但其究竟如何影响企业全要素生产率还有待论证。

新古典经济学理论将生产活动中的交易行为设定为理想状态,因而此时生产活动不存在投入和产出外的任何成本。但企业在现实中的交易行为必然广泛面临与公共制度、市场主体的摩擦。其中,与公共制度的摩擦主要体现在繁杂的行政审批、有偏的司法保护、非市场化的资源配给等制度障碍(宋渊洋和黄礼伟,2014);而与市场主体的摩擦主要是信息不对称造成的管理者有限理性、交易对象机会主义、市场环境不确定性等。此时,企业需要支付额外的交易费用以保证生产活动顺利进行,但这无疑会挤占生产资源,从而抑制全要素生产率提高(夏杰长和刘诚,2017;钱雪松等,2018)。在本文的研究情景中,政府数据开放影响企业全要素生产率的理论逻辑如下。

从与公共制度的摩擦来看,政府数据开放有助于优化营商环境,减少非公平竞争、政府干预、政策低效造成的生产资源挤占,从而提高全要素生产率。更高的透明度和更强的问责机制有利于营造公平的市场竞争环境并抑制政府干预企业活动的行为,如更高效的行政审批、更公正的司法保护、更市场化的资源配置。政府数据开放涉及的政府数据不仅种类丰富、量级庞大,而且是能够充分传递真实信息的原始数据,这在客观上能极大地提高政府运行的透明度(Conradie 和 Choenni,2014;Park 和 Gil-Garcia,2022)。与此同时,企业基于开放政府数据不仅能更有效地监督政府决策结果,也能更有效地监督政府决策过程,因而政府数据开放能强化问责机制(Zuiderwijk 和 Janssen,2014)。特别地,企业对开放政府数据的利用反馈将形成双向互动治理模式(段盛华等,2020;宋烁,2021),继而有利于促进政务服务体制、市场运行机制、资源配置措施等公共政策创新(Park 和 Gil-Garcia,2022),从而改善政策效率,如提高政策执行力度和有效性、抑制政策频繁变更。

从与市场主体的摩擦来看,政府数据开放有助于增强企业能力,减少管理能力、营运能力、融资能力不足造成的生产资源挤占,从而提高全要素生产率。对企业而言,政府数据开放能提高外部数据的可获得性与多样性,而其对丰富政府数据的分析将减少信息不对称以增强管理能力,如克服企业管理者的有限理性、限制交易对象的机会主义行为(李唐等,2020;Farboodi 和 Veldkamp,2021)。与此同时,开放领域内关键性、基础性、权威性政府数据能提高企业对市场需求、客户价值、供应商风险的评估效率(李唐等,2020;Magalhaes 和 Roseira,2020),增强营运能力,从而促进合作伙伴寻找、工厂和仓库选址、经营目标优化等(Ghasemaghahi 和 Calic,2019、2020),而这些都是全要素生产率提高的关键表现。不仅如此,管理能力和营运能力增强也会随之增加企业在融资活动中的谈判筹码,加之金融机构同样可以通过分析客户信用数据、工商行政处罚等开放政府数据提高对信贷风险评估的准确性,因而政府数据开放也能有效增强企业的融资能力。

此外,政府数据开放还有助于促进高质量创新,从而提高全要素生产率。全要素生产率的提高

虽很大部分源自创新,但创新不仅面临较高的不确定性,还需要强力的资源承诺和高昂的资本投入(Bhattacharya等,2017),尤其是高质量创新。政府数据开放能缓解企业与公共制度、市场主体的摩擦,减少对生产资源的挤占,这便有利于促使企业将更多资源投入高质量创新。特别地,内源融资的有限性使得企业进行高质量创新往往需要借助外部资金,但高质量创新却因存在更高的不透明性和不可核实性而难以获得外部资金支持(Aboody和Lev,2000)。在政府数据开放缓解企业与市场主体的摩擦中,融资能力的增强既有利于降低创新融资成本,也有利于扩大融资规模,以支持企业开展更多高质量创新。与此同时,政府数据开放也包括开放产品市场数据,这意味着企业对开放政府数据的分析可以更为快速地掌握市场情况,从而捕捉高质量创新机会。

本文据此提出研究假说:其他条件不变时,政府数据开放会显著提高企业全要素生产率。

### 三、研究设计

#### (一)模型设定与变量定义

理论上,地方政府数据开放时间和内容基本是由地方政府决定,不会受到辖区企业的影响,而现有文献也的确未曾发现企业层面的因素会显著影响政府数据开放。从实践情况来看,上海市和北京市虽最先开通,但经济较发达的天津市、江苏省反而落后于经济欠发达的贵州省、江西省。由此可见,地方政府数据开放相对企业全要素生产率具有较强的外生性和随机性。因此,本文将通过双重差分模型考察政府数据开放对企业全要素生产率的影响。具体基础双重差分模型设定为:

$$TFP_{i,t+1} = \beta_0 + \beta_1 OPEN_{it} + \beta_2 X_{it} + \mu_t + \nu_i + \varepsilon_{i,t+1} \quad (1)$$

其中, $i$ 和 $t$ 分别代表企业和年份; $TFP$ 代表企业全要素生产率; $OPEN$ 代表政府数据开放,其系数 $\beta_1$ 为政府数据开放对企业全要素生产率的净影响, $\beta_1$ 显著为正则表明政府数据开放能提高企业全要素生产率; $X$ 代表可能影响企业全要素生产率的财务特征、治理特征、地区经济特征; $\mu$ 和 $\nu$ 分别代表年份固定效应和企业固定效应; $\varepsilon$ 代表回归残差。考虑到企业全要素生产率变化相对政府数据开放可能存在明显的滞后,因而对其采用 $t+1$ 期的数据计量。特别地,为了缓解残差序列相关和异方差造成的估计偏差,所有回归标准误均会进行地区层面的聚类调整。

各变量的具体测度方式如下:

第一,关于自变量政府数据开放( $OPEN$ )的衡量。本文的主要回归均是通过双重差分模型实现,因而 $OPEN$ 将被定义为虚拟变量。即若样本处于办公地(省、自治区、直辖市)开通政府数据开放平台之后,则对 $OPEN$ 取值1,否则对 $OPEN$ 取值0。其中,上海市和北京市是2012年开通,浙江省是2015年开通,广东省和贵州省是2016年开通,河南省、江西省、山东省、陕西省、宁夏回族自治区是2018年开通,天津市、福建省、江苏省、四川省、海南省、青海省、新疆维吾尔自治区是2019年开通,重庆市、湖北省、湖南省、广西壮族自治区是2020年开通,安徽省、河北省、甘肃省是2021年开通。鉴于虚拟变量所包含的信息不如连续变量丰富,因而本文在拓展分析中将利用复旦大学数字与移动治理实验室构建的中国开放数林指数进行补充检验。

第二,关于因变量全要素生产率( $TFP$ )的衡量。借鉴刘莉亚等(2018)、胡海峰等(2020)、宋敏等(2021)、吴敏等(2022)的做法,同时以现有文献使用最为广泛的Levinsohn和Petrin(2003)、Olley和Pakes(1996)提供的方法估算,分别记为 $TFPLP$ 、 $TFPOP$ ,其优点是能较好地克服联立性、同时性、选择性偏误造成的估计偏差。具体估算模型设定为:

$$Y_{it} = \alpha_0 + \alpha_1 K_{it} + \alpha_2 L_{it} + \alpha_3 M_{it} + \alpha_4 AGE_{it} + \delta_{it} \quad (2)$$

其中, $i$ 和 $t$ 分别代表企业和年份; $Y$ 代表产出,以营业收入的自然对数衡量; $K$ 代表资本投入,以



固定资产账面价值的自然对数衡量; $L$ 代表劳动投入,以员工数量的自然对数衡量; $M$ 代表中间投入,采用 Levinsohn 和 Petrin(2003)估算时以营业收入与增加值之差的自然对数衡量,其中增加值为折旧、劳动者报酬、生产税净额和营业盈余四项之和(刘莉亚等,2018),采用 Olley 和 Pakes(1996)估算时以投资金额的自然对数衡量,投资金额为购建固定资产、无形资产和其他长期资产支付的现金(胡海峰等,2020); $AGE$ 代表企业年龄,以成立年数的自然对数衡量; $\delta$ 代表全要素生产率的对数形式信息(刘莉亚等,2018)。

第三,关于控制变量( $X$ )的选取及衡量。借鉴胡海峰等(2020)、宋敏等(2021)、吴敏等(2022)等的做法,主要控制可能影响企业全要素生产率的财务特征、治理特征、地区经济特征。其中,财务特征包括资产报酬率( $ROA$ ,息税前利润与资产总额之比)、现金比率( $CASH$ ,现金及现金等价物余额与资产总额之比)、资产负债率( $LEV$ ,负债总额与资产总额之比)、营收增长率( $SGR$ ,营业收入增长率)、市账率( $MTB$ ,股东权益市值与账面价值之比)、资产规模( $SIZE$ ,资产总额的自然对数);治理特征包括机构持股比例( $INST$ ,机构持股数量与总股本数量之比)、设立年龄( $AGE$ ,成立年限的自然对数);地区经济特征包括经济水平( $GDP$ ,人均生产总值的自然对数)、教育水平( $EDU$ ,每百万常住人口拥有的高等学校数量)、人口规模( $POP$ ,常住人口数量的自然对数)。

(二)数据来源与样本选择

本文选取 2007~2021 年 A 股企业为初始研究样本,并借鉴钱雪松等(2018)、胡海峰等(2020)、宋敏等(2021)的做法,剔除金融保险行业的样本和基础回归涉及变量存在数据缺失的样本。文中所使用的地方政府数据开放相关数据取自复旦大学数字与移动治理实验室在 2017~2021 年发布的《中国地方政府数据开放报告》,其中政府数据开放平台的开通时间利用各地政府数据开放平台信息及百度搜索予以确认和补充;计算全要素生产率所需营业收入、购建固定资产支出的现金、固定资产账面价值、员工数量以及其余企业财务和治理数据均取自 Wind 金融终端;地区生产总值、地区常住人口、地区高等学校数量取自 Wind 金融终端和中国研究数据服务平台;企业专利申请数量、分类号、引用信息等专利相关数据取自 CnOpenData;此外,中央和地方产业政策数据取自中国研究数据服务平台,该数据是由平台从国家和地方政府颁布的产业规划中整理获得。

(三)主要变量描述性统计

为了规避极端值对实证分析结果的干扰,本文所有连续变量均会进行双边 1% 水平的缩尾处理。表 1 列示了全要素生产率和政府数据开放两个变量的描述性统计结果。<sup>①</sup> $TFPLP$ 的均值为 12.355,中位数为 12.297; $TFPOP$ 的均值为 10.026,中位数为 10.070,表明样本企业的全要素生产率不存在明显的左偏或者右偏,且估计值也与刘莉亚等(2018)、胡海峰等(2020)的估算结果较为接近; $OPEN$ 的均值为 0.375,表明约 37.5% 的样本受到政府数据开放的影响,因而样本具有较强代表性。

表 1 主要变量描述性统计

变量	均值	标准差	25 分位数	中位数	75 分位数	样本量
$TFPLP$	12.355	0.992	11.701	12.297	12.941	36725
$TFPOP$	10.026	1.479	9.065	10.070	11.054	36725
$OPEN$	0.375	0.484	0.000	0.000	1.000	36725

① 由于篇幅限制,其余变量的描述性统计结果参见《数量经济技术经济研究》网站附录。

## (四)政府数据开放外生性

本文在模型设定中从理论和实践层面阐述了政府数据开放外生的合理性。为了进一步排除政府数据开放相对企业全要素生产率的非外生性担忧,本文将借鉴 Beck 等(2010)的做法利用 Weibull Hazard 模型直接验证政府数据开放相对企业全要素生产率是否外生,即考察企业全要素生产率是否会影响政府数据开放时间。若结果显示企业全要素生产率不会影响政府数据开放时间,则表明政府数据开放相对企业全要素生产率是外生的。具体检验模型设定为:

$$LY_{dt} = \lambda_0 + \lambda_1 TFP_{dt} + \lambda_k X_{dt} + \mu_t + \omega_d + \tau_{dt} \quad (3)$$

其中, $d$ 和 $t$ 分别代表地区和年份; $LY$ 代表政府数据开放的预期时间,以政府数据开放年份与观测年份之差的自然对数衡量,若某地区在样本期内始终未开通政府数据开放平台,则其预期时间为样本结束年份与观测年份之差的自然对数; $TFP$ 代表地区企业的平均全要素生产率; $X$ 代表系列地区层面的控制变量,包括经济水平( $GDP$ ,人均地区国内生产总值的自然对数)、教育水平( $EDU$ ,每百万常住人口拥有的高等学校数量)、人口规模( $POP$ ,常住人口数量的自然对数); $\mu$ 和 $\omega$ 分别代表年份固定效应和地区固定效应;此处的回归标准误仍在地区层面进行聚类调整。

表2列示了地区企业全要素生产率均值与政府数据开放预期时间的回归结果。第(1)、(2)列结果显示,在不加任何控制变量时, $TFP_{MLP}$ 、 $TFP_{MOP}$ 与 $LY$ 的回归系数均不显著,表明地区企业全要素生产率不会影响政府数据开放时间。第(3)、(4)列结果显示,加入控制变量后, $TFP_{MLP}$ 、 $TFP_{MOP}$ 与 $LY$ 的回归系数仍不显著,这与单变量回归的结果相符; $GDP$ 、 $EDU$ 、 $POP$ 与 $LY$ 的回归系数不显著,表明地区经济水平、地区教育水平、地区人口规模也不会影响政府数据开放时间。以上检验结果再次支持了政府数据开放的相对企业全要素生产变化外生的假定,亦即较好地满足了外生性假设。

表2 政府数据开放的外生性

变量	单变量回归		全变量回归	
	(1) $LY$	(2) $LY$	(3) $LY$	(4) $LY$
$TFP_{MLP}$	-0.106 (0.375)		0.016 (0.388)	
$TFP_{MOP}$		-0.071 (0.371)		0.062 (0.410)
$GDP$			-0.659 (0.567)	-0.676 (0.533)
$EDU$			0.546 (0.355)	0.562 (0.335)
$POP$			-1.717 (1.574)	-1.682 (1.743)
地区固定	是	是	是	是
年份固定	是	是	是	是
样本量	379	379	379	379
组间 $R^2$ 值	0.001	0.001	0.124	0.124

注: \*、\*\*、\*\*\*分别代表在10%、5%、1%水平上显著;括号内数值为地区聚类标准误;常数项已省略。



#### 四、基准回归分析

##### (一)基准回归检验

表3列示了政府数据开放与企业全要素生产率的回归结果。第(1)、(2)列为仅包括年份固定效应和企业固定效应的单变量回归结果,无论是 $TFPLP$ 还是 $TFPOP$ 作为因变量, $OPEN$ 的回归系数均在1%水平上显著为正,表明企业在办公地开通政府数据开放平台后的全要素生产率显著高于预期水平。该结果初步支持了政府数据开放在提高企业全要素生产率方面的积极作用,这也意味着政府数据开放有助于优化营商环境、增强企业能力、促进高质量创新的理论逻辑可能成立。

考虑到财务状况、治理特征、地区经济特征也可能影响企业全要素生产率,因而本文在前述单变量回归的基础上控制了资产报酬率、现金比率、资产负债率、营收增长率、市账率、资产规模、机构持股比例、企业成立年龄、地区经济水平、地区教育水平、地区人口规模。第(3)、(4)列全变量回归结果显示, $OPEN$ 与 $TFPLP$ 、 $TFPOP$ 的回归系数分别为0.044、0.033,均在至少5%水平上显著;且各控制变量的回归系数符号及显著性也与现有文献(钱雪松等,2018;胡海峰等,2020;宋敏等,2021;赵宸宇等,2021)的检验结果基本相似。由此可见,政府数据开放能显著提高企业全要素生产率的假说获得充分的证据支持。就经济显著性而言,由于 $TFPLP$ 、 $TFPOP$ 为对数全要素生产率,因而以 $TFPLP$ 作为因变量时全要素生产率较预期提高了约4.50%( $e^{0.044}-1$ ),以 $TFPOP$ 作为因变量时全要素生产率较预期提高了约3.36%( $e^{0.033}-1$ ),具有较强的经济意义。

此外,不同行业的发展趋势存在差异,如行业生命周期引起的兴衰更替,而这可能增加组间全要素生产率的固有差异,因而本文在全变量回归的基础上增加了行业趋势控制。第(5)、(6)列结果显示, $OPEN$ 与 $TFPLP$ 、 $TFPOP$ 的回归系数在至少5%水平上显著为正,表明基准回归结论并非源自行业趋势存在差异。进一步,部分地方政策也可能影响企业全要素生产率,从而增加回归估计偏误。为了消除地方政策混杂效应对研究结论的影响,本文在全变量回归的基础上增加了地区趋势控制。第(7)、(8)列结果显示, $OPEN$ 与 $TFPLP$ 、 $TFPOP$ 的回归系数在至少5%水平上显著为正,表明地区趋势差异也不能解释基准回归结论。以上检验结果意味着,政府数据开放对企业全要素生产率的提高作用并非行业和地区趋势差异导致,这有力地佐证了基准回归结论的可靠性。

表3 政府数据开放对企业全要素生产率的影响

变量	单变量回归		全变量回归		控制行业趋势		控制地区趋势	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	$TFPLP$	$TFPOP$	$TFPLP$	$TFPOP$	$TFPLP$	$TFPOP$	$TFPLP$	$TFPOP$
$OPEN$	0.087*** (0.022)	0.071*** (0.023)	0.044*** (0.012)	0.033** (0.013)	0.037*** (0.013)	0.031** (0.014)	0.031*** (0.010)	0.029** (0.011)
控制变量	否	否	是	是	是	是	是	是
企业固定	是	是	是	是	是	是	是	是
年份固定	是	是	是	是	是	是	是	是
行业趋势	否	否	否	否	是	是	否	否
地区趋势	否	否	否	否	否	否	是	是
样本量	36725	36725	36725	36725	36725	36725	36725	36725
组间 $R^2$ 值	0.003	0.002	0.258	0.207	0.256	0.214	0.248	0.197

注:同表2。

## (二)平行趋势检验

依据平行趋势假设,如果政府数据开放事件没有发生,那么开放地区和未开放地区企业的全要素生产率会呈现相同变化趋势。然而,政府数据开放并没有被取消的案例,致使无法通过反事实法直接验证。对此,本文借鉴钱雪松等(2018)的做法通过动态双重差分模型考察企业全要素生产率在政府数据开放前的组间变化趋势。若政府数据开放地区和未开放地区企业在开放前的全要素生产率变化趋势不存在显著差异,则表明平行趋势假设满足。动态双重差分模型设定为:

$$TFP_{it} = \eta_0 + \eta_m EVENT_{it}^n + \eta_j X_{it} + \mu_i + \nu_t + \varepsilon_{it} \quad (4)$$

其中, $i$ 和 $t$ 分别代表企业和年份; $EVENT_{it}^n$ 是样本所处年份相对于政府数据开放年份的时间距离变量, $n$ 取值0时代表开放当年, $n$ 取负(正)值时代表开放前(后);其余变量的定义与基础双重差分模型相同。为了规避时间距离变量间的多重共线性,此处以政府数据开放前1期作为基期,即在回归中省略 $EVENT^{-1}$ 。若平行趋势假设成立,则在政府数据开放前时间距离变量的回归系数均不会显著偏离零轴,即开放地区和未开放地区企业的全要素生产率变化趋势同步。

图1展示了90%置信区间下政府数据开放对企业全要素生产率的动态作用情况。其中,横轴为样本所处年份相对政府数据开放年份的时间距离,即模型中 $n$ 的取值;纵轴为 $EVENT$ 的回归系数,亦可理解为处理组和对照组样本的全要素生产率差异。无论是以 $TFPLP$ 还是 $TFPOP$ 作为因变量, $EVENT$ 的回归系数在政府数据开放前均在零轴附近波动且未出现显著情况,而仅仅是在政府数据开放后才出现正向显著偏离零轴,表明双重差分模型的应用基本满足平行趋势假设。与此同时, $EVENT$ 的回归系数稳定持续到第3年时仍显著异于零,表明政府数据开放的全要素生产率提高作用存在较长期的影响。此外, $EVENT$ 的回归系数在政府数据开放后呈现先增大后减小的趋势,表明政府数据开放对企业全要素生产率的影响存在动态变化。总之,图1结果充分说明政府数据开放提高企业全要素生产率的因果关系并非事前趋势导致。

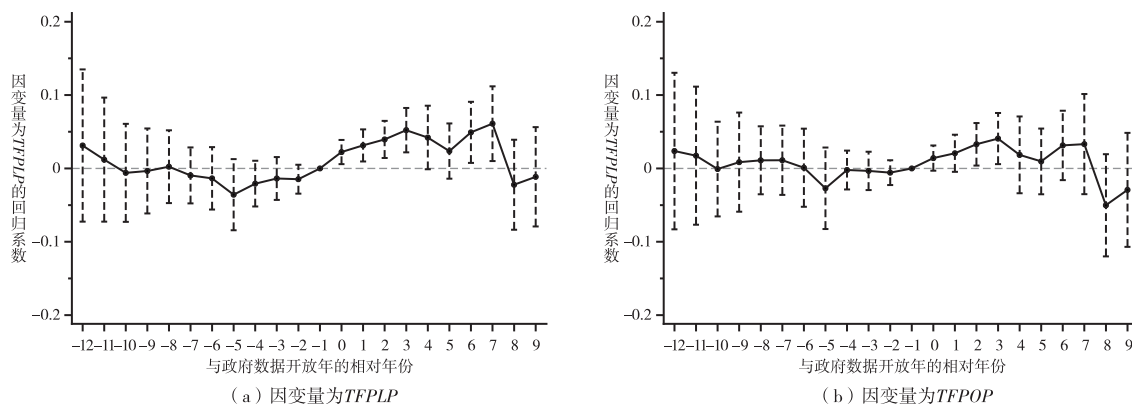


图1 作用效果的动态变化

## (三)安慰剂检验

遗漏地区和时间层面关键因素可能得出政府数据开放提高企业全要素生产率的有偏结论。考虑到地区和时间层面诸多因素难以直接测度,因而此处通过随机设定政府数据开放的地区和年份进行安慰剂测试。鉴于最终样本有20个地区开通了政府数据开放平台,因而本文首先将样本地区随机排序后抽取前20个地区设定为伪政府数据开放地区;其次对地区一年份样本随机排序后抽取

每个地区层面的第1个地区一年份样本;再次将随机抽取的地区和地区一年份样本合并且仅保留伪政府数据开放地区对应的地区一年份样本;然后构建伪政府数据开放(*FAKE OPEN*)变量并进行回归;最后记录*FAKE OPEN*的回归系数和t统计量。此处的安慰剂检验将重复3000次。若基准回归结论并非地区和时间层面遗漏因素导致,则回归系数和t统计量均是趋于均值为零的正态分布。

图2展示了*FAKE OPEN*回归系数和t统计量的核密度函数分布。图(a)显示,*FAKE OPEN*与*TFPLP*、*TFPOP*的回归系数均基本呈现均值为0的正态分布,实际均值分别为0.000、0.000,远小于表3第(3)、(4)列中*OPEN*的回归系数数值(0.044、0.033)。图(b)显示,*FAKE OPEN*与*TFPLP*、*TFPOP*的t统计量也基本呈现均值为0的正态分布,实际均值分别为-0.014、-0.016,远小于表3第(3)、(4)列中*OPEN*的t统计量(3.667、2.583)。以上检验结果表明,无论是从回归系数还是从t统计量来看,随机设定的伪政府数据开放均不会显著影响企业全要素生产率。由此可见,政府数据开放对企业全要素生产率的提高作用并非地区和时间层面遗漏因素导致。

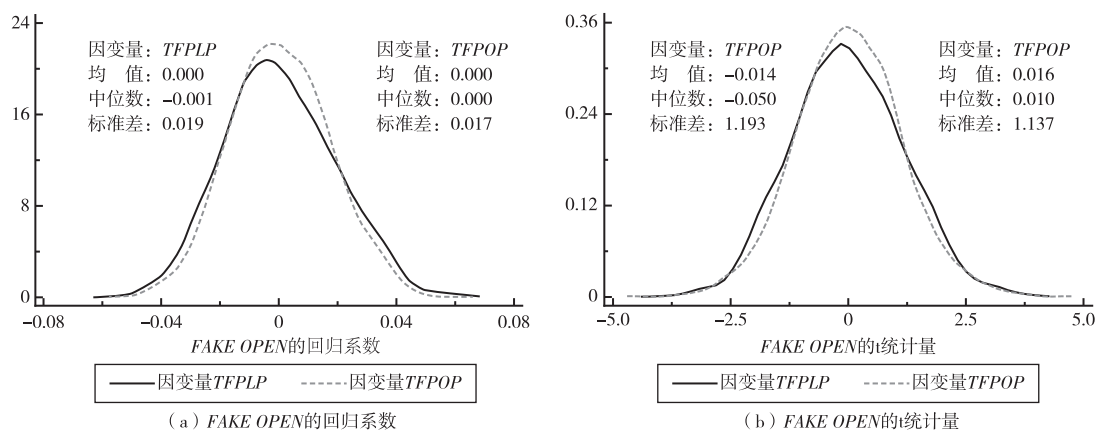


图2 安慰剂检验

注:为了更直观地判断显著性,此处选择展示t统计量而非标准误;t统计量等于回归系数除以标准误。

#### (四)倾向得分匹配回归

政府数据开放虽相对企业全要素生产率充分外生,但样本特征在组间不可比引发的选择性偏差也可能造成政府数据开放提高企业全要素生产率的现象。为了消除选择性偏差的干扰,本文将利用倾向得分匹配控制样本的组间特征差异。具体而言,本文首先将样本期内办公地开放政府数据的企业与办公地始终未开放政府数据的企业在政府数据开放前1年以基准回归中的控制变量为匹配变量进行1:1的分行业、分年份匹配;然后通过前述基础双重差分模型对匹配成功企业对应的样本进行回归。基于倾向得分匹配样本的检验结果显示,无论是以*TFPLP*还是以*TFPOP*作为因变量,*OPEN*的回归系数均在至少5%水平上显著为正,这表明基准回归结果并非选择性偏差导致。<sup>①</sup>

#### (五)其他稳健性检验

本文还进行了双重聚类标准误、剔除2012年以前样本、控制产业政策支持的控制变量,结果显示政府数据开放对企业全要素生产率的提高作用未发生显著改变。

<sup>①</sup> 由于篇幅限制,详细检验结果参见《数量经济技术经济研究》网站附录,下同。

第一,采用双重聚类标准误。鉴于样本时间跨度较长,基准回归不仅可能受到地区截面相关因素的影响,也可能受到时间序列自相关问题的影响,因而本文对估计标准误进行了年份和地区层面的双重聚类。检验结果显示,*OPEN*与*TFPLP*、*TFPOP*的回归系数均在至少5%水平上显著为正,从而排除了标准误聚类层级差异对基准回归结论的影响。

第二,剔除2012年以前样本。自2012年以后的营改增等系列重大政策可能增加与此前样本的全要素生产率差异。与此同时,中国对政府数据开放的探索源自2012年上海市试运行上海政府数据服务网,但此前政府数据开放鲜有讨论。对此,本文剔除了2012年以前的样本。检验结果显示,*OPEN*与*TFPLP*、*TFPOP*的回归系数仍在5%水平上显著为正,再次证明基准回归结论稳健可靠。

第三,控制产业政策支持。考虑到产业政策也可能影响企业全要素生产率,因而本文依据中央(地方)产业规划内容增加企业是否属于中央(地方)产业政策重点鼓励发展行业的控制变量。检验结果显示,无论是控制中央还是地方产业政策支持,*OPEN*与*TFPLP*、*TFPOP*的回归系数均在至少5%水平上显著为正,表明基准回归结论并非源自遗漏控制产业政策支持。

五、作用机制分析

政府数据开放对企业全要素生产率的提高作用虽已充分证实,但其作用机制是否为优化营商环境、增强企业能力、促进高质量创新还有待深入检验。对此,本文接下来将依次展开分析。

(一)优化营商环境机制

政府数据开放释放了海量与政府运行密切相关的原始数据(郑磊,2015;宋烁,2021),这为企业更加全面地了解政府运行情况、监督政府决策过程和结果提供了契机,并形成双向互动治理模式。而政府数据开放对营商环境的优化作用可能主要体现在政府透明度提高、问责机制增强、公共政策创新导致更公平的竞争、更少的政府干预、更高的政策效率。基于此,此处将通过检验政府数据开放对公平竞争、政府干预、政策效率的影响验证优化营商环境机制。其中,公平竞争以营业收入的地区市场集中度(*MKT*)衡量(牛志伟等,2023),集中度越小则公平竞争越强;政府干预以财政赤字水平(*DEFT*,财政净支出与生产总值之比)衡量(马连福和曹春方,2011),赤字水平越高则政府干预越强;政策效率以Yu等(2021)开发的地方经济政策不确定性指数标准化后的自然对数(*EPU*)衡

表4 政府数据开放与优化营商环境

变量	公平竞争	政府干预	政策效率
	(1)	(2)	(3)
	<i>MKT</i>	<i>DEFT</i>	<i>EPU</i>
<i>OPEN</i>	-0.037*	-0.087***	-0.365*
	(0.018)	(0.031)	(0.205)
控制变量	是	是	是
企业固定	是	是	是
年份固定	是	是	是
样本量	36725	36725	21817
组间R <sup>2</sup> 值	0.130	0.437	0.028

注:同表2。地方经济政策不确定指数可用数据仅至2017年,因此第(3)列样本量减少。



量,不确定性指数越大则政策效率越低。表4列示的结果显示,*OPEN*与*MKT*、*DEFT*、*EPU*的回归系数分别在10%、1%、10%水平上显著为负,表明政府数据开放能降低地区市场集中度、地方政府财政赤字水平、地方经济政策不确定性,即促进公平竞争、减少政府干预、提高政策效率。因此,优化营商环境是政府数据开放提高企业全要素生产率的重要作用机制。

(二)增强企业能力机制

现有文献的研究表明,以管理能力、营运能力、融资能力为代表的企业能力是全要素生产率的重要决定因素(周茂等,2015;宋敏等,2021;吴敏等,2022)。增强企业能力机制可能存在的主要原因是,企业在生产活动中纳入对政府数据开放的分析不仅有助于减少信息不对称以增强管理能力,还有助于提高市场需求、客户价值、供应商风险的评估效率以增强营运能力。此外,金融机构对客户信用等开放政府数据的分析能降低信贷风险溢价,叠加企业管理能力和营运能力增强带来的谈判筹码增加,显然会对企业的融资能力产生积极影响。基于此,此处将通过考察政府数据开放对管理能力、营运能力、融资能力的影响验证增强企业能力机制。其中,管理能力以管理费用率(*COST*,人均管理费用的自然对数值)衡量(周茂等,2015),管理费用率越低则管理能力越强;营运能力以资产周转率(*TURN*,营业收入与平均资产总额之比)衡量,资产周转率越高则营运能力越强;融资能力以Kaplan和Zingales(1997)开发的用于测度融资约束程度的KZ指数(*KZINDEX*)衡量,KZ指数越小则融资能力越强。表5列示的结果显示,*OPEN*与*COST*、*TURN*、*KZINDEX*的回归系数分别在5%水平上显著为负、5%水平上显著为正、1%水平上显著为负,表明政府数据开放能降低管理费用率、提高资产周转率、缓解融资约束,即增强管理能力、营运能力、融资能力。因此,增强企业能力是政府数据开放提高企业全要素生产率的重要作用机制。

表5 政府数据开放与增强企业能力

变量	管理能力	营运能力	融资能力
	(1)	(2)	(3)
	<i>COST</i>	<i>TURN</i>	<i>KZINDEX</i>
<i>OPEN</i>	-0.035** (0.013)	0.025** (0.011)	-0.013*** (0.003)
控制变量	是	是	是
企业固定	是	是	是
年份固定	是	是	是
样本量	36725	36725	36725
组间R <sup>2</sup> 值	0.011	0.088	0.276

注:同表2。

(三)促进高质量创新机制

高质量创新既能提高产出效率,也能获得高额垄断租金以增加产出价值,其作为企业全要素生产率提高的机制已获得充分论证(黄群慧等,2019;吴敏等,2022)。但高质量创新不仅面临知识正外部性和产出不确定性导致的垄断租金减少,更会因强力的资源承诺、高昂的资本投入以及不透明性和不可核实性而存在严重的融资约束(Aboody和Lev,2000;Bhattacharya等,2017)。政府数据开放引致的生产资源挤占减少、融资能力增强对缓解创新资金不足有积极作用,而企业基于开放政府数据的分析还能快速寻找创新方向,以促进高质量创新。为了验证促进高质量创新机制,此处借鉴

Chang等(2019)的做法以探索式专利申请数量的自然对数(*EXPLR*)、产品专利申请数量的自然对数(*PRDCT*)测度高质量创新。其中,探索式专利定义为至少引用60%非创造者过往专利的专利,由于引用越多他人专利表明越不是基于既有知识的创新,因而探索式专利越多则高质量创新越多;产品专利定义为国际专利分类中非B01类别的专利,其旨在创造新产品而非改进生产工艺,因而产品专利越多则高质量创新越多。<sup>①</sup>表6列示的结果显示,*OPEN*与*EXPLR*、*PRDCT*的回归系数均在5%水平上显著,表明政府数据开放能促使企业申请更多以探索式专利、产品专利,即政府数据开放通过促进高质量创新提高企业全要素生产率的作用机制存在。

表6 政府数据开放与促进高质量创新

变量	探索式专利	产品专利
	(1) <i>EXPLR</i>	(2) <i>PRDCT</i>
<i>OPEN</i>	0.067** (0.032)	0.144** (0.062)
控制变量	是	是
企业固定	是	是
年份固定	是	是
样本量	36725	36725
组间R <sup>2</sup> 值	0.043	0.069

注:同表2。

## 六、拓展分析

政府数据开放有助于提高企业全要素生产率的研究假说虽已充分证实,但其提高作用在不同特征的企业存在怎样的差异、政府数据开放质量是否会影响提高作用、哪些维度起到决定性作用等问题仍有待深入分析。对此,本文接下来将基于企业特征异质性、政府数据开放质量及维度对政府数据开放与企业全要素生产率提高的积极关系进行拓展分析。

### (一)企业特征异质性检验

相较于成熟企业、国有企业,年轻企业、非国有企业不仅内部经营管理结构和外部伙伴合作关系存在较大的不稳定性,而且普遍面临较为严重的融资约束(张思涵等,2022)。与此同时,年轻企业、非国有企业对公共制度的认知也不如成熟企业、国有企业全面和深刻(陶锋等,2017),如其全要素生产率会更受益于以“电视问政”为代表的服务型政府建设(张思涵等,2022)。由此可见,政府数据开放的全要素生产率提高作用在年轻企业、非国有企业中可能会更为显著。基于此,此处将对年轻企业和成熟企业、非国有企业和国有企业年龄进行分组检验。

第一,基于年轻企业和成熟企业的分组检验。为了比较政府数据开放的全要素生产率提高作用在年轻企业和成熟企业中的差异,此处借鉴陶锋等(2017)、张思涵等(2022)的做法将成立年数小于地区一年度中位数的样本划分为年轻企业组,其余划分为成熟企业组。表7第(1)~(4)列结果显示,*OPEN*与*TFPLP*、*TFPOP*的回归系数均在年轻企业组显著为正,而在成熟企业组不显著;系数

① 外观设计专利并不属于严格意义上的技术创新,因而两种测度方式均未包括外观设计专利。

差异检验显示,*OPEN*与*TFPLP*、*TFPOP*在年轻企业组的回归系数值显著大于成熟企业组。由此可见,政府数据开放对全要素生产率提高作用在年轻企业更为显著,前述理论预期得到充分支持。

第二,基于国有企业和非国有企业的分组检验。为了比较政府数据开放的全要素生产率提高作用在国有企业和非国有企业中的差异,此处借鉴陶锋等(2017)、张思涵等(2022)的做法依据企业产权性质将样本划分为国有企业组和非国有企业组。表7第(5)~(8)列结果显示,*OPEN*与*TFPLP*、*TFPOP*的回归系数均在非国有企业组显著为正,而在国有企业组不显著;系数差异检验显示,*OPEN*与*TFPLP*、*TFPOP*在非国有企业组的回归系数值显著大于国有企业组。以上检验结果表明,政府数据开放对全要素生产率提高的促进作用在非国有企业更为显著,这与前述理论预期相符。

表7 基于企业特征的分组检验

变量	年轻企业 (1)	成熟企业 (2)	年轻企业 (3)	成熟企业 (4)	非国有企业 (5)	国有企业 (6)	非国有企业 (7)	国有企业 (8)
	<i>TFPLP</i>	<i>TFPLP</i>	<i>TFPOP</i>	<i>TFPOP</i>	<i>TFPLP</i>	<i>TFPLP</i>	<i>TFPOP</i>	<i>TFPOP</i>
<i>OPEN</i>	0.062*** (0.011)	0.027 (0.017)	0.064*** (0.014)	0.006 (0.015)	0.044*** (0.014)	0.033 (0.021)	0.035*** (0.011)	0.030 (0.024)
系数差异	0.035***		0.058**		0.012**		0.005*	
控制变量	是	是	是	是	是	是	是	是
企业固定	是	是	是	是	是	是	是	是
年份固定	是	是	是	是	是	是	是	是
样本量	16602	20123	16602	20123	22804	13921	22804	13921
组间R <sup>2</sup> 值	0.271	0.244	0.219	0.198	0.257	0.209	0.213	0.161

注:同表2。

## (二)政府数据开放质量及维度检验

政府数据开放质量也是现有文献讨论的重点,其普遍认为仅有高质量的政府数据开放才可能发挥积极的价值创造作用。如前所述,政府数据开放的实践形式是通过在线平台向社会公众提供数据,因而政府数据开放能创造多大的价值主要是开放数据质量、平台体系建设、政策保障力度等多维度因素综合作用的结果(Nikiforova和Lnenicka,2021)。特别地,中国政府数据开放事业主要是由地方政府推进,但各地政府数据开放质量参差不齐,部分地方政府数据开放平台对数据的更新较慢,抑或提供的数据量级较少,甚至部分地区的政府数据开放平台存在只有平台而无数据的异常现象。由此可见,深入考察政府数据开放质量对企业全要素生产率的影响也有重要意义。

基于此,本文将利用复旦大学数字与移动治理实验室构建的中国开放数林指数进行拓展检验。该指数对各地政府数据开放平台进行了较为全面的客观评价,除综合指数外还包括数据层指数、平台层指数、准备度指数等细分指数。其中,数据层指数是对开放数据质量的评价,如开放范围、数据规范、数据数量;平台层指数是对平台建设质量的评价,如用户体验、互动反馈、数据发现与获取;准备度指数是对政策保障力度的评价,如法规政策效力与内容、标准规范制定、组织与实施。此处以综合指数的自然对数测度政府数据开放的整体质量(*QUALITY*),而数据层指数、平台层指数、准备度指数的自然对数则用于测度开放数据质量(*DATA*)、平台建设质量(*PLAT*)、政策保障力度(*POLICY*)。本文预期,政府数据开放质量越高、开放数据质量越高、平台建设质量越高、政策保障

力度越强则企业全要素生产率越高。

表8列示了政府数据开放质量及维度与企业全要素生产率的回归结果。第(1)、(2)列结果显示, *QUALITY* 与 *TFPLP*、*TFPOP* 的回归系数均在至少 5% 水平上显著为正, 表明政府数据开放质量越高则全要素生产率越高, 这也再次支持了基准回归结论。细分指数维度后, 第(3)、(4)列结果显示, *DATA* 与 *TFPLP*、*TFPOP* 的回归系数均在至少 5% 水平上显著为正, 表明开放数据质量越高则全要素生产率越高; 第(5)、(6)列结果显示, *PLAT* 与 *TFPLP*、*TFPOP* 的回归系数均在 1% 水平上显著为正, 表明平台体系建设质量越高则全要素生产率越高; 第(7)、(8)列结果显示, *POLICY* 与 *TFPLP* 的回归系数在 1% 水平上显著为正, 与 *TFPOP* 的回归系数接近在 10% 水平上显著为正, 表明政策保障力度越强则全要素生产率越高。以上检验结果意味着, 政府数据开放不可忽视开放数据质量和平台建设质量的重要性, 且还需强有力的政策制度来保障政府数据开放。

表 8 政府数据开放质量及维度与企业全要素生产率								
变量	政府数据开放质量		开放数据质量		平台建设质量		政策保障力度	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	<i>TFPLP</i>	<i>TFPOP</i>	<i>TFPLP</i>	<i>TFPOP</i>	<i>TFPLP</i>	<i>TFPOP</i>	<i>TFPLP</i>	<i>TFPOP</i>
<i>QUALITY</i>	0.016*** (0.006)	0.016** (0.006)						
<i>DATA</i>			0.021*** (0.007)	0.020** (0.008)				
<i>PLAT</i>					0.016*** (0.005)	0.016*** (0.004)		
<i>POLICY</i>							0.013** (0.006)	0.012 (0.007)
控制变量	是	是	是	是	是	是	是	是
企业固定	是	是	是	是	是	是	是	是
年份固定	是	是	是	是	是	是	是	是
样本量	14495	14495	14495	14495	14495	14495	11123	11123
组间 R <sup>2</sup> 值	0.152	0.129	0.153	0.130	0.152	0.130	0.090	0.072

注: 同表 2。开放数林指数仅 2017 年起有数据可用, 其中准备度指数仅 2018 年起有数据可用, 因此回归样本量减少; 所有因变量均较自变量滞后 1 期。

七、结论与政策建议

数据要素在推动经济高质量发展中的重要作用已在全球达成共识, 尤其是领域内具有关键性、基础性、权威性特征的海量政府数据, 但政府数据能发挥多大程度的积极作用取决于其开放程度及社会主体的利用程度。为了激励社会主体利用政府数据以推动经济高质量发展, 政府数据开放逐渐成为各国治理政府数据的重要制度性工具, 即通过在线平台主动向社会主体免费提供可共享且无版权的原始政府数据。尽管诸多研究表明企业的生产经营活动越来越倚重开放政府数据, 且广泛预计政府数据开放将创造巨大经济社会价值, 但现有研究多是在制度基础、开放程度、决定因素等就开放论开放的范畴内展开分析, 而其如何甚至是否创造价值仍模糊不清, 尤其是对企业全要素生产率的影响严重制约了对政府数据开放收益、成本及障碍的认识。



基于此,本文利用地方开通政府数据开放平台的准自然实验,通过双重差分模型考察政府数据开放对企业全要素生产率的影响,以揭示宏观政府数据开放在微观企业层面的价值创造作用。研究发现,政府数据开放通过优化营商环境、增强企业能力、促进高质量创新显著提高了企业全要素生产率;但政府数据开放对企业全要素生产率的提高作用主要存在于年轻企业、非国有企业;此外,政府数据开放质量越高时对企业全要素生产率的提高作用越明显,且开放数据质量、平台建设质量、政策保障力度决定着政府数据开放对企业全要素生产率提高作用的大小。以上检验结果充分说明,政府数据开放在企业层面有效发挥了价值创造作用,但不同情景下存在些许差异,且开放数据质量、平台体系建设、政策保障力度都会深刻影响政府数据开放的价值创造作用。

依据前述研究结论,本文提出如下政策建议。

第一,在全面推进中国式现代化的过程中,各级政府应充分重视政府数据对推进数字经济发展和数字政府建设的重要作用。数据已然成为比肩土地和资本的新型生产要素,且无论是数字经济发展还是数字政府建设,数据都是其底层逻辑支撑。特别地,作为国家基础性战略资源,以领域内关键性、基础性、权威性数据为主要特征的政府数据占据着全社会数据资源80%以上的份额,因而充分发挥政府数据的生产要素和治理要素属性不仅有助于加快培育数据要素市场,更能强化政府的决策科学化水平和管理服务效率。有鉴于此,主管部门应充分重视政府数据对推进数字经济发展和数字政府建设的重要作用,积极推进和引导职能部门和社会主体加强对政府数据的开发、共享、利用,尤其是探索出符合中国式现代化的发展路径。如疏通职能部门间的数据烟囱以促进各部门协同履行公共职责,构建以政府数据为底层逻辑支撑的政府治理体制机制,积极宣传政府数据以提高社会主体对政府数据的利用效率。

第二,在推动经济高质量发展战略中,主管部门应将推进政府数据高质量开放作为重要抓手。企业在生产活动中与公共制度、市场主体的摩擦是经济运行效率低下的重要原因。本文分析发现政府数据开放通过优化营商环境、增强企业能力、促进高质量创新提高企业全要素生产率,且开放数据质量、开放平台建设、政策保障力度是提高效用的重要决定维度。由此可见,高质量开放政府数据对提高经济运行效率以实现经济发展质量具有重要积极作用。但当前政府数据开放机制还存在诸多缺陷,不仅许多地区未开通政府数据开放平台,而且既有政府数据开放平台提供的数据也存在有平台无数据、数据杂乱无章等形式重于实质的现象。如山西省尚未开通省级政府数据开放平台;陕西省虽已开通省级政府数据开放平台,但却仅提供141条数据且数据更新非常滞后;浙江省虽开放了非常多的数据,但多是杂乱无章的简单发布。对此,主管部门首先应鼓励地方政府依法主动开放政府数据,尤其是社会发展迫切需要的高价值政府数据;其次需要完善政府数据开放平台体系建设,如提高数据更新频率、完善用户交流和反馈机制;再次需要推出中央层面的政府数据分类分级管理、政府数据开放原则等规律,为政府数据开放提供强有力的政策保障;最后需要加快建设国家层面的政府数据统一开放平台,实现政府数据跨区域流动,以促进区域协调发展、均衡发展。

第三,主管部门应对年轻企业、非国有企业给予更多关注。年轻企业、非国有企业无疑是推动中国经济社会高质量发展的重要力量,如工信部在2018年的国新办新闻发布会上宣布民营企业贡献了80%的就业、70%以上的技术创新成果、60%以上的国内生产总值、50%以上的税收。但相较于成熟企业、国有企业,年轻企业、非国有企业在治理结构、治理能力、市场规则和公共制度认知等方面存在不足,甚至可能遭遇金融机构和职能部门的歧视。本文分析发现,政府数据开放对全要素生产率提高作用在年轻企业、非国有企业中更为显著。结合前述优化营商环境、增强企业能力、促进高质量创新的作用机制,主管部门首先需要营造公平的市场竞争环境、抑制对企业活动的干预以

及持续提高政策效率;其次需要引导年轻企业、非国有企业通过积极开发利用开放政府数据增强自身能力。这既有助于完善政府数据开放体系,也有助于形成互动治理模式以创新公共政策,从而实现良性循环,最终增强年轻企业、非国有企业的竞争力。

### 参考文献

- [1]胡海峰,窦斌,王爱萍.企业金融化与生产效率[J].世界经济,2020,43(1):70~96.
- [2]刘莉亚,金正轩,何彦林,朱小能,李明辉.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[J].经济学(季刊),2018,17(4):1329~1360.
- [3]陈尚龙.论政府数据开放的理论基础[J].理论与改革,2016(6):104~107.
- [4]蔡继明,刘媛,高宏,陈臣.数据要素参与价值创造的途径——基于广义价值论的一般均衡分析[J].管理世界,2022,38(7):108~121.
- [5]蔡继明,曹越洋,刘乐易.论数据要素按贡献参与分配的价值基础——基于广义价值论的视角[J].数量经济技术经济研究,2023,(8):5~24.
- [6]董松柯,刘希章,李娜.数字化转型是否降低企业研发操纵?[J].数量经济技术经济研究,2023,(4):28~51.
- [7]段盛华,于凤霞,关乐宁.数据时代的政府治理创新——基于数据开放共享的视角[J].电子政务,2020,(9):74~83.
- [8]韩啸,吴金鹏.政府数据开放水平的驱动因素:基于跨国面板数据研究[J].电子政务,2020,(6):98~106.
- [9]胡业飞,刘梦露.创新激励目标下的契约设计:基于用户协议的政府数据开放平台治理研究[J].电子政务,2021,(10):15~28.
- [10]黄群慧,余泳泽,张松林.互联网发展与制造业生产率提升:内在机制与中国经验[J].中国工业经济,2019,(8):5~23.
- [11]李万利,刘虎春,龙志能,汤旭东.企业数字化转型与供应链地理分布[J].数量经济技术经济研究,2023,(8):90~110.
- [12]李唐,李青,陈楚霞.数据管理能力对企业生产率的影响效应——来自中国企业—劳动力匹配调查的新发现[J].中国工业经济,2020,(6):174~192.
- [13]马连福,曹春方.制度环境、地方政府干预、公司治理与IPO募集资金投向变更[J].管理世界,2011,(5):127~139+148+188.
- [14]牛志伟,许晨曦,武瑛.营商环境优化、人力资本效应与企业劳动生产率[J].管理世界,2023,39(2):83~100.
- [15]钱雪松,康瑾,唐英伦,曹夏平.产业政策、资本配置效率与企业全要素生产率——基于中国2009年十大产业振兴规划自然实验的经验研究[J].中国工业经济,2018,(8):42~59.
- [16]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.
- [17]宋烁.政府数据开放是升级版的政府信息公开吗?——基于制度框架的比较[J].环球法律评论,2021,43(5):52~66.
- [18]宋渊洋,黄礼伟.为什么中国企业难以国内跨地区经营?[J].管理世界,2014,(12):115~133.
- [19]唐浩丹,方森辉,蒋殿春.数字化转型的市场绩效:数字并购能提升制造业企业市场势力吗?[J].数量经济技术经济研究,2022,(12):90~110.
- [20]陶锋,胡军,李诗田,韦锦祥.金融地理结构如何影响企业生产率?——兼论金融供给侧结构性改革[J].经济研究,2017,52(9):55~71.
- [21]王康,赵蕊,苏盖美.数字化如何助力全国统一大市场建设——基于企业资本跨地区流动视角[J].中国经济学,2023,(2):46~72+352~353.

- [22]吴非,胡慧芷,林慧妍,任晓怡.企业数字化转型与资本市场表现——来自股票流动性的经验证据[J].管理世界,2021,37(7):130~144+10.
- [23]吴敏,曹婧,毛捷.地方公共债务与企业全要素生产率:效应与机制[J].经济研究,2022,57(1):107~121.
- [24]夏杰长,刘诚.行政审批改革、交易费用与中国经济增长[J].管理世界,2017,(4):47~59.
- [25]徐翔,赵墨非.数据资本与经济增长路径[J].经济研究,2020,55(10):38~54.
- [26]袁淳,肖土盛,耿春晓,盛誉.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.
- [27]张思涵,张明昂,王雨坤.服务型政府建设与企业高质量发展[J].财经研究,2022,48(9):109~123.
- [28]赵宸宇,王文春,李雪松.数字化转型如何影响企业全要素生产率[J].财贸经济,2021,42(7):114~129.
- [29]郑磊.开放不等于公开、共享和交易:政府数据开放与相近概念的界定与辨析[J].南京社会科学,2018,(9):83~91.
- [30]周茂,陆毅,陈丽丽.企业生产率与企业对外直接投资进入模式选择——来自中国企业的证据[J].管理世界,2015,(11):70~86.
- [31]朱峥.政府数据开放的权利基础及其制度构建[J].电子政务,2020,(10):117~128.
- [32]Aboody D., Lev B., 2000, *Information Asymmetry, R&D, and Insider Gains* [J], *Journal of Finance*, 55 (6), 2747~2766.
- [33]Beck T., Levine R., Levkov A., 2010, *Big Bad Banks? The Winners and Losers from Bank Deregulation in the United States* [J], *Journal of Finance*, 65 (5), 1637~1667.
- [34]Bhattacharya U., Hsu P., Tian X., Xu Y., 2017, *What Affects Innovation More: Policy or Policy Uncertainty?* [J], *Journal of Financial and Quantitative Analysis*, 52 (5), 1869~1901.
- [35]Chang X., Chen Y. Y., Wang S. Q., Zhang K., Zhang W. R., 2019, *Credit Default Swaps and Corporate Innovation* [J], *Journal of Financial Economics*, 134 (12), 474~500.
- [36]Chen D. Q., Preston D. S., Swink M., 2015, *How the Use of Big Data Analytics Affects Value Creation in Supply Chain Management* [J], *Journal of Management Information Systems*, 32 (4), 4~39.
- [37]Cong L. W., Xie D., Zhang L., 2021, *Knowledge Accumulation, Privacy, and Growth in a Data Economy* [J], *Management Science*, 67 (10), 6480~6492.
- [38]Conradie P., Choenni S., 2014, *On the Barriers for Local Government Releasing Open Data* [J], *Government Information Quarterly*, 31 (S1), S10~S17.
- [39]Donker F. W., Loenen B. V., 2017, *How to Assess the Success of the Open Data Ecosystem?* [J], *International Journal of Digital Earth*, 10 (3), 284~306.
- [40]Farboodi M., Veldkamp L., 2021, *A Growth Model of the Data Economy* [R], NBER Working Paper, No. 28427.
- [41]Ghasemaghaei M., Calic G., 2019, *Does Big Data Enhance Firm Innovation Competency? The Mediating Role of Data-Driven Insights* [J], *Journal of Business Research*, 104, 69~84.
- [42]Ghasemaghaei M., Calic G., 2020, *Assessing the Impact of Big Data on Firm Innovation Performance: Big Data Is Not Always Better Data* [J], *Journal of Business Research*, 108, 147~162.
- [43]Jetzek T., Avital M., Bjorn-Andersen N., 2019, *The Sustainable Value of Open Government Data* [J], *Journal of the Association for Information Systems*, 20 (6), 702~734.
- [44]Kaplan S. N., Zingales L., 1997, *Do Investment-Cash Flow Sensitivities Provide Useful Measures of Financing Constraints?* [J], *Quarterly Journal of Economics*, 112 (1), 169~215.
- [45]Levinsohn J., Petrin A., 2003, *Estimating Production Functions Using Inputs to Control for Unobservables* [J], *Review of Economic Studies*, 70 (2), 317~341.
- [46]Magalhaes G., Roseira C., 2020, *Open Government Data and the Private Sector: An Empirical View on Business*

*Models and Value Creation* [J], *Government Information Quarterly*, 37 (3), 101248.

[47] Moro A., Fink M., 2013, *Loan Managers' Trust and Credit Access for SMEs* [J], *Journal of Banking & Finance*, 37 (3), 927~936.

[48] Müller O., Fay M., Vom Brocke J., 2018, *The Effect of Big Data and Analytics on Firm Performance: An Econometric Analysis Considering Industry Characteristics* [J], *Journal of Management Information Systems*, 35 (2), 488~509.

[49] Nikiforova A., Lnenicka M., 2021, *A Multi-Perspective Knowledge-Driven Approach for Analysis of the Demand Side of the Open Government Data Portal* [J], *Government Information Quarterly*, 2021, 38 (4), 101622.

[50] Olley G. S., Pakes A., 1996, *The Dynamics of Productivity in the Telecommunications Equipment Industry* [J], *Econometrica*, 64 (6), 1263~1297.

[51] Park S., Gil-Garcia J. R., 2022, *Open Data Innovation: Visualizations and Process Redesign as a Way to Bridge the Transparency-Accountability Gap* [J], *Government Information Quarterly*, 39 (1), 101456.

[52] Wamba S. F., Gunasekaran A., Akter S., Ren S. J., Dubey R., Childe S. J., 2017, *Big Data Analytics and Firm Performance: Effects of Dynamic Capabilities* [J], *Journal of Business Research*, 70, 356~365.

[53] Wang H., Lo J., 2020, *Factors Influencing the Adoption of Open Government Data at the Firm Level* [J], *IEEE Transactions on Engineering Management*, 67 (3), 670~682.

[54] Yu J., Shi X., Guo D., Yang L., 2021, *Economic Policy Uncertainty (EPU) and Firm Carbon Emissions: Evidence Using a China Provincial EPU Index* [J], *Energy Economics*, 94, 105071.

[55] Zhao Y., Fan B., 2021a, *Effect of an Agency's Resources on the Implementation of Open Government Data* [J], *Information & Management*, 58 (4), 103465.

[56] Zhao Y., Fan B., 2021b, *Understanding the Key Factors and Configurational Paths of the Open Government Data Performance: Based on Fuzzy-Set Qualitative Comparative Analysis* [J], *Government Information Quarterly*, 38 (3), 101580.

[57] Zhao Y., Liang Y., Yao C., Han X., 2022, *Key Factors and Generation Mechanisms of Open Government Data Performance: A Mixed Methods Study in the Case of China* [J], *Government Information Quarterly*, 39 (4), 101717.

[58] Zuiderwijk A., Janssen M., 2014, *Open Data Policies, Their Implementation and Impact: A Framework for Comparison* [J], *Government Information Quarterly*, 31 (1), 17~29.

## The Value Creation Role of Open Government Data: The Perspective of Firm's Total Factor Productivity

PENG Yuanhuai

(School of Economics, Jinan University)

**Summary:** Government data that are obtained in accordance with the law during the fulfillment of public duties not only have a rich variety and large volume but also contain crucial, fundamental, and authoritative data in the field. It plays an important role in promoting the development of the digital economy and the construction of digital government. Data can only achieve their maximum effectiveness when they are fully utilized, but a large amount of government data is not fully utilized due to the reluctance or difficulty of the supervising units to open them to the public. To promote the utilization of government data by the public, governments worldwide have begun to actively promote open government data in recent years, which are providing original government data to the public free of charge and indiscriminately.

In parallel with the active promotion of open government data by practitioners, the academic



community widely anticipates that it will create enormous economic and social value. Previous literature has mostly relied on the research paradigm of public administration to analyze open government data in terms of the institutional basis, openness level, and determining factors, but how it creates value or even if it creates value is unclear. Specifically, the fundamental goal of open government data is to promote high-quality economic development. Against the backdrop of slowing economic SGR and rising factor costs, improving total factor productivity is the main driving force for promoting high-quality economic development. However, only a few studies have discussed the value-creation role of open government data from the perspective of total factor productivity. As micro-entities participating in market economic activities, enterprises are the main subjects encouraged to use open government data to create value. This raises the following question: how does open government data affect the total factor productivity of enterprises?

The practical form of open government data is for functional departments to place data that need to be opened in various common data formats on online platforms, allowing the public to freely search, preview, and download or retrieve them through data interfaces. In view of this, based on a quasi-natural experiment of the provincial government's data opening platform, this study uses the difference-in-difference model to test the impact of open government data on the total factor productivity of enterprises. The results reveal that open government data significantly increase total factor productivity. The main reason is that open government data help optimize the business environment, enhance enterprise capabilities, and promote high-quality innovation. Moreover, the effect of open government data on total factor productivity is mainly in young firms and non-state-owned enterprises. Additionally, the higher the level of open government data, the higher the total factor productivity. Further, open data quality, platform construction quality, and policy support are important dimensions that determine the magnitude of total factor productivity improvement.

The main contributions of this study are as follows. First, although open government data have been widely expected to create enormous economic and social value, the existing literature has limited its research to discussing openness in terms of institutional foundation, openness level, and determining factors. This study focuses on the impact of open government data on micro-entities and supplements the theoretical and empirical research on the value creation of open government data. Second, the essential characteristics of the digital economy are data-driven, and the direct effect of open government data is to improve the availability and diversity of data. Therefore, examining the impact of open government data on the total factor productivity of enterprises can help deepen our understanding of the economic benefits of data. Third, government data are significantly different from enterprise data and personal data due to their legal status, institutional foundation, information content, and other factors. The empirical analysis of the economic consequences of open government data can help enrich the research on the value creation of heterogeneous data elements.

**Keywords:** Open Government Data; Total Factor Productivity; Value Creation; Digital Government

**JEL Classification:** D24; D23

(责任编辑:焦云霞)