

开放政府数据的经济增长效应研究

张晨^{1,2}, 万相昱^{2,3}, 姜智超², 唐亮⁴

- (1. 山东财经大学财政税务学院, 山东 济南 250002; 2. 中国社会科学院大学, 北京 100006;
3. 中国社会科学院数量经济与技术经济研究所, 北京 100732;
4. 东北师范大学经济与管理学院, 吉林 长春 130117)

摘要: 新一代信息通信技术的飞速发展使得数据要素成为驱动经济增长的重要力量。作为全社会数据资源的主要掌握者政府, 无论考虑公众知情权还是经济社会效益, 都应当适度公开政府数据, 用以提升政府治理效能和创造经济价值。关注开放政府数据和经济增长的关系, 理论分析指出公共数据开放不仅可以为社会经济活动提供生产、创新等所需数据要素, 直接拉动经济增长, 而且可以通过提高政府治理效能间接发挥经济增长效应。利用 2004—2019 年 1 911 个国家样本, 实证发现开放政府数据可以显著促进经济增长。中介效应检验结果表明, 虽然政府治理效能的间接影响显著, 但是直接经济增长效应占据主导地位。异质性分析发现, 信息技术越先进, 开放政府数据对经济增长的促进作用越大, 意味着“技术+数据”是发挥政府数据要素价值的重要方式。

关键词: 开放政府数据; 经济增长; 政府治理效能; 信息技术

中图分类号: F062.4 文献标识码: A 文章编号: 1005-0566(2023)02-0001-11

Economic growth effects of open government data

ZHANG Chen^{1,2}, WAN Xiangyu^{2,3}, JIANG Zhichao², TANG Liang⁴

- (1. School of Public Finance & Taxation, SDUFE, Jinan 250002, China;
2. University of Chinese Academy of Social Sciences, Beijing 100006, China;
3. Institute of Quantitative & Technological Economics, CASS, Beijing 100732, China;
4. Business School, Northeast Normal University, Changchun 130117, China)

Abstract: The rapid development of the new generation of information and communication technology makes data elements become an important force driving economic growth. As the main holder of social data resources, the government should properly disclose government data to improve government governance efficiency and create economic value, regardless of the public's right to know or economic and social benefits. This paper focuses on the relationship between open government data and economic growth. Theoretical analysis points out that open public data, on the one hand, provides production, innovation and other required data elements for social and economic activities, and directly drives economic growth. On the other hand, the economic growth effect can be brought into play indirectly by improving the efficiency of government governance. Using a sample of 1 911 countries from 2004 to 2019, the empirical study finds that open government data significantly boosts economic growth. The mediating effect test results show that although the indirect effect of government governance efficiency is significant, the direct economic growth effect plays a dominant role.

收稿日期: 2022-10-16 修回日期: 2023-02-08

基金项目: 国家社科基金项目“推动数字经济和实体经济深度融合研究”(21AZD022)。

作者简介: 张晨(1992—), 男, 河南平顶山人, 山东财经大学财政税务学院讲师, 博士, 研究方向为数字经济与收入分配。

Heterogeneity analysis results show that the more advanced ICT is, the greater the promotion effect of open government data on economic growth is, which means that “Technology + Data” is an important way to give full play to the value of government data elements.

Key words: open government data; economic growth; government governance efficiency; information communication technology

2015 年 8 月 31 日,国务院印发《促进大数据发展行动纲要》(以下简称《纲要》),指出信息技术与经济社会的交汇融合引发了数据迅猛增长,数据已成为国家基础性战略资源,大数据正日益对全球生产、流通、分配、消费活动以及经济运行机制、社会生活方式和国家治理能力产生重要影响。《纲要》提出加快政府数据开放共享,推动资源整合,是提升政府治理能力的重要途径。大力推动政府部门数据共享、稳步推动公共数据资源开放是大数据发展的主要任务。为何开放政府数据(Open Government Data, OGD)如此重要?从政府角度来看,政府数据开放与共享将极大提升政府治理能力与治理效能。第一,通过建立国家宏观调控数据体系,互联网数据与政务数据的关联分析与融合利用将提高政府在金融、消费、投资、节能减排等领域宏观调控的科学化。第二,推动政府部门和企事业单位将生产经营、市场监管、违法失信、投诉举报等数据进行聚合分析,将加强政府事中事后监管和服务,提升政府决策和风险防范能力,推动政府治理实现精准化。第三,利用公民、企业社会信用大数据,依托国家间信用信息共享交换平台,运用大数据手段,简化注册登记、市场准入等商事服务流程,提高商事服务的便捷化。第四,政府部门间数据流通将加强社会治理领域数据的搜集、挖掘和分析,强化处理重大突发公共事件的数据支持,促进安全保障高效化。对企业而言,数据资源向数据要素的转变成为驱动企业创新发展的新力量。政府数据作为一项重要战略资产,从汽车生产商到数字平台,许多公司都可以使用这些数据,用于生产和投资决策、营销和库存管理以及长期战略规划。并且在评估银行的金融健康状况和国家金融体系的整体稳定性方面也发挥了重要的作用。

本文认为深入研究特别是实证评估开放政府

数据价值具有较强的理论价值与现实意义。从实践来看,政府对于公共数据如何开放、哪些数据具有经济价值等核心问题并不十分清晰。已开放数据往往难以满足商业化需求,数据质量、种类、开放延续性难以得到保障,严重制约了开放政府数据的经济社会价值挖掘^[1]。与此同时,遗憾的是学界鲜有文献针对以上实践问题进行深度探讨。付熙雯等^[2]指出现有研究缺乏对 OGD 效果和价值维度的测量,包括规范的研究设计、可操作化的测量指标以及一手采集的实证数据,因此关于 OGD 预期价值的评估与验证将是未来研究的重点方向。但是经验证据的缺乏使得相关研究难以为 OGD 实践提供具有借鉴意义的参考。基于理论与现实需求,本文尝试针对公共数据开放的价值进行评估,具体关注 OGD 的经济增长效应。第一,开放获取的公共数据成为社会经济活动所必须的数据资源或数据要素,直接影响企业生产、销售、研发创新等行为,拉动经济增长。第二,政府通过开放数据提高自身透明度,来自媒体、民众等外部对象的监督将提高政府治理效能。政府部门间数据的共享与关联,为政府在公共服务提供、监督管理、法治建设、打击腐败等犯罪活动领域提供强有力的数据支持和效率提升工具。政府治理效能数据开放过程中得以提升,进而产生间接的经济增长效应。围绕上述内容,本文将从理论和实证两个方面分析和评估 OGD 对于经济增长的影响。

一、文献综述

(一) 开放政府数据的历史发展

对于 OGD 的理解需要结合其历史发展路径。二战之后,美国新闻界精英发起了“信息自由”运动,他们认为正是由于缺乏新闻自由,一些专制国家发动战争,人民有权利了解政府运作流程。随后伴随着“知情权”(Right to Know)、世界信息自由委员会(The Committee on World Freedom of In-

formation)等一系列概念和组织出现,社会各界对于政府信息公开的需求日益强烈。美国于1967年和1976年相继出台了《信息自由法案》(Freedom of Information Act)和《阳光政府法案》(Government in the Sunshine Act),信息自由运动开始席卷全球。伴随信息技术的飞速发展,计算机领域的技术精英为了打破“专有软件”(Proprietary Software)的商业模式,倡议公布原始代码和远程协作编程。1998年“开源”(Open Source)风潮从美国硅谷兴起,日益成为影响计算机领域技术发展的一股重要力量。数据是软件工程的重要组成部分,加之“知情权”深入人心,美国社会各界对于公共数据产生了浓厚兴趣。2007年12月,公共数据开放的标准和准则在美国加州诞生,包括完整性、原始性、及时性、可读性、机器可处理性、无歧视性、非专有性和无须授权八大原则。随着互联网、云计算等新一代信息通信技术广泛应用,政府数据在规模和价值方面获得极大提升。2009年,美国奥巴马政府出台了《开放政府指令计划》(Open Government Directive Plan),同年上线全球第一个政府数据开放平台“data.gov”。紧跟美国步伐,加拿大、英国、澳大利亚等国家相继制定和实施OGD计划。我国也于2015年8月出台了《促进大数据发展行动纲要》,将OGD纳入建设数据强国战略当中。与最初信息自由运动相比,近年来OGD的目的除了提高政府透明度外,额外增加了经济效益维度,导致开放数据的内容和使用对象均发生了较大改变。正如Ubalidi^[3]所言,OGD不仅有利于实现社会包容与更高参与性的民主制度,还将允许第三方,如企业、社会组织和个人,利用公共数据创新产品和服务,从而刺激经济增长。

(二) 开放政府数据的价值及评估

根据OGD的目的,可以将其价值划分为政府治理和经济效益,前者涉及政府通过开放数据增进公众对于其运行机制和流程的了解,提高在公共服务、打击腐败、政策监管等方面的治理效能。后者则意味着公众利用数据进行产品创新和服务创新,实现政府数据的经济效益。付熙雯等^[2]详细梳理了过往研究关于OGD的价值维度划分,将

其归纳为政治价值、经济价值、社会价值、技术价值、战略价值、科学研究价值、精神文化价值以及环境生态价值,认为政治价值和经济价值是最重要的两个方面。作为被研究最多的维度,政治价值指的是OGD为公共治理所带来的收益,如提升政府部门回应公共需求的能力、提高公众参与公共治理的积极性以及改善公共服务的质量与效率^[4]。经济价值的讨论频率仅次于政治价值,指的是OGD将为政府、企业和公众带来经济效益。如OGD降低了企业获取数据与信息的成本,海量公共数据将优化企业决策、实现商业创新、改善服务和产品质量^[5]。

诸多文献针对OGD的多维价值进行了理论分析,但是其本质上仍然属于一种实践运动,重点在于OGD是否确实发挥了理论所指出的各种价值,意味着需要对OGD的真实效应进行评估。国外学界围绕OGD价值评估进行了大量研究。Charalabidis等^[6]以用户利用为输出变量,探究OGD的价值模型,对开放政府数据平台给用户和数据提供者带来的影响进行考察。通过问卷调查,利用该模型对欧盟OGD项目的社会和商业价值进行评估。Stuermer等^[7]采用社会投资回报的方法,对OGD的投入产出、效果影响进行区分,并基于此研究,提出了测量OGD在社会发展与生态环境方面的价值评估框架。Magalhaes等^[8]收集了178个使用政府数据的美国公司信息,总结了OGD在商业中包括营销、咨询、服务研发等的11种利用模式。与国外相比,国内学界和社会对于OGD的价值评估成果相对较少,仍然停留在理论分析,鲜有实证文献。

(三) 文献评述

上述文献为本文的研究提供了有益借鉴。特别是通过梳理OGD的历史发展历程,明晰其目的从一开始“知情权”所要求的政府透明度逐步转向新一代信息技术飞速发展所带来的数据要素需求,为本文的理论分析构筑了坚实基础。围绕OGD的目的即政府治理和社会经济效益,国外学界、企业和政府从不同角度对其价值进行了量化评估,同样对本文的实证研究具有较大启发。遗

憾的是国内学者关于 OGD 价值评估的研究并无太多成果,政府和企业同样如此。综上所述,本文将在以下 3 个方面作出贡献。第一,将 OGD 对政府治理的影响以及其直接经济效益统一到经济增长的分析框架中,探究 OGD 的直接经济增长效应和通过政府治理所发挥的间接经济影响。第二,为 OGD 的价值评估提供更多的经验证据。采用多种计量模型科学评估 OGD 的经济增长效应。进一步引入政府治理变量分离直接和间接作用,探究何种效应占据主导地位。第三,引入技术因素为 OGD 经济增长效应的异质性提供解释,强调技术与数据匹配的重要性。

二、理论分析

OGD 的发展历史告诉我们,其价值体现在政府治理和社会经济效益两个维度,即使政府治理,其最终目的仍在于服务经济发展,因此本文将从经济增长角度评估 OGD 价值。一方面 OGD 为企业生产、研发创新等经济活动直接提供数据资源,数据要素成为推动经济增长的直接力量。另一方面,通过开放获取数据,政府透明度和治理效能将获得显著提升,间接经济增长效应不可忽视。

(一) 开放政府数据与经济增长

OGD 影响经济增长最为直接的方式就是政府数据转变为企业生产、销售、物流等环节的数据资源,成为除资本、劳动力、创新外可能最为重要的投入要素。数据要素承载了大量有价值的信息,能够提高生产过程中劳动、资本等其他传统要素间的协同性,进而提升产出效率。此外,数据要素的非竞争性、非排他性和低成本复制特征将在宏观层面产生规模效应^[9],支撑以数据信息为基础的模式创新,提高资源产品配置效率。本文认为 OGD 将通过以下 3 种机制直接促进经济增长。

第一,OGD 将激励企业从经验型决策转向数据驱动型决策。McAfee 等^[10]认为数据驱动型决策充分利用经验数据而非商业直觉,提高企业决策的科学性和准确性。案例研究发现,采用数据驱动型决策的企业有着更好的财务状况和更高的生产效率。以交通运输业为例,短期业务决策涉及对汽车需求状况和近期前景的建模,这对装配

厂提出生产效率非常重要。结合政府数据与企业内部数据,利用同一领域的外部专家与内部职能团队,开展销售预测、期望库存水平管理、竞争对手行为预判和定价,以便在装配层面作出有关生产的商业决策。长期来看,装配厂扩建或选址的投资决策需要依赖政府数据进行供应链管理、物流、测绘和路线规划。

第二,OGD 可以通过信息挖掘和协同创新提高企业生产效率。海量政务数据为企业经营决策的科学化提供了可能。借助大数据技术,通过信息挖掘企业能够洞察未来短期甚至长期的人口、经济、环境等变化趋势,为战略决策提供支撑。同时利用市场、人口等数据刻画复杂且完整的客户画像,有利于制定针对性的产品和服务,提升生产效率,此时政府数据将成为决定企业生产效率的信息资产。协同创新是 OGD 影响企业生产效率的另一机制,创新不仅需要技术,而且需要数据作为基础的投入要素。OGD 通过开放公共数据在政府、科研院所以及企业之间形成信息共享,各类创新主体共同挖掘政府数据价值,广泛形成协同创新,降低试错成本,提高创新效率^[11]。

第三,OGD 通过降低信息摩擦,提高金融资源配置效率与企业投资绩效。Zhu^[12]认为降低信息获取成本将提高金融产品价格的信息含量,有助于投资者精确锚定潜力企业,金融资源配置效率获得提升。OGD 可以帮助金融公司评估寻求融资的企业,从初创企业、小企业到大公司,为涉及公司、特定行业、房地产、货币、商品和其他资产的各种投资决策提供信息。信贷公司使用企业、专业执照、财产、法庭记录等数据发现欺诈并降低风险。政府数据还有助于企业研究本国和全球经济前景,以了解消费者行为、识别和量化风险,并优化其战略布局。Farboodi 等^[13]发现大数据技术的引入显著提高了美国大企业股票价格的信息含量,高效率的定价引导资本流入更具发展潜力的行业和企业。

(二) 开放政府数据、治理效能与经济增长

大量研究证明政府治理能力的高低是决定经济增长的关键要素^[14-15]。OGD 至少可以从两个

方面提高政府治理效能。第一,政府数据开放获取意味着政府透明度提升,而透明度是影响公共治理行为的关键因素。第二,从数据使用层面,海量的政务数据将为公共政策的制定和实施提供更为科学的依据和工具,提升政府自身的决策效率。下面将从政府透明度和决策效率层面详细阐述 OGD 如何提高政府治理效能。

1. 开放政府数据与政府透明度

更多的公开信息意味着更高的政府透明度,来自公众的广泛监督对提高治理效能至关重要。在公共信息获取通畅的国家,人们对于政府治理能力有着更高的要求,而政府也有更大的动力做得更好。即使在非民主国家,透明的制度设计和政府运作也会提高公众对于政府的监督力度,政策制定者出于对自身声誉的关心,将更加警惕错误的发生,有责任去制定更好的经济社会政策。Besley 等^[16]针对印度的研究发现媒体更加活跃的地区同时也是干旱期间最不可能遭受饥荒的地区。这是因为媒体影响力较大的地区,政治领导人知道他们的表现将会受到公众监督,从而影响连任的可能性,因此对选民更加负责。Kaufmann 等^[17]对财政透明度和政府治理能力的关系进行了研究,发现透明度的提升将会显著增强政府治理水平。正所谓“阳光是最好的防腐剂,灯泡是最有效的警察”。信息透明一直是约束腐败行为的有效手段。Cordis 等^[18]发现,美国《信息自由法》的颁布和实施提高了腐败行为被揭露的可能。黄寿峰等^[19]基于博弈模型和实证检验发现提高我国财政透明度能够显著遏制腐败活动。

2. 开放政府数据与政府决策效率

可信和及时的数据是政策制定和实施改革的基础。以一家企业为例,经理的首要目标是提高利润。为了实现这一点,业绩要以历史为基准,并与竞争对手进行比较。同时必须对抵押品进行评估和杠杆化,以获得资金进行新的投资。如果没有数据,这些目标中的许多都无法实现。Bloom 等^[20]发现在发展中国家,帮助厂商改善库存记录保存的简单干预措施能够显著提高经营绩效。伴随着新一代信息通信技术的发展,借助人工智能、

物联网等先进技术,数据要素的深度价值挖掘成为可能,广泛可用的政府数据将会改善政府部门之间的协调性,政策设计、实施和调整将会更有效率^[21]。如税务部门的收入数据将会帮助提高财政预算的精准性。基于大数据的政策设计、模拟和评估帮助决策者锚定政策目标,并评估替代政策和措施。传统的公共政策制定和实施通常只能赋予政府领导者一些定性的判断,随着更多的数据介入,政策的预期效果将被逐步量化,科学性也会大大提升。

三、变量选择与描述性统计

(一) 变量选择与度量

1. 开放政府数据度量

本文采用世界银行的统计能力指数(Statistical Capacity Indicators)度量 OGD 水平,记为 *OGDI*。该指数是世界银行为了评估各国统计系统质量,利用一个涉及 25 项数据的评估框架所合成的一个综合性指数,不仅包括政府统计数据的数量、频率,还包括统计数据的质量,被广泛用于评估政府透明度和公共数据开放程度。具体涉及以下 3 个维度:一是源数据(Source Data)。反映一个国家是否定期进行数据收集活动,从政务系统获得的数据是否可用于统计分析,以及结论是否可靠,记为 *SDA*。二是统计方法(Methodology)。衡量一个国家遵守国际标准和公认方法的情况,对宏观经济、社会数据进行编制过程中所遵循的程序和规则进行评估,涉及统计数据质量评价,记为 *MA*。三是统计数据的周期性和时效性(Periodicity and Timeliness)。衡量关键社会经济指标的可用性和及时性,记为 *PTA*。

2. 政府治理效能度量

本文认为除了直接经济增长效应外,OGD 将通过提高政府治理效能间接促进经济增长,因此需要考虑对于机制变量即政府治理效能的合理度量。本文采用世界治理指数(Worldwide Governance Indicators)度量政府治理效能,记为 *WGI*。具体选择其政府效率(*GE_S*)、监管质量(*RQ_S*)、法治水平(*RL_S*)以及遏制腐败(*CC_S*)4 个子维度的算数平均度量政府治理效能。没有选择话语权

和问责、政治稳定和杜绝暴力进行分析的原因在于两者所描述的内容过于极端且政治色彩浓厚,容易受到不同国家和地区政治制度和文化传统等因素的影响,评估结果的严谨性和可靠性不易保证。而所选维度的界定和评估在不同国家之间已经形成了统一认识,也与政府治理关注于政府运作的内涵相一致。

3. 其他变量度量

经济增长采用 GDP($\ln GDP$, 2010 年不变价美元,取自然对数)进行度量。此外参考 Levine 等^[22],实证过程中引入劳动($\ln L$,劳动力总数的自然对数)、资本($\ln K$,固定资本的自然对数,2010 年不变价美元)、金融环境(Fin ,私人部门信贷占比)、产业结构($Indu$,工业增加值占 GDP 比重)、国际贸易($Itrade$,进出口额占 GDP 比重)和外商投资(FDI ,外国直接投资净流入占 GDP 比重)控制影响经济增长的其他因素。

(二)描述性统计

本文所需数据均来自世界银行数据库^①。实证分析的主要任务是验证 OGD 与经济增长的关系。为了保证实证样本充分,在对两者进行直接回归分析时不考虑政府治理数据,共获得 1 911 个实证样本。随后机制检验引入政府治理效能,共计获得 1 312 个实证样本。表 1 呈现了变量的基本统计特征,比例变量 $Indu$ 、 $Itrade$ 、 Fin 和 FDI 单位均为%,其中 FDI 为外商投资净流入,等于年度流入外资减去流出外资,因此可能出现负值。源数据 SDA 反映了政府定期搜集数据情况及其可靠性。从统计结果来看,国家之间差异十分明显,均值为 0.672,标准差达到了 0.192。与源数据 SDA 具有相似特征的是 MA ,反映了政府编制统计数据所依据和使用的规则、方法的科学性。平均值为 0.581,标准差为 0.215,说明不同国家在统计规则和统计方法上存在较大差异。 PTA 反映了政府数据公布的周期性和及时性,均值为 0.827,说明样本在此维度的评价较高。标准差为 0.098,意味着样本差异性较小。基于三者简单平均后的综合指

标 $OGDI$ 反映了开放政府数据的程度,平均开放水平为 0.693,标准差为 0.140,样本间的区别性十分明显。

表 1 变量统计特征描述

变量	观测数	均值	标准差	最小值	最大值
$\ln GDP$	1 911	25.192	2.667	19.720	31.106
$\ln L$	1 911	16.445	2.462	10.573	21.773
$\ln K$	1 911	23.762	2.748	17.708	29.957
MA	1 911	0.581	0.215	0.000	1.000
PTA	1 911	0.827	0.098	0.267	1.000
SDA	1 911	0.672	0.192	0.000	1.000
$OGDI$	1 911	0.693	0.140	0.167	0.989
$Indu$	1 911	27.753	10.215	1.305	84.349
$Itrade$	1 911	72.626	33.453	0.200	311.354
Fin	1 911	40.705	31.552	1.095	164.664
FDI	1 911	3.975	6.064	-37.155	103.337

四、实证结果与分析

(一)开放政府数据与经济增长关系检验

1. 基准回归结果

本文参考 Islam^[23]的宏观经济增长模型,在包含劳动力和资本的传统生产函数中引入 OGD,控制影响经济增长的可观测因素,包括产业结构、金融环境、国际贸易和外商投资,尽可能消除不可观测因素的影响,包括不随个体和时间变动的因素。表 2 汇报了 OGD 与经济增长关系的检验结果,第(1)列所呈现的内容清楚表明 OGD 产生了显著的经济增长效应。 $OGDI$ 系数在 1% 水平下显著大于零,说明开放政府数据对经济增长的促进作用非常显著。一方面,政府所拥有的数据可以直接作为交通运输、能源部门和金融行业等的直接生产要素,即数据要素。另一方面,开放数据将通过提高政府治理效能间接促进经济增长。更大力度的数据开放意味着政府透明度不断提升,公众可以利用公开信息实现对政府更为严格的监督。同时,开放数据不仅可以为私营部门所使用,政府部门也将成为政务数据价值开发的受益者。借助大数据、人工智能等先进技术,开放数据将全面提高政府运行效率、公共服务质量,保证政策制定的科学性和实施精准性。表 2 后

① 开放政府数据来自 <https://datatopics.worldbank.org/statisticalcapacity/>;政府治理数据来自 <http://info.worldbank.org/governance/wgi/#home>;控制变量数据来自 <https://data.worldbank.org/cn/>。

三列呈现了 OGD 的 3 个子维度(源数据、统计方法、周期性和及时性)对经济增长影响的检验结果。源数据反映了政府数据搜集活动的频率和搜集结果的可靠性。统计方法反映了政府编制统计数据过程中所依据的准则、方法的科学性,即统计数据质量。周期性和及时性则反映出政府数据的可用性和公开及时性。与加总指标 *OGDI* 完全一致,*MA*、*PTA* 和 *SDA* 的系数估计结果均在 1% 水平下显著大于零,说明三者和经济增长之间同样存在显著的正相关关系,意味着政府数据搜集、统计数据编制和数据公开过程均会显著影响经济增长。

表 2 OGD 与经济增长关系的基准回归结果

变量	(1)	(2)	(3)	(4)
<i>OGDI</i>	0.138 *** (4.180)	—	—	—
<i>MA</i>	—	0.056 *** (2.726)	—	—
<i>PTA</i>	—	—	0.086 *** (3.090)	—
<i>SDA</i>	—	—	—	0.053 *** (2.890)
<i>Cons</i>	18.847 *** (31.553)	18.651 *** (31.249)	18.795 *** (31.532)	18.867 *** (31.856)
控制变量	控制	控制	控制	控制
年度效应	控制	控制	控制	控制
个体效应	控制	控制	控制	控制
<i>N</i>	1 911	1 911	1 911	1 911
<i>R</i> ²	0.489	0.485	0.485	0.486
<i>F</i>	141.944	138.388	138.581	143.235

注:括号内为 *t* 值,基于稳健标准误计算。*** 表示在 $p < 0.01$ 时有统计学意义。

2. 稳健性检验

为了说明表 2 结果的稳健性,采用 4 种方式重新检验 OGD 和经济增长的关系。

第一,随着统计技术和统计能力的不断提升,世界银行会将更多的社会、经济、环境等领域数据纳入评估体系中,保证评估体系随着时间推移而不断完善。李磊等^[24]认为综合指标评估过程中不可避免地会出现评估方法和评估系统的动态调整,在应用评估结果时可以采用基于评分的样本排序反映国家和地区之间的差异。本文遵循上述操作,基于 *OGDI*、*MA*、*PTA* 和 *SDA* 分数逐年进行排名形成 *OGDI_R*、*MA_R*、*PTA_R* 和 *SDA_R* (数值越大,分数越高),重新估计 OGD 对经济增长的影

响。根据表 3 所呈现的估计结果,*OGDI_R*、*MA_R*、*PTA_R* 和 *SDA_R* 系数均在 1% 水平下显著大于零,说明即使在规避评估方法和评估体系变动影响后,OGD 依然具有显著的经济增长效应。

表 3 基于排名度量的稳健性检验

变量	(1)	(2)	(3)	(4)
<i>OGDI_R</i>	0.472 *** (5.006)	—	—	—
<i>MA_R</i>	—	0.266 *** (3.209)	—	—
<i>PTA_R</i>	—	—	0.273 *** (4.879)	—
<i>SDA_R</i>	—	—	—	0.174 *** (2.645)
<i>Cons</i>	18.737 *** (31.357)	18.596 *** (31.222)	18.940 *** (31.644)	18.773 *** (31.558)
控制变量	控制	控制	控制	控制
年度效应	控制	控制	控制	控制
个体效应	控制	控制	控制	控制
<i>N</i>	1 911	1 911	1 911	1 911
<i>F</i>	144.753	140.103	143.438	143.845
<i>R</i> ²	0.491	0.486	0.489	0.484

注:括号内为 *t* 值,基于稳健标准误计算。*** 表示在 $p < 0.01$ 时有统计学意义。为了便于展示,*OGDI_R*、*MA_R*、*PTA_R* 和 *SDA_R* 系数均乘以 1 000。

第二,由于经济增长可能存在显著的收敛性,意味着 GDP 具有自回归特征,在进行实证检验时需要在方程右侧引入前期经济增长水平,相应的普通面板数据模型将转变为动态面板模型,进而需要处理滞后被解释变量作为解释变量所引发的内生性问题。本文参考郑世林等^[25]在经济增长回归方程右侧引入 $\ln GDP$ 的一阶和二阶滞后,应用系统广义矩估计方法进行检验,表 4 汇报了 OGD 对经济增长影响的重新估计结果。*OGDI*、*MA*、*PTA* 和 *SDA* 的系数在 1% 水平下显著大于零,说明在考虑经济收敛特征和内生性问题后,OGD 依然能够显著促进经济增长。

第三,由于本文使用的是国家样本,除了常见的异方差和自相关外,国家之间的经济关联度有可能对表 2 系数推断产生影响。此处使用 Driscoll 等^[26]提出的 DK 标准误进行调整,结果如表 5 所示。由于仅对标准误进行了调整,系数估计结果不发生改变,可以看到即使考虑了自相关、异方差和截面相关,表 3 实证发现依然十分稳健,OGD 对经济增长的促进作用的确十分显著。

表 4 考虑经济收敛特征和内生性的稳健性检验

变量	(1)	(2)	(3)	(4)
<i>OGDI</i>	0.028 ** (2.531)	—	—	—
<i>MA</i>	—	0.023 *** (2.587)	—	—
<i>PTA</i>	—	—	0.051 *** (2.963)	—
<i>SDA</i>	—	—	—	0.020 ** (2.373)
<i>Cons</i>	0.150 *** (4.802)	0.168 *** (5.460)	0.109 *** (2.923)	0.148 *** (4.571)
控制变量	控制	控制	控制	控制
年度效应	控制	控制	控制	控制
个体效应	控制	控制	控制	控制
<i>N</i>	1 076	1 076	1 076	1 076
<i>Hansen (P)</i>	0.435	0.454	0.496	0.437
<i>AR(1)</i>	0.003	0.003	0.003	0.003
<i>AR(2)</i>	0.684	0.694	0.720	0.655
<i>Wald chi²</i>	1.341e+09	1.038e+09	1.476e+09	1.550e+09

注:括号内为 *z* 值,基于稳健标准误计算。***、** 分别表示在 $p < 0.01$ 、 $p < 0.05$ 时有统计学意义。

表 5 同时考虑异方差、自相关和截面相关的稳健性检验

变量	(1)	(2)	(3)	(4)
<i>OGDI</i>	0.138 *** (6.139)	—	—	—
<i>MA</i>	—	0.057 *** (3.052)	—	—
<i>PTA</i>	—	—	0.086 *** (3.245)	—
<i>SDA</i>	—	—	—	0.053 *** (4.671)
<i>Cons</i>	18.632 *** (35.602)	18.438 *** (38.952)	18.579 *** (36.342)	18.649 *** (37.423)
控制变量	控制	控制	控制	控制
年度效应	控制	控制	控制	控制
个体效应	控制	控制	控制	控制
<i>N</i>	1 911	1 911	1 911	1 911
<i>F</i>	3.361e+06	4.935e+06	5.143e+06	1.092e+06
<i>R²</i>	0.489	0.485	0.485	0.486

注:括号内为 *t* 值,基于 DK 标准误计算。*** 表示在 $p < 0.01$ 时有统计学意义。

第四,本文采用 Chernozhukov 等^[27]所发展的双重机器学习方法(Double Machine Learning)控制混淆因素和非线性关系。原因在于实证分析过程中,同时影响解释变量和被解释变量的混淆因素将干扰解释变量的系数估计结果,可能产生伪相关关系。另外,传统计量模型通常假设线性关系存在,但是不可否认的是,经济变量之间广泛存在非线性相关,模型误设可能带来内生性问题。此处将放松资本、劳动力等变量和经济增长线性相关的假设。模型设定如下:

$$Y = \theta_0 OGD + g_0(X) + U, \quad E[U|X, OGD] = 0$$
$$OGD = m_0(X) + V, \quad E[V|X] = 0 \quad (1)$$

式中, θ_0 为 OGD 对经济产出 *Y* 的边际影响。*X* 为控制变量集合,不仅包括单独影响经济增长和数据开放的因素,也包括同时影响两者的混杂因素。此处并未设定函数 g_0 和 m_0 的具体形式,而是放松了传统线性假设,采用机器学习算法估计。用于本文实证分析时,采用随机森林学习函数 g_0 和 m_0 ,控制变量集合 *X* 共计 166 个变量,包括 *lnK*、*lnL*、*Indu*、*Itrade*、*Fin*、*FDI*、国家虚拟变量(144 个)和年份虚拟变量(16 个)。表 6 给出了双重机器学习估计结果,相关系数至少在 5% 水平下显著大于零,再一次说明提高政府数据开放水平能够显著促进经济增长。

表 6 考虑混淆因素和非线性关系的稳健性检验

变量	系数	标准误	<i>t</i> 值	<i>p</i> 值
<i>OGDI</i>	1.884 ***	0.210	8.989	0.000
<i>MA</i>	1.612 ***	0.133	12.130	0.000
<i>SDA</i>	0.552 ***	0.153	3.599	0.000
<i>PTA</i>	0.642 **	0.292	2.197	0.028

注:*t* 值基于稳健标准误计算。***、** 分别表示在 $p < 0.01$ 、 $p < 0.05$ 时有统计学意义。

(二) 机制检验与效应分解

理论分析指出,OGD 对经济发展具有直接效应和间接效应。直接效应指的是政府数据将作为金融、能源、交通等领域发展的数据要素,直接拉动经济增长。间接效应则意味着数据开放将通过提高政府透明度和运行效率提升治理效能,间接促进经济增长。实证分析第一部分只是对 OGD 的经济增长效应进行了检验,进一步需要回答以下两个问题。第一,OGD 通过提高政府治理效能促进经济增长的渠道是否存在。第二,对于直接增长效应和间接增长效应,哪一种占据主导地位。此处将遵循温忠麟等^[28]通过中介效应检验流程进行实证分析。讨论实证结果之前需要对模型估计过程进行说明。首先,此处所使用的实证样本与前文不同,原因在于为了保证样本充足性,在对 OGD 和经济增长关系进行检验时并未引入政府治理数据,缺失值导致此处的实证样本少于前文。其次,传统中介效应检验通常只关注最简单的截面回归,对于更复杂模型比如面板数据模型并未涉及(需要同时控制时间效应和个体效应),因此并不能直接计算 Sobel *Z* 值。为了避免这一问题,

本文首先将样本数据减去组内均值,剔除不可观测的时间效应和个体效应。随后利用截面回归方法进行中介效应检验和计算 *Sobel Z* 值。

表 7 和表 8 汇报了中介效应检验结果,可以发现:第一,OGD 通过政府治理效能促进经济增长的中介渠道显著存在。在同时引入开放政府数据和政府治理效能后,*OGDI*、*MA* 和 *PTA* 系数在显著性不变的情况下,绝对值均有所降低,说明存在部分中介效应。*SDA* 的系数减少,显著性水平甚至超过 10%,意味着此处存在完全中介效应。对应的 *Sobel Z* 值至少在 5% 水平下显著,说明政府治理效能能在 OGD 和经济增长之间充当了显著的中介角色。第二,根据系数变化程度,可以发现除了 *SDA* 外,*OGDI*、*MA* 和 *PTA* 系数减小程度并不大,是否意味着开放政府数据的经济增长效应主要体现为直接效应而非间接影响呢?根据系数估计结果可以测算 OGD 关于经济增长的直接效应和间接效应大小,*OGDI* 直接效应占比为 77.717%,*MA* 为 86.400%,*PTA* 为 68.687%,*SDA* 为 64.444%,可以发现直接效应占据绝对主导地位。

表 7 OGD 促进经济增长的中介效应检验一

变量	OGDI			MA		
	(1)	(2)	(3)	(1)	(2)	(3)
OGDI	0.184 *** (5.360)	0.228 *** (4.247)	0.143 *** (4.318)	—	—	—
MA	—	—	—	0.125 *** (5.565)	0.091 *** (2.582)	0.108 *** (5.025)
WGI	—	—	0.178 *** (10.467)	—	—	0.181 *** (10.687)
控制变量	控制	控制	控制	控制	控制	控制
年度效应	控制	控制	控制	控制	控制	控制
个体效应	控制	控制	控制	控制	控制	控制
N	1 312	1 312	1 312	1 312	1 312	1 312
F	640.695	24.243	621.968	642.090	22.432	624.888
R ²	0.775	0.115	0.792	0.775	0.107	0.793
Sobel Z	3.935 ***			2.509 **		

注:括号内为 *t* 值,基于稳健标准误计算。***、** 分别表示在 $p < 0.01$ 、 $p < 0.05$ 时有统计学意义。

表 8 OGD 促进经济增长的中介效应检验二

变量	PTA			SDA		
	(1)	(2)	(3)	(1)	(2)	(3)
PTA	0.099 *** (2.967)	0.166 *** (3.214)	0.068 ** (2.129)	—	—	—
SDA	—	—	—	0.045 ** (2.392)	0.085 *** (2.914)	0.029 (1.619)
WGI	—	—	0.183 *** (10.755)	—	—	0.184 *** (10.813)
控制变量	控制	控制	控制	控制	控制	控制
年度效应	控制	控制	控制	控制	控制	控制
个体效应	控制	控制	控制	控制	控制	控制
N	1 312	1 312	1 312	1 312	1 312	1 312
F	628.334	23.016	612.601	626.423	22.723	611.466
R ²	0.771	0.110	0.790	0.771	0.109	0.790
Sobel Z	3.079 ***			2.813 ***		

注:括号内为 *t* 值,基于稳健标准误计算。***、** 分别表示在 $p < 0.01$ 、 $p < 0.05$ 时有统计学意义。

(三)异质性分析:信息通信技术的作用

信息通信技术的发展将从两个方面增强 OGD 的经济增长效应。第一,信息通信技术极大促进政府数据开放进程。一方面,伴随互联网,特别是移动互联的飞速发展,公众参政议政的积极性有了显著提升,对公共事务更加关注,要求获得更多的“知情权”和“参与权”。公众诉求将倒逼政府在开放数据上作出更大迈步。另一方面,信息通信技术为政府搜集数据、统计数据和公开数据提供了平台和技术保障,显著扩大数据开放的范围和提高开放数据的质量。借助互联网平台,政府拓宽了信息搜集渠道,提高了数据搜集频率,保证数据公开的及时性。如自 2020 年新冠病毒感染疫情暴发以来,在疫情信息发布、患者行程追踪等方面,信息通信技术完美展现了其在应对重大公共卫生安全事件中所发挥的巨大作用。第二,信息通信技术显著提升政府、企业挖掘数据要素价值的的能力。谢康等^[29]指出,数据资源向数据要素转变需要以企业自身的数据分析能力、配套信息通信技术基础设施为必要条件。信息通信技术除了带来了大量数据资源,同时为数据要素的价值挖掘提供了工具。“数据+AI”早已成为材料科学、量子物理、生物医学等领域提高研发效率的标准配置^[30],推动知识创造效率持续提升,而知识又是决定经济增长中产出效率(全要素生产率)的关键要素。除此之外,利用信息通信技术,政府在气象信息发布、交通规划、经济预测等方面的能力即数据要素价值开发能力也将获得较大提升,进一步提高治理效能。

为了识别信息通信技术在 OGD 促进经济增长过程中扮演的重要作用,此处利用国家信息通信产品出口(信息通信产品出口占产品出口总量百分比)度量一个国家信息通信技术能力,根据中位数将样本分为高信息通信技术和低信息通信技术水平,分组估计开放政府数据和经济增长的关系,表 9 呈现了相关系数估计结果。当采用 *OGDI* 度量开放政府数据时,高信息通信技术水平样本系数为 0.225,低水平样本系数为 0.190,虽然均在 1% 水平下显著大于零,但是前者明显大于后者,说明信息通信技术水平较高的国家,开放数据对经济增长的促进作用更大。相似的发现还体现在以 *MA* 和 *PTA* 度量数据开放的回归中,前者两组系

数分别为 0.104 和 0.047,后者为 0.255 和 0.044(未获得统计显著性)。值得注意的是,源数据维度展

现了完全相反的实证结果,低信息技术水平的样本,源数据对于经济增长的促进作用更大。

表 9 信息通信技术影响 OGD 经济增长效应的检验结果

变量	OGDI		MA		PTA		SDA	
	高	低	高	低	高	低	高	低
OGDI	0.225*** (3.383)	0.190*** (4.159)	—	—	—	—	—	—
MA	—	—	0.104* (2.101)	0.047* (2.103)	—	—	—	—
PTA	—	—	—	—	0.255*** (5.158)	0.044 (1.164)	—	—
SDA	—	—	—	—	—	—	0.034 (1.506)	0.098*** (5.200)
Cons	17.863*** (20.263)	15.123*** (22.574)	17.190*** (22.048)	15.331*** (23.288)	17.606*** (21.881)	15.388*** (24.166)	17.504*** (19.915)	15.263*** (25.205)
控制变量	控制	控制	控制	控制	控制	控制	控制	控制
年度效应	控制	控制	控制	控制	控制	控制	控制	控制
个体效应	控制	控制	控制	控制	控制	控制	控制	控制
N	798	788	798	788	798	788	798	788
F	2.882e+06	8.195e+04	1.495e+06	1.042e+05	3.699e+05	5.045e+05	1.290e+06	1.692e+05
R ²	0.948	0.922	0.947	0.919	0.949	0.919	0.945	0.922

注:括号内为 t 值,基于 DK 标准误计算。***、* 分别表示在 $p < 0.01$ 与 $p < 0.1$ 时有统计学意义。

五、研究结论与政策建议

本文从经济增长角度分析和评估开放政府数据的价值。理论分析认为:一方面,OGD 为社会经济活动提供生产、研发等所必须的数据要素,直接推动经济增长;另一方面,通过公开数据,政府透明度和治理效率将获得显著提升,所产生的政府治理效能间接促进经济增长。为了对理论判断进行检验,本文采用世界银行评估数据度量政府数据开放水平,利用 2004—2019 年 1 911 个国家样本进行实证检验,获得以下主要研究结论。第一,开放政府数据显著促进经济增长,直接经济增长效应占据主导地位。直接检验结果表明,OGD 是拉动经济增长的重要因素。这一发现在一系列稳健性检验下均保持一致。中介效应检验结果表明 OGD 不仅具有直接经济增长效应,还将通过提高政府治理效能间接促进经济增长,但是前者明显更为显著。第二,信息通信技术显著影响 OGD 的经济增长效应。本文研究发现,信息通信技术越发达,OGD 的经济增长效应越强,意味着政府数据开放进程中,需要关注技术与数据的匹配性。因此,根据研究内容与发现提出 4 点政策建议。

第一,加快推动开放数据立法,完善公共数据开放的制度建设。加快推动开放政府数据方面的立法工作,是持续深入完善公共数据开放,发挥政府数据要素价值的前提所在。一方面,要重视公众的信息自由权,在确保公众获取信息自由以及信息隐私权方面进行立法保障。在信息获取和公

开方面,要对企业与政府行为进行严格规定,明确划定收集信息的界限,并规定公民可以自由获取的信息范围,对相关信息进行依法公开。另一方面,建立健全的政府数据开放秩序与规定。明确规定可开放的政府数据目录,搭建数据存储库,进行数据在线开放。

第二,加快培育数据要素市场,挖掘公共数据要素价值。一方面,需要发展多元的数据供给主体,加快构建面向企业、个人等社会主体的公共数据库。另一方面,需要刺激市场对于数据要素的需求,充分挖掘数据的市场价值。政府应该因地制宜,根据地方产业布局采取针对性措施,刺激不同行业的数据需求,实现数据价值链的市场化,通过对接企业数据需求特征,制定差异化数据供给策略,助力企业数字化转型。

第三,加强数据安全治理,防范公共数据开放风险。伴随数据要素价值的愈加凸显,政府数据开放特别是经济数据公开面临的安全风险与日俱增,突出表现为数据贩卖引致隐私权被侵害、数据跨境流动引发国家安全隐患、高附加值数据泄露频发等特征。鉴于《中华人民共和国数据安全法》推出时间不长,未来我国数据安全防范仍然面临较大改进空间。长期来看,完善数据安全法制建设、加强数据安全治理、推动数据安全领域发展、推进全球数据安全治理以及提升国际话语权将成为我国数据安全治理的重要着力点。

第四,完善数据开放平台建设,保障公共数据

获取通畅与质量。一方面需要设置统一的元数据标准,包括收集标准、存储标准、下载标准和使用标准,减轻平台终端压力的同时,提高数据获取便利性和保证数据质量。另一方面,各级政府需定期将各业务部门提交的数据集进行整合,集中设置数据目录,在各目录下进一步提供数据资源获取链接,实现数据源头可追溯。同时,要对数据目录以及各类数据进行动态更新,保证数据的时效性。

参考文献:

- [1] SAXENA S, MUHAMMAD I. Barriers to use open government data in private sector and NGOs in Pakistan [J]. Information discovery and delivery, 2018, 46(1): 67-75.
- [2] 付熙雯,郑磊. 开放政府数据的价值:研究进展与展望[J]. 图书情报工作, 2020, 64(9): 122-132.
- [3] UBALDI B. Open government data: towards empirical analysis of open government data initiatives [Z]. OECD Working Paper, 2013.
- [4] FUNG A. Infotopia: unleashing the democratic power of transparency [J]. Politics & society, 2013, 41(2): 183-212.
- [5] 谭海波,张楠. 政府数据开放:历史、价值与路径[J]. 学术论坛, 2016, 39(6): 31-34, 53.
- [6] CHARALABIDIS Y, LOUKIS E, ALEXOPOULOS C. Evaluating second generation open government data infrastructures using value models [C] // 2014 47th Hawaii international conference on system sciences. New Jersey: IEEE Press, 2014: 2114-2126.
- [7] STUERMER M, DAPP M M. Measuring the promise of open data: development of the impact monitoring framework [C] // 2016 Conference for e-democracy and open government (CeDEM). New Jersey: IEEE Press, 2016: 197-203.
- [8] MAGALHAES G, ROSEIRA C. Open government data and the private sector: an empirical view on business models and value creation [J]. Government information quarterly, 2020, 37(3): 34-44.
- [9] JONES C I, TONETTI C. Nonrivalry and the economics of data [J]. American economic review, 2020, 110(9): 2819-2858.
- [10] MCAFEE A, BRYNJOLFSSON E, DAVENPORT T H, et al. Big data: the management revolution [J]. Harvard business review, 2012, 90(10): 60-68.
- [11] AKCIGIT U, LIU Q. The role of information in innovation and competition [J]. Journal of the European economic association, 2016, 14(4): 828-870.
- [12] ZHU C. Big data as a governance mechanism [J]. Review of financial studies, 2019, 32(5): 2021-2061.
- [13] FARBOODI M, MATRAY A, VELDKAMP L. Where has all the big data gone? [Z]. SSRN Working Paper, 2018.
- [14] PISTOR K, RAISER M, GELFER S. Law and finance in transition economies [J]. Economics of transition, 2000, 8(2): 325-368.
- [15] 刘小鲁. 知识产权保护、自主研发比重与后发国家的技术进步[J]. 管理世界, 2011(10): 10-19, 187.
- [16] BESLEY T, BURGESS R. The political economy of government responsiveness: theory and evidence from India [J]. The quarterly journal of economics, 2002, 117(4): 1415-1451.
- [17] KAUFMANN D, KRAAY A. Growth Without Governance [Z]. World Bank Working Paper, 2002.
- [18] CORDIS A S, WARREN P L. Sunshine as disinfectant: the effect of state freedom of information act laws on public corruption [J]. Journal of public economics, 2014, 115: 18-36.
- [19] 黄寿峰,郑国梁. 财政透明度对腐败的影响研究——来自中国的证据[J]. 财贸经济, 2015(3): 30-42.
- [20] BLOOM N, MAHAJAN A, MAS A. Do management interventions last? evidence from India [J]. American economic journal: applied economics, 2020, 12(2): 198-219.
- [21] BINSWANGER J, OECHSLIN M. Disagreement and learning about reforms [J]. The economic journal, 2015, 125(584): 853-886.
- [22] LEVINE R, NORMAN L, THORSTEN B. Financial intermediation and growth: causality and causes [J]. Journal of monetary economics, 2000, 46(1): 31-77.
- [23] ISLAM N. Growth empirics: a panel data approach [J]. The quarterly journal of economics, 1995, 110(4): 1127-1170.
- [24] 李磊,马欢. 数字政府能否留住外资? [J]. 中山大学学报(社会科学版), 2020, 60(4): 183-194.
- [25] 郑世林,周黎安,何维达. 电信基础设施与中国经济增长[J]. 经济研究, 2014, 49(5): 77-90.
- [26] DRISCOLL J C, KRAAY A C. Consistent covariance matrix estimation with spatially dependent panel data [J]. Review of economics and statistics, 1998, 80(4): 549-560.
- [27] CHERNOZHUKOV V, CHETVERIKOV D, DEMIRER M, et al. Double/debiased machine learning for treatment and structural parameters [J]. The econometrics journal, 2018, 21(1): 1-68.
- [28] 温忠麟,叶宝娟. 中介效应分析:方法和模型发展[J]. 心理科学进展, 2014, 22(5): 731-745.
- [29] 谢康,夏正豪,肖静华. 大数据成为现实生产要素的企业实现机制:产品创新视角[J]. 中国工业经济, 2020(5): 42-60.
- [30] AGRAWAL A, MCHALE J, OETTL A. Finding needles in haystacks: artificial intelligence and recombinant growth [C] // The economics of artificial intelligence: an agenda. Chicago: University of Chicago Press, 2018: 149-174.

(本文责编:辛 城)