

公共数据开放与乡村振兴

——来自政府数据平台上线的准自然实验

辛冲冲¹ 袁一杰² 韩韵格³

(1.北京市社会科学院《北京社会科学》编辑部,北京 100101;2.中国社会科学院大学 应用经济学院,北京 102488;3.湖北文理学院 经济管理学院,湖北 襄阳 441053)

摘要:政府数据要素作为一种重要的公共资源,具有高效整合信息的优势,能够有效破解乡村振兴中的诸多难题。本文基于2010—2022年我国地级市面板数据,以政府数据平台上线为准自然实验,采用双差分方法实证检验公共数据开放对乡村振兴的影响及其作用路径。研究结果表明,公共数据开放释放了数字红利,可以有效促进乡村振兴,多项稳健性检验支持了这一结论。机制分析显示,公共数据开放通过增强民生性财政支出强度和推动市场一体化,推动乡村振兴。异质性分析发现,在第一产业占比较高、数字基础设施相对完善的地区,公共数据开放对乡村振兴的促进作用更大。基于此,本文提出持续推动政府数据平台的建设和优化、增加民生性财政支出和推动市场一体化、因地制宜地打造乡村特色产业和配套数字基础设施建设等发展思路,为推动乡村振兴提供政策参考。

关键词:公共数据开放;乡村振兴;民生性财政支出;市场一体化;数据要素

中图分类号:F327 **文献标识码:**A **文章编号:**1003-5230(2025)03-0123-12

一、引言

2020年颁布的《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》明确将数据要素与土地、劳动力、资本和技术并列列为五大生产要素。作为新型生产要素,数据要素已成为我国数字经济建设和经济高质量发展的基础性资源和战略性资源。乡村振兴是实现经济高质量发展的关键环节,不仅关系到农业农村现代化的全局进程,更是推动区域协调发展、缩小城乡差距的重要一环。有效利用数据这一新型生产要素已成为推动乡村振兴的重要抓手。数据要素的开放和应用不仅有助于提升乡村产业的发展水平和优化乡村治理体系,而且能通过信息化提升农业生产力 and 重塑生产关系,优化资源配置,特别是信用服务、财税金融、环境与资源保护等多类数据要素的整合与供给,可以大幅降低农民获取信息的成本,提高农业生产的科学性,为乡村振兴注入强大动力。长期以来,

收稿日期:2024-12-17

基金项目:北京市社会科学基金项目“疫后‘减税政策’对北京中小企业价值形成的作用机制研究”(20GLC038)

作者简介:辛冲冲(1988—),男,河北邢台人,北京市社会科学院《北京社会科学》编辑部助理研究员;

袁一杰(1997—),男,四川成都人,中国社会科学院大学应用经济学院博士生;

韩韵格(1991—),女,湖北襄阳人,湖北文理学院经济管理学院讲师,本文通讯作者。

农村产业具有周期长、回报率低等特征,资本难以在农村依靠数据投入获得丰厚回报,进而影响数据要素在农村的流动性。既有研究表明第一产业利用数据要素进行生产难以弥补其投入成本,且较长的生产周期还会进一步压缩乡村产业的获利空间^[1],所以仅依靠私有数据要素推动乡村振兴仍面临较大挑战。具有公共资源属性的政府数据或许能成为解决这一问题的关键所在,其能够有效缓解数据要素在乡村供给不足的问题。据此,本文关注政府公共数据开放对乡村振兴的影响及其作用机制。

目前学术界对政府公共数据开放的研究还处于起步阶段。方锦程等认为公共数据开放有助于破除区域信息壁垒和弥合资源禀赋差距,进而促进城市和乡村的协调发展^[2]。蔡运坤等的研究表明,公共数据开放能够激发地级市创业活力^[3];而沈艳等的研究表明,县级城市同样能够借助数据开放推动数字创业,继而促进经济高质量发展^[4]。此外,王晓丹等指出,各地根据城市发展差异来优化政府公共数据平台内容,可以提升政府数据开放的质量与效率,进而有效推动数字经济与实体经济融合^[5]。数字经济与农业经济的有效联动,可以为农业等传统行业注入数字动能^[6]。综上可知,政府公共数据开放在乡村振兴中的作用研究尚不充分,但其所蕴含的潜力和红利不容忽视,对此,研究政府公共数据开放如何充分发挥数据要素的优势,推动乡村振兴极具现实意义。

本文将政府数据平台上线作为准自然实验,采用双重差分法实证检验公共数据开放对乡村振兴的影响及其作用机制,理由是:一方面据统计,我国超过80%的公共数据被政府掌握^①,这表明政府数据开放在释放数据红利、产生经济效益的过程中具有天然的主体优势;另一方面,自2012年以来,各地级市政府陆续上线数据平台,这为本文识别公共数据开放的实际影响以及探索其推动乡村振兴的具体路径提供了契机。本文的边际贡献在于以下三个方面。第一,本文将数据要素与乡村振兴问题相联系,为推动乡村振兴提供了新的研究视角。现有研究主要关注劳动力、资本、技术和土地等传统生产要素对乡村振兴的影响,而对数据要素与乡村振兴之间关系的探讨相对不足,且缺乏实证研究。实际上,公共数据要素因其独特的公共资源属性,理论上是推动乡村振兴的重要力量,应成为研究乡村问题的重要关注点。第二,本文构建了公共数据开放影响乡村振兴的理论框架,并揭示了其作用机制,丰富了乡村振兴的研究内容。第三,本文进行了多种异质性检验,探讨了地区主要产业形态和数字基础设施建设水平对公共数据开放效果的影响,从而为不同地区因地制宜地优化政策提供了直接依据。

二、政策背景与理论分析

(一)政策背景

为推动数据要素的广泛应用并充分发挥其价值红利,2022年《中共中央 国务院关于构建数据基础制度更好发挥数据要素作用的意见》要求合理降低市场主体获取数据的门槛,增强数据要素的共享性和普惠性。实际上,早在2015年发布的《促进大数据发展行动纲要》就已指明数据的重要性,并针对我国数据发展中的问题提出了诸多推动政府数据开放共享的要求;2016年出台的《政务信息资源共享管理暂行办法》对非涉密数据进行了界定,这加快了政府数据公开的步伐;2019年颁布的《国家政务信息化项目建设管理办法》要求公开区市级部门的政务信息,扩大了数据覆盖范围;2022年出台的《国务院关于加强数字政府建设的指导意见》明确要求加强政府数据平台建设,为数据要素赋能数字经济发展提供了制度保障。这些政策充分体现出我国对数据要素的高度重视,也彰显了国家推动政府数据平台建设的决心。

政府数据平台是指将政府信息库以数据形态通过政府授权网站集中进行呈现的媒介。公共数据开放则是指放开数据集合使用的权限,让社会个体和企业均能够根据需求无偿从政府数据平台获取数据资源。政府数据平台开放的内容涉及面广,包括经济建设、信用服务、财税金融、交通服务、政府机构与社会团队、环境与资源保护、农业农村等主题,逐步推进数据资源共享和政务信息化建设。最早在2012年,北京市和上海市相继进行政府数据平台试点,长沙市、佛山市和深圳市等城市紧随其后,政府数据平台上线数量逐年增长。《促进大数据发展行动纲要》和《政务信息资源共享管理暂行办

法》对推动政府数据平台上线发挥了重要作用。截至 2023 年底,我国共有 204 个城市推出了政府数据平台^②。由此可见,以政府数据平台为载体,公共数据开放可以向社会提供大量数据要素,对促进经济发展具有重要意义。

(二)理论分析

乡村振兴是实现共同富裕的重要路径,党的十九大报告提出了乡村振兴战略的二十字方针,即“产业兴旺、生态宜居、乡风文明、治理有效、生活富裕”。因此,研究公共数据开放对乡村振兴的影响,应结合产业、生态、治理等多维视角,全面探讨其在推动乡村振兴中的作用。

1.直接效应

第一,公共数据开放有利于推动乡村产业的科学发展。一方面,政府数据平台公布了自然资源、环保、气象等方面的数据,这有利于农业生产者提前应对潜在风险,使乡村经济活动降低因自然因素受到的破坏,提高农业生产效率。农业生产高度依赖自然条件和生物的生长特性,而龙卷风和洪涝灾害等自然灾害频发对农业生产具有不可忽视的冲击^[7]。公共数据要素共享不仅包括年度数据,而且提供实时数据、发布检测报告,为农业生产者提供全面的数据参考。例如,宿迁市公共数据开放平台同时整合了天气状况、地震和干旱等自然灾害的预警及应对措施等信息^③。另一方面,政府数据平台通过整合土壤、水文和气象等数据,构建智慧农业体系,使农业生产更具科学性。例如,江苏省丰县采集涉农数据信息,构建了智慧农村可视化联动管控、农产品价值检测预警等数据系统,收集本地涉农数据 3.4 亿条。这种农业数据的共享互通,有利于帮助农民优化种植模式,实现农业生产的精细化管理^④。

第二,公共数据开放有利于改善乡村生态环境建设。政府数据平台将分散于不同政府工作报告中的治理数据进行整合和公布,提升了地区生态环境信息的透明度。例如,成都市、茂名市和深圳市等^⑤多个官方公共数据开放平台均对各区(县)绿化程度、农村公共厕所建设和改造情况及生活污水处理进度等信息进行了公示。为避免公开数据披露的短板引发社会广泛关注或质疑,各级政府通常会加大对生态宜居工程的投入力度,主动提高绿化覆盖率、公共厕所建设数量,并提升生活污水处理能力和效率,为乡村振兴提供了坚实的生态基础。同时,当前政府数据平台开放的海量公共数据,涵盖了乡村生态系统的各个方面,包括生产智能化、灾害预警与风险管理、供应链优化、农村治理与公共服务等,以更高定位、更高标准、更大数量级,将各类数据汇集共享,把各种应用场景纳入,从而为政府相关部门在制定乡村生态保护政策、规划和措施时提供全面、准确、科学的依据,使决策更加符合乡村生态环境的实际情况和发展需求,提高决策的科学性和有效性。

第三,公共数据开放有利于促进乡村文明进步。当前,大数据平台下的“数商兴农”已成为推动农业农村现代化发展的重要动力,扩展满足农民需求的数字乡村应用场景。农民利用网络新业态,重塑乡村文化自信与创造力,将乡村文化融入文化产业,通过推动乡村旅游、手工艺品销售等传统业务向平台经济转化,既促进了农村电子商务的发展,也实现了乡村文化的数字化传播。同时,政府数据开放也会促进乡村治理的多元主体参与,吸引更多流入人口、社会资本和外来企业等其他主体参与乡村文明建设,所营造的氛围也会进一步强化乡村农民的主动参与意识,激发他们参与乡村建设的积极性与主动性,从而丰富乡村文明,为实现乡村全面振兴提供了重要支撑。

第四,公共数据开放有利于提高乡村治理效率。乡村地区面临“信息孤岛”“数据壁垒”等问题,缺乏有效的数据共享平台,数据流通不畅,难以实现互联互通,从而制约数据价值的充分发挥。政府数据平台可以有效破解数字资源整合共享难题,通过整合各类公共服务资源,提供一站式服务,有利于方便群众办事,实现“数据多跑路,群众少跑腿”,从而提升服务效率和便捷度,进而增强乡村人民群众的获得感和幸福感。例如,广东省“粤省事”平台建设了三农服务专区,以“便农惠农”为目标,涵盖三农信息发布、业务办理、信息查询、金融服务等业务,提供惠农便农服务。同时,线上平台还能实现治理任务的合理派发与跟踪管理,确保各项政策与措施得到有效执行。政府数据平台通过为乡村治理

提供全面、准确的数据支撑,推动治理模式从“经验驱动”向“数据驱动”转变,不断提升治理决策的科学性和精确性,从而为乡村振兴提供有力支撑。

第五,公共数据开放有利于提升乡村生活富裕水平。政府数据平台的上线,一方面接入农机补贴、涉农政策等信息,帮助农民通过政策文件列表及时了解最新涉农政策和福利,改善农户生产条件。另一方面通过实时记录全国农产品市场各时期不同品类的价格等信息,打破了地理因素造成的信息壁垒,缓解了销售过程中的信息不对称问题。例如,武汉市不仅公布了各类水产品、蔬菜、生猪和鸡蛋等的出厂价格,还对最大的“菜篮子”白沙洲市场中的农副产品价格进行了详细披露;无锡市对各个区的最大市场中的农副产品价格也进行了详细披露^⑥。这类农产品价格信息的透明化,使农业生产者能够更为科学地规划生产决策,从而提高收入。

根据以上分析,本文提出研究假设 1:公共数据开放有利于乡村振兴。

2.间接效应

第一,公共数据开放有利于增强民生性财政支出强度,助力乡村振兴。成都市、武汉市和无锡市等均在公共数据平台公布了市本级和其辖区区县的各项财政收入和支出。这类数据要素主要通过两个途径发生作用:一是通过公众监督强化民生性财政支出;二是通过政府间行为模仿提高民生性财政支出的效率和规模。一方面,公共数据开放降低了公众监督政府行为的信息成本,有助于优化民生性财政支出。公众能够以较低成本获得这些数据,从而履行监督政府行为的权利^[8]。民生性财政支出直接关系到公众福利,所以公众对其规模、用途及占比尤为关注。政府数据平台提高了财政透明度、缓解了信息不对称,提升了公众行使监督权的效率,并对政府行为偏好产生了积极导向作用^[9]。面对公众需求,政府将更加重视民生问题,着力于优化财政资源配置,开展共建共治共享的乡村治理,注重财政支出的精准发力。另一方面,公共数据开放还能推动政府间的行为模仿,进而优化民生性财政支出。地方官员的晋升机制注重实绩原则,这有助于鼓励官员积极履职,服务于民生。而民生性财政支出直接关系到社会治理的成效,是地方政府工作绩效考核的重要组成部分。为应对晋升锦标赛带来的压力,地方政府会学习其他政府在财政支出方面的成功经验以优化自身财政支出结构^[10]。政府数据平台上线为这种政府间的模仿行为提供了良好契机,使地方政府能够迅速获取其他地区的财政支出结构数据,尤其是民生性财政支出的具体分配细节。这样地方政府便能及时调整支出策略,提高民生性财政支出的占比。此外,地方政府间的竞争还会形成“攀比效应”^[11],有助于推动民生性财政支出的进一步增长。

民生性财政支出是政府实现收入分配职能的重要手段,也是实现乡村振兴的基础保障。福利经济学理论强调社会资源的公平分配,通过增加民生性财政支出,能有效减少城乡之间的教育、医疗和社会保障差异,进而推动乡村振兴。具体来看,教育支出是推动农村教育供给侧改革的主要资金来源,不仅有助于优化乡村劳动力市场^[12],还通过培育高素质农民,推动乡风文明建设,为乡村振兴提供了人才支持^[13];医疗支出聚焦于强化农村公共卫生服务和建设健康乡村,有效助力乡村实现生活富裕,为乡村振兴提供了健康支持^[14];社会保障和就业支出则为乡村劳动力提供了后备支持,激发了乡村劳动力的就业积极性,为实现乡村生活富裕提供重要基础,进而有利于推动乡村振兴^[15]。

第二,公共数据开放有助于加速市场一体化的进程,助力乡村振兴。市场一体化指包括城市与乡村在内的各种生产要素实现统一流动与高效配置。政府数据平台将各类重要交易信息整合后公开,有利于补齐乡村信息极度匮乏的短板,推动了城市和乡村市场的紧密联系,为构建市场一体化奠定了基础。该平台不仅提供了农产品价格信息,还公布了企业经营异常名录、知识产权类案件和商事主体基本信息等与企业信誉相关的重要数据。例如,深圳市公布了行政处罚信息、企业经营异常名录和法人等基本信息;成都市公布了行政处罚信息、企业经营异常名录等基本信息。一方面,透明化的农产品价格信息有利于优化营商环境,进而有助于市场一体化。在传统交易中,农产品常因信息匮乏而被城市企业压低价格,而城市企业也会因获取信息不畅可能遭遇不合理的抬价行为。而政府数据平台提供了实时、透明的农产品价格信息,使双方能够基于公开的价格信息合理交易,避免了价格博

弈中信息不对称引致的不公平行为,进而激发城市和农村共同构筑市场一体化的积极性。另一方面,企业信誉信息的透明化能帮助农业经营主体规避交易风险,进一步提高了合作的稳定性。农产品期货交易具有长周期的特性,在实际交割阶段才结算交易金额,因而收购方的信誉问题可能导致卖方承受较大风险。政府数据平台通过整合并披露企业经营异常信息和信誉状况,有利于农业经营主体在交易前准确评估对方的履约能力,降低了未知风险造成的经济损失。这种信息透明既有利于增强农业经营主体的市场信心,又能进一步激发其参与市场合作的意愿,为城市和乡村产业合作发展注入动力。

推动市场一体化发展对乡村振兴具有重要意义。城乡市场一体化建设有助于乡村经济追赶城市经济,进而实现乡村产业兴旺和生活富裕。一方面,市场一体化扩大了农产品的有效需求,推动农民增加农业经营规模和经济作物种植面积^[16],实现乡村产业兴旺,带动农民收入增长,推动乡村振兴。另一方面,市场一体化促进了农村消费增加,改善了农村居民的生活质量,创造了消费新场景,促进农村居民消费升级^[17],从消费端实现生活富裕,进而推动乡村振兴。据此,本文提出以下研究假设。

假设 2:公共数据开放通过增强民生性财政支出强度推动乡村振兴。

假设 3:公共数据开放通过促进市场一体化推动乡村振兴。

三、研究设计

(一)实证策略与模型设定

本文借助各个城市政府数据平台上线,研究公共数据开放对乡村振兴的作用效果。具体地,将各个城市政府数据平台上线的时间和个体差异看作准自然实验,采用地级市层面的样本,基于双重差分模型进行实证分析。具体模型设定如下:

$$Ruralrev_{i,t} = \alpha + \beta Pubdata_{i,t} + \theta Control_{i,t} + \gamma_i + \delta_t + \epsilon_{i,t} \quad (1)$$

式(1)中, $Ruralrev_{i,t}$ 表示 t 年 i 地级市的乡村振兴指数; $Pubdata_{i,t}$ 表示 t 年 i 地级市是否为公开数据城市且政府数据平台已上线,其系数 β 衡量的是政府数据平台上线带来的净效应,为本文的核心关注点; $Control_{i,t}$ 表示地级市层面的一系列控制变量。 γ_i 为地区固定效应, δ_t 为年份固定效应, $\epsilon_{i,t}$ 为随机误差项。

(二)变量选取

1.被解释变量:乡村振兴指数(Ruralrev)

本文参考徐雪和王永瑜的方法测算乡村振兴水平:将产业兴旺、生态宜居、乡风文明、治理有效和生活富裕分别作为二级子标,再分别下设三级子标,采用熵值法将上述指标进行汇总计算得到乡村振兴指数^{[18] ⑦}。

2.解释变量:公共数据开放(Pubdata)

公共数据开放地区和时间是构成公共数据开放变量的重点,为确保该解释变量具有可信度,本文通过多个步骤对相关信息进行搜集。鉴于复旦大学数字与移动治理实验室和华中师范大学信息管理学院分别发布了《2023 中国地方公共数据开放利用报告(城市)》和《中国政府开放数据利用研究报告》,我们进行如下信息整理:首先,交叉参考已公布的两项资料对政府数据平台上线的地区和时间进行整理和核对;其次,将整理结果逐一通过政府官网进行核实,保障被包含在内的每个地区政府数据平台已上线以及上线时间的正确性;最后,在百度等搜索引擎中采用城市名+“数据开放”“数据资源”“公共数据”,对排除在外的城市进行系统性搜索。截至 2022 年,最终获得共 187 个城市的政府数据平台上线时间(包括地级市、副省级城市和直辖市),用 $treated$ 区分某地区是否为处理组,若在样本年限截止前政府数据平台上线则为处理组并设定为 1,否则设定为 0;用 $post$ 的取值区分平台是否上线,若在政府数据平台上线当年及之后则设定为 1,否则设定为 0。 $Pubdata$ 变量为 $treated$ 和 $post$ 的交乘项,若该地区是公开数据城市且政府数据平台已上线, $Pubdata$ 为 1,否则为 0。

3.控制变量

为缓解遗漏变量可能带来的影响及提高回归模型的估计精度,本文参考现有研究的普遍做法,控制了乡村振兴的主要影响因素,具体包括:外商投资水平(Fdi),用当年实际利用外资金额与 GDP 的比值表示;信息基础设施建设(Infrom),用邮政业务总量的自然对数表示;金融业发展水平(Finandev),用金融机构年末存贷款余额与 GDP 的比值表示;人力资本水平(Human),用在校大学生数与年末人口总数的比值表示;财政支持强度(Govinter),用一般性财政支出与 GDP 的比值表示;产业结构(Instru),用第三产业增加值与第二产业增加值的比值表示。

(三)数据说明

本文采用的数据年份为 2010—2022 年。为使分析更具有可信度,剔除了数据缺失较多的地级市,最终本文获得了 287 个城市的数据,样本量为 3731。政府数据平台上线地区和时间相关数据来自《(2023 中国地方公共数据开放利用报告(城市))》《中国政府开放数据利用研究报告》以及百度引擎搜索,将得到的结果进行交叉验证。其余经济变量均来源于《中国农村统计年鉴》《中国城乡建设统计年鉴》《中国民政统计年鉴》和历年各个地级市统计局公布的统计年鉴、国民经济和社会发展公报以及地方政府工作报告等。表 1 为本文主要变量的描述性统计结果。

表 1 主要变量的描述性统计结果							
变量分类	变量名称	变量符号	样本量	平均值	标准差	最小值	最大值
被解释变量	乡村振兴	Ruralrev	3731	0.336	0.118	0.059	0.845
	公共数据开放	Pubdata	3731	0.239	0.427	0.000	1.000
	外商投资水平	Fdi	3627	0.161	0.169	0.000	1.978
控制变量	信息基础设施建设	Infrom	3724	10.853	1.369	7.209	16.795
	金融业发展水平	Finandev	3725	2.552	1.282	0.588	21.301
	人力资本水平	Human	3640	0.019	0.024	0.000	0.145
	财政支持强度	Govinter	3718	0.202	0.103	0.035	0.872
	产业结构	Instru	3800	1.070	0.620	0.109	5.929

四、实证分析

(一)基准回归结果分析

表 2 为政府公共数据开放对乡村振兴的基准回归结果,无论是否引入控制变量,核心解释变量 Pubdata 的回归系数均在至少 5%的水平上显著为正,这表明公共数据开放带来了信息红利,给乡村生产和生活注入新动能,促进了乡村振兴,本文研究假设 1 成立。

(二)平行趋势检验

平行趋势检验通过是双重差分法有效识别因果效应的重要前提条件。为此,借鉴 Chen 等、高跃光等的思路^{[19][20]},设定如下模型:

$$\begin{aligned} \text{Ruralrev}_{i,t} = & \alpha + \beta_t \sum_{t=-5}^6 \text{test}_{i,t} + \theta \text{Control}_{i,t} \\ & + \gamma_i + \delta_t + \epsilon_{i,t} \end{aligned} \tag{2}$$

式(2)中,test_{i,t}为 i 地级市距离政府数据平台最初上线时刻的相对时间,t=0 为上线当年。本文将距离数据平台上线 5 期以前的变量累积到第 5 期,数据平台上线 6 期以后的变量累积到第 6 期,且均以上线前一期作为基期。根据图 1 发现,与基准年份相比,公共数据开放前几年的回归系数均在 90%的置信区间显著接近于 0。这表明,在公共数据开放之前,上线地区和非上线地区的乡村振兴水平具有相同的变化趋势,通过了平行趋势检验。同时,从图

表 2 基准回归结果

变量	(1)	(2)
	Ruralrev	
Pubdata	0.004 ** (0.002)	0.007 *** (0.002)
Fdi		- 0.014 ** (0.006)
Infrom		0.003 ** (0.001)
Finandev		0.006 * (0.003)
Human		- 0.086 (0.139)
Govinter		0.065 *** (0.022)
Instru		0.001 (0.003)
地区固定效应	控制	控制
年份固定效应	控制	控制
观测值	3731	3613
R ²	0.956	0.957

注:***、**、* 分别表示在 1%、5%和 10%水平上显著,括号内的数值为 t 值,下表同。本文回归分析中的稳健标准误均聚类到城市层面。

1 还可以发现,公共数据开放对乡村振兴的促进作用具有一定的时滞,从上线后的第 1 期(即第 2 年)才开始显现。这可能是因为农业生产周期较长,数据要素发挥作用需要一定时间。此外,公共数据开放的效果具有较长的持续性,在第 6 年依旧具有推动乡村振兴的效果。

(三)内生性检验

实际上,政府数据平台上线不具有完全随机性,可能与地区数字经济的发展水平有关,所以采用双重差分法进行研究,仅能克服部分内生性问题。对此,本文寻找工具变量以克服这一问题。本文参考黄群慧等的研究^[21],从历史角度选取 1984 年邮电业务总量作为工具变量(IV)的基础性指标。这样构造工具变量具有一定的合理性:一是满足相关性,地方历史邮电业务总量与数字基础设施建设水平相关,同时也反映出当地政府对数字基础设施建设的重视程度,地方政府对其的重视程度越高越有可能搭建政府数据平台;二是满足外生性,1984 年的邮电业务数据不直接与近年乡村振兴相关,且不会受到乡村振兴的影响。此外,各地级市的邮电业务总量数据不具有时变特征,因而本文将 1984 年邮电业务总量与年份变量交乘处理后再进行回归分析。表 3 汇报了工具变量回归结果。第(1)列为将 1984 年邮电业务总量与年份的交互项作为工具变量的第一阶段回归结果,工具变量对公共数据开放的影响显著为正,这符合本文预期。同时,第一阶段回归的 F 统计量远大于 10,C-D Wald F 统计量远大于 16.39,表明工具变量与内生变量具有密切的关系,不存在弱工具变量问题。第(2)列为第二阶段回归结果,可以发现,公共数据开放的回归系数依旧显著为正,证明本文核心结论稳健有效。

(四)稳健性检验

1.PSM-DID 方法

尽管公共数据开放对乡村振兴的影响相对外生,但考虑到公共数据开放的城市可能与未开放的城市存在特征差异,从而造成反向因果,进而引发内生性偏误问题。同时,现实中可能存在诸多因素会影响基准回归结果而本文未能将其控制,从而造成遗漏变量问题^[22]。为避免上述问题影响结果的稳健性,以下基于倾向得分匹配法,将外商投资水平、信息基础设施建设、金融业发展水平、人力资本水平、财政支持强度和产业结构作为匹配变量,以乡村振兴为结果变量,使用核密度匹配法对处理组和控制组进行匹配,得到 2009—2022 年的倾向匹配得分数据。经过平衡性检验发现,匹配结果较为有效,然后进一步利用匹配的数据进行双重差分回归。回归结果如表 4 列(1)所示,可以发现核心变量的回归系数和显著性水平与基准回归几乎保持一致,表明本文结果较为稳健。

2.交叠的异质性处理效应

在多期双重差分模型中,政策变量存在多个时间点,这使较早实施的处理组可能被当作较晚实施处理组的控制组,造成个体和时间维度上存在异质性处理效应,进而产生“负权重”问题。在“负权重”问题下,采用双向固定效应进行多时点双重差分模型估计会产生偏误^[23]。为了避免这一问题对估计结果的有效性造成干扰,本部分拟采用两种方法解决:一是基于插补估计量的视角,Gardner 提出通过两阶段估计方法来规避将政策实施较早的处理组样本纳入控制组进行稳健性检验^[24];二是基于堆

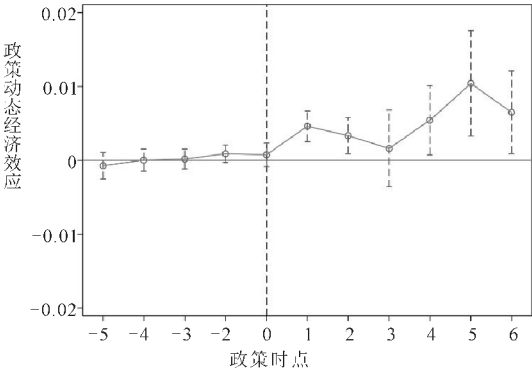


图 1 平行趋势检验

表 3 内生性检验		
变量	(1)	(2)
	一阶段	二阶段
	Pubdata	Ruralrev
IV	0.033 *** (0.08)	
Pubdata		0.104 *** (0.031)
控制变量	控制	控制
地区固定效应	控制	控制
年份固定效应	控制	控制
观测值	2898	2898
F 统计量	16.840	—
C-D Wald F	32.803	—

叠回归估计量的视角,Cengizet 等提出使用堆叠回归估计量进行回归分析^[25],每个堆叠包括来自同一时间段内实验组和从未接受处理的控制组的所有观察结果。回归结果见表 4 列(2)和列(3),结果表明核心解释变量的回归系数仍旧显著为正。综合来看,可以认为异质性处理效应对本文核心结论的影响较小。

表 4 PSM-DID 检验和交叠的异质性处理效应的回归结果

变量	(1)	(2)	(3)
		Ruralrev	
Pubdata	0.007 *** (0.002)	0.013 ** (0.006)	0.016 *** (0.005)
控制变量	控制	控制	控制
地区固定效应	控制	控制	控制
年份固定效应	控制	控制	控制
观测值	3507	—	8163
R ²	0.956	—	0.967

注:列(1)为 PSM-DID 方法的回归结果;列(2)采用 did2s 命令进行回归分析,该命令得到的结果不报告除核心解释变量外的其他变量的系数;列(3)为使用 Cengizet 等(2019)的方法进行的回归,该方法通过堆叠数据进行检验,其观测值应大于基准回归,包含 8163 个观测值。

3.安慰剂检验

由于现实经济运行存在诸多因素的干扰,实证回归结果可能存在误差^{[26][27]}。本文拟通过安慰剂检验来确保得到的核心结论具有稳健性,其核心思想为随机分配公共数据开放的地区和时间。具体而言,首先,确定公共数据开放的地区数量并从样本中随机抽取与该数量一致的地区数量,将剩余地区作为对照组,同时随机赋予处理组地区公共数据开放的时间。其次,将该过程重复 500 次,将得到的“新政策变量”与本文的被解释变量进行回归分析。图 2 为安慰剂检验结果,将“新政策变量”带入双重差分模型得到的估计系数大部分位于 0 附近或不显著,且不与本文基准回归的估计系数重叠,验证了公共数据开放的有效性,表明本文核心结论较为稳健。

除了上述稳健性检验外,本文还考虑了合成指标方法差异、省级公共数据开放和城市行政等级差异以及其他干扰政策可能带来的影响,将这些因素分别加入回归模型中进行实证分析,所得结果均与基准回归保持一致^⑦,这进一步验证了本文核心结论的有效性。

五、机制检验与异质性分析

(一)机制检验

根据上文理论分析,公共数据开放通过增强民生性财政支出强度和促进市场一体化两个途径推动乡村振兴,下面分别验证这两个途经的存在性。

1.增强民生性财政支出强度

根据上文理论分析,公共数据开放能够使公众和政府以较低成本获得地方民生性财政支出规模和结构等详细信息,继而通过监督与模仿行为推动民生性财政支出规模提高和结构优化,这表明公共数据开放释放的数据要素红利,能推动政府资源配置效率提高。据此,本部分重点考察公共数据开放影响民生性财政支出强度的具体过程:首先,考察公共数据开放能否推动民生性财政支出的绝对量增加,从整体上提高财政对民生的支持力度;其次,考察公共数据开放能否推动民生性财政支出的相对量增加,即提高民生性财政支出在财政总支出中的占比;最后,从细化的角度考察公共数据开放对教育、医疗以及社会保障和就业等单项民生性财政支出的影响,这样有利于揭示公共数据开放在民生性财政分项支出优化中的表现。具体变量选取如下:民生性财政支出规模(Abspubfisex)采用教育、医疗以及社会保障和就业等财政支出之和的自然对数度量;民生性财政支出占比(Pubfisex)采用教育、医疗以及社会保障和就业等财政支出之和与公共财政支出总额的比值度量;教育财政支出占比

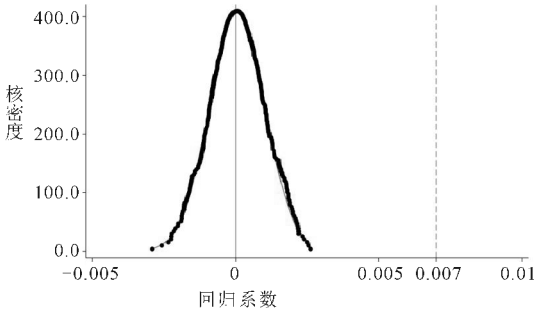


图 2 安慰剂检验

(Edufisex)采用教育财政支出与公共财政支出总额的比值度量;医疗财政支出占比(Medfisex)采用医疗财政支出与公共财政支出总额的比值度量;社会保障和就业财政支出占比(Insfisex)采用社会保障和就业财政支出与公共财政支出总额的比值度量。回归结果如表 5 列(1)~(5)所示,在将民生性财政支出规模、民生性财政支出占比、教育财政支出占比、医疗财政支出占比以及社会保障和就业财政支出占比作为被解释变量的回归结果中,公共数据开放的回归系数均至少在 10%的水平上显著为正,表明公共数据开放有利于督促政府增强民生性财政支出的强度,假设 2 得到验证。

2.促进市场一体化

根据上文理论分析,公共数据开放有利于缓解信息不对称问题,通过公布的市场价格信息和企业信誉信息,降低农业经营主体与城市企业的合作风险,促进城市与乡村形成产业链条,进而推动市场一体化。对此,本部分从价格的视角计算市场整合指数作为市场一体化(Marinteg)的衡量指标,借鉴吕冰洋和贺颖的研究采用价格法测算市场分割指标^[28],因市场整合程度与市场分割指数呈反向关系,再按照毛其淋和盛斌的方法,进一步得到市场整合指数^[29]。回归结果如表 5 列(6)所示,公共数据开放的回归系数显著为正,表明公共数据开放能有效缓解信息不对称问题,进而推动市场一体化发展,假设 3 得到验证。

表 5 机制检验结果						
变量	(1)	(2)	(3)	(4)	(5)	(6)
	Abspubfisex	Pubfisex	Edufisex	Medfisex	Insfisex	Marinteg
Pubdata	0.071 *** (0.014)	0.007 ** (0.003)	0.004 ** (0.002)	0.003 *** (0.001)	0.003 * (0.002)	0.294 ** (0.149)
控制变量	控制	控制	控制	控制	控制	控制
地区固定效应	控制	控制	控制	控制	控制	控制
年份固定效应	控制	控制	控制	控制	控制	控制
观测值	3330	3330	3613	3336	3465	3327
R ²	0.986	0.788	0.550	0.581	0.581	0.451

(二)异质性分析

1.产业形态

地方产业形态差异可能会对地方政府数据平台的服务方向产生重要影响,继而使公共数据开放对乡村振兴的促进效果不同。具体而言,若地区第一产业占比相对较高,地方政府对第一产业的重视程度通常也会相对较高,其搭建的数据平台也将更倾向于为第一产业提供针对性的数据信息服务,并持续充实相关数据集从而为第一产业生产活动提供支持。同时,第一产业占比较高的地区还意味着具备较为成熟的农业发展理念,更易于实现数据要素与生产环节的有机融合,提升生产效率和资源配置水平,进而推动第一产业高质量发展,为乡村振兴提供强劲的动力。对此,为进一步探讨地区主要产业形态对公共数据开放促进乡村振兴效果的影响,本文按照地区第一产业占比的均值和中位数对处理组样本进行分组,接着使用相同的控制组对分组样本进行回归检验。其中,第一产业占比的衡量方式为第一产业产值与 GDP 的比值。表 6 表明无论是以中位数还是平均值为划分标准的分样本回归,公共数据开放的回归系数均为正且至少在 5%的水平上显著,但第一产业占比较高地区的公共数据开放对乡村振兴具有更大的促进效果。

2.数字基础设施建设

数据要素的作用效果与地区数字基础设施建设密切相关,因而公共数据开放对乡村振兴的促进效果在很大程度上受到数字基础设施建设水平的影响。一方面,数字基础设施建设有助于将乡村各类产业信息全面纳入政府数据平台,推动数据要素与生产环节的有机融合,进而促进农村产业转型,为乡村振兴提供有力支持。另一方面,乡村地区数字基础设施可以提高数据要素与生产环节的融合效率,加快农村产业的转型升级进程,以更高效的方式发挥数据要素在推动乡村振兴中的重要作用,进而放大公共数据开放的政策红利。为进一步研究数字基础设施建设是否会影响公共数据开放对乡村振兴的促进效果,本文按照地区数字基础设施建设水平的均值和中位数对处理组样本进行分组,接

着采用相同的控制组对分组样本进行回归分析。其中,数字基础设施建设的衡量方式借鉴潘为华等的做法^[30],综合考虑数据可得性和相关性,本文选取本地电话年末用户数(万户)、移动电话年末用户数(万户)、互联网用户数以及开通互联网宽带业务的行政村比重四个指标,采用主成分分析法计算得到数字基础设施建设指标。表 7 显示无论是以中位数还是平均值为划分标准的分组样本回归,公共数据开放的回归系数均为正且大部分显著,但在数字基础设施建设水平较高地区公共数据开放对乡村振兴具有更大的促进效果。

表 6

异质性分析 1:产业形态

划分标准	(1)	(2)	(3)	(4)
	中位数		平均值	
样本组别	第一产业占比较高	第一产业占比较低	第一产业占比较高	第一产业占比较低
Pubdata	0.011 *** (0.003)	0.007 ** (0.003)	0.011 *** (0.003)	0.007 ** (0.003)
控制变量	控制	控制	控制	控制
地区固定效应	控制	控制	控制	控制
年份固定效应	控制	控制	控制	控制
观测值	2235	2288	2274	2249
R ²	0.9561	0.9624	0.9558	0.9627

表 7

异质性分析 2:数字基础设施建设

划分标准	(1)	(2)	(3)	(4)
	中位数		平均值	
样本组别	数字基础设施建设水平较高	数字基础设施建设水平较低	数字基础设施建设水平较高	数字基础设施建设水平较低
Pubdata	0.013 *** (0.003)	0.005 (0.003)	0.016 *** (0.004)	0.007 ** (0.003)
控制变量	控制	控制	控制	控制
地区固定效应	控制	控制	控制	控制
年份固定效应	控制	控制	控制	控制
观测值	2287	2236	1560	2963
R ²	0.961	0.957	0.967	0.953

六、结论与政策建议

本文基于 2010—2022 年我国城市面板数据,采用熵值法从五个方面综合测度了乡村振兴指数,并基于政府数据平台上线这一准自然实验,采用多期双重差分法实证检验了公共数据开放对乡村振兴的影响及作用机制。结果表明:(1)公共数据开放有效推动了乡村振兴,在经过平行趋势检验、工具变量法、PSM-DID 方法、安慰剂检验等一系列稳健性检验后,这一结果仍然成立;(2)公共数据开放主要通过增强民生性财政支出强度和推动市场一体化两个途径促进乡村振兴;(3)公共数据开放对乡村振兴的促进作用在第一产业占比和数字基础设施建设水平较高地区具有更大的效果。根据以上研究结论,本文提出如下政策建议。

第一,持续推动政府数据平台的建设和优化,充分发挥数据要素对乡村振兴的促进作用。一方面,对于还未上线政府数据平台的地区,应将其作为重要任务,通过上线公共数据平台提高政府透明度、优化资源配置。另一方面,对于已上线政府数据平台的地区,需要进一步加强对农村产业的信息扶持力度,尤其是在自然资源、环保、气象等方面提供实时数据信息,便于农村生产安排。此外,应推动政府数据平台向智慧化方向发展,既要降低农村用户获取数据的门槛,又要简化操作流程,为农民提供便捷高效的使用体验。

第二,增加民生性财政支出和推动市场一体化,探索多维路径助力乡村振兴。各地政府要以提高政府公共数据资源开放水平为目的,注重公共数据资源的共享与配置,从而充分发挥数据要素推动乡村振兴的积极效果。一方面,要大力优化财政支出结构,更加注重民生性支出,切实提高财政资金使

用效果,完善财政投入保障机制,将财政资金进一步向农业农村倾斜,改善农业生产环境和居住环境,为乡村振兴注入活力。另一方面,要注重市场一体化建设,扩大公共数据的供给,提高公共数据资源的使用效率,深化数据要素在农村产业中应用的潜力,推动数字红利的充分释放。

第三,因地制宜地打造乡村特色产业和配套数字基础设施建设,为乡村振兴创造更为有利的条件。一方面,借助政府数据平台的共享数据,精准定位各地产业发展规划,注重特色农产品、乡村旅游等品牌塑造,打破传统产业边界和业态模式,积极培育乡村新产业新业态,打造乡村特色产业,从而推动乡村产业提质增效。另一方面,要提升数字基础设施建设水平和质量,提升数据传输速度和容量,确保数据流通更高效、更准确,避免数据开放因基础设施不足而流通受阻,从而深化数据要素与农村产业的有机融合,助力乡村振兴。

注释:

- ①数据来源:https://www.guancha.cn/economy/2016_05_15_360316.shtml。
- ②参考复旦大学数字与移动治理实验室发布的《2021 中国地方政府数据开放报告(城市)》。以此为借鉴,笔者进一步手工整理计算。
- ③宿迁市公共数据开放平台:<https://data.suqian.gov.cn/oportal/catalog/index?domainId=domain-19&page=1>。
- ④徐州市政府数据平台:<http://data.gxj.xz.gov.cn/#/DataDirectory>;丰县政府快报:<https://www.ourjiangsu.com/a/20220102/1641110412671.shtml>。
- ⑤成都市公共数据开放平台:<https://data.chengdu.gov.cn/oportal/catalog/>;茂名市公共数据开放平台:<https://gddata.gd.gov.cn/opdata/base/collect?chooseValue=collectForm&deptCode=31&t=1737619643498t>;深圳市公共数据开放平台:<https://opendata.sz.gov.cn/data/dataSet/toDataSet>。
- ⑥武汉市公共数据开放平台:http://data.wuhan.gov.cn/page/search/data_search.html;无锡市公共数据开放平台:<https://data.wuxi.gov.cn/data/search/index.htm>。上述数据平台为本文随机抽取并列举,其余数据平台公布数据情况基本一致。
- ⑦篇幅所限,相关指标体系和指数不再详细披露,留存备案。

参考文献:

[1] 张文珂,张琳雪,万立全,等.数字经济促进乡村共同富裕的现实路径[J].南开经济研究,2024(5):14—30.

[2] 方锦程,刘颖,高昊宇,等.公共数据开放能否促进区域协调发展?——来自政府数据平台上线的准自然实验[J].管理世界,2023(9):124—142.

[3] 蔡运坤,周京奎,袁旺平.数据要素共享与城市创业活力——来自公共数据开放的经验证据[J].数量经济技术经济研究,2024(8):5—25.

[4] 沈艳,冯冬发,陶云清.公共数据开放与中国县域数字创业[J].城市问题,2024(5):38—49.

[5] 王晓丹,石玉堂,刘达.公共数据开放能促进数字经济与实体经济融合吗?——来自政府数据平台上线的准自然实验[J].南方经济,2024(9):25—44.

[6] 鲁钊阳,杜雨潼,邓琳钰.数字经济促进乡村产业振兴的影响机理及实证研究[J].江南大学学报(人文社会科学版),2024(1):37—51.

[7] Berg, M. Household Income Strategies and Natural Disasters: Dynamic Livelihoods in Rural Nicaragua[J]. Ecological Economics,2010,69(3):592—602.

[8] 陈蔚.马克思主义政治意识形态制度保障力面临的挑战剖析[J].当代世界与社会主义,2011(3):117—120.

[9] Edmiston, K. D. State and Local E-Government among Municipalities: Prospects and Challenge[J]. American Review of Public Administration, 2002,33(1):20—45.

[10] 唐曼,王刚.行为模仿:地方政府行为的一个分析框架——基于多案例的研究[J].公共行政评论,2021(3):158—175.

[11] 卢洪友,龚锋.政府竞争、“攀比效应”与预算支出受益外溢[J].管理世界,2007(8):12—22.

[12] 赖德胜,陈建伟.人力资本与乡村振兴[J].中国高校社会科学,2018(6):21—28.

[13] 彭超.高素质农民培育政策的演变、效果与完善思路[J].理论探索,2021(1):22—30.

[14] 白描.乡村振兴背景下健康乡村建设的现状、问题及对策[J].农村经济,2020(7):119—126.

[15] 聂建亮,吴玉锋.社会保障助力乡村振兴:基础、路径与提升策略[J].农村经济,2021(12):10—17.

[16] 邹宝玲,曹壹帆,肖亚成.共同富裕视域下市场一体化的农民增收效应研究[J].财经问题研究,2023(11):115—129.

[17] 黄贇琳,秦淑悦.市场一体化对消费升级的影响——基于“量”与“质”的双重考察[J].中国人口科学,2021(5):18—31.

[18] 徐雪,王永瑜.中国乡村振兴水平测度、区域差异分解及动态演进[J].数量经济技术经济研究,2022(5):64—83.

- [19] Chen, Y., Fan, Z., Gu, X., Zhou, L. A. Arrival of Young Talents; Send-Down Movement and Rural Education in China[J]. The American Economic Review, 2020, 110(11): 3393—3430.
- [20] 高跃光, 范子英, 冯晨. 义务教育专项融资与教育投入: 基于开征地方教育附加的研究[J]. 管理世界, 2023(2): 72—82.
- [21] 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.
- [22] 石虹, 余少龙. 数据要素集聚如何驱动企业新质生产力发展? [J]. 技术经济, 2024(12): 35—46.
- [23] Goodman-Bacon, A. Difference-in-differences with Variation in Treatment Timing[J]. Journal of Econometrics, 2021, 25(2): 254—277.
- [24] Gardner, J. Two-stage Differences in Differences[Z]. Working Paper, 2021.
- [25] Cengizet, D., Dube, A., Lindner, A., et al. The Effect of Minimum Wages on Low-wage Jobs[J]. The Quarterly Journal of Economics, 2019, 134(3): 1405—1454.
- [26] 辛冲冲, 席鹏辉, 唐世芳. 区域协调发展促进了就业扩容吗? ——基于京津冀协同发展战略实施的准自然实验[J]. 经济与管理研究, 2024(12): 88—105.
- [27] 陈艳利, 刘亚. 数据要素市场化配置与全要素生产率——来自数据交易平台设立的证据[J]. 中南财经政法大学学报, 2024(6): 131—143.
- [28] 吕冰洋, 贺颖. 迈向统一市场: 基于城市数据对中国商品市场分割的测算与分析[J]. 经济理论与经济管理, 2020(4): 13—25.
- [29] 毛其淋, 盛斌. 对外经济开放、区域市场整合与全要素生产率[J]. 经济学(季刊), 2012(1): 181—210.
- [30] 潘为华, 贺正楚, 潘红玉. 中国数字经济发展的时空演化和分布动态[J]. 中国软科学, 2021(10): 137—147.

Public Data Opening and Rural Revitalization: Quasi-natural Experiment on the Government Data Platform

XIN Chongchong¹ YUAN Yijie² HAN Yunge³

(1. Editorial Department of Beijing Social Sciences, Beijing Academy of Social Sciences, Beijing 100101, China;

2. School of Applied Economics, China Academy of Social Sciences, Beijing 102488, China;

3. College of Economics & Management of Hubei University of Arts and Science, Xiangyang 441053, China)

Abstract: As an important public resource, government data elements have the advantage of efficiently integrating information and can effectively solve many problems in rural revitalization. Based on the panel data of China's prefecture-level cities from 2010 to 2022, this paper empirically tests the influence of public data opening on rural revitalization and its realization path by using the difference in difference method. The research results show that the opening of public data releases digital dividends and can effectively promote rural revitalization, and a number of robustness tests support this conclusion. The mechanism analysis shows that the openness of public data can promote rural revitalization by enhancing the intensity of people's livelihood financial expenditure and promoting market integration. Heterogeneity analysis shows that the opening of public data plays a greater role in promoting rural revitalization in areas with relatively high primary industry and relatively perfect digital infrastructure. Based on this, this paper puts forward some development ideas, such as continuously promoting the construction and optimization of government data platform, continuously increasing people's livelihood fiscal expenditure and promoting market integration, and building rural characteristic industries and supporting digital infrastructure construction according to local conditions, so as to provide policy reference for promoting rural revitalization.

Key words: Open Public Data; Rural Revitalization; People's Livelihood Fiscal Expenditure; Market Integration; Data Elements

(责任编辑: 易会文)