

DOI:10.3969/j.issn.1674-8131.2024.05.001

国家大数据综合试验区建设的 家庭增收效应研究

王乾坤

(中国政法大学商学院,北京100088)

摘要:国家大数据综合试验区建设促进了新产业、新业态、新模式的产生和发展,有利于家庭通过参与更多更高质量的生产活动来增加收入。采用中国家庭金融调查2013年、2015年、2017年、2019年4期面板数据的分析表明:国家大数据综合试验区建设显著提高了试验区内的家庭收入,该作用主要通过增加工资性收入和经营性收入来实现,且对收入水平较低家庭的增收效应更强;国家大数据综合试验区建设可以通过促进数字普惠金融发展和创业发展两条路径来增加家庭收入;国家大数据综合试验区建设对家庭总收入和工资性收入的提升作用在城市家庭、社会资本和人力资本水平较高家庭中更显著,而对经营性收入的提升作用在农村家庭、社会资本和人力资本水平较低家庭中更显著;城市家庭、社会资本水平和人力资本水平较高家庭的增收效应主要通过增加工资性收入来实现,而农村家庭、社会资本和人力资本水平较低家庭的增收效应主要通过增加经营性收入来实现。因此,应当积极推进国家大数据综合试验区建设,充分发挥大数据发展的家庭增收效应和家庭收入差距缩小作用。

关键词:大数据;家庭收入;数字普惠金融;创业发展;工资性收入;经营性收入

中图分类号:F124.7;F207 **文献标志码:**A **文章编号:**1674-8131(2024)05-0001-17

引用格式:王乾坤. 国家大数据综合试验区建设的家庭增收效应研究[J]. 西部论坛,2024,34(5):1-17.
WANG Qian-kun. Household income increasing effect of the national big data comprehensive pilot zone[J].
West Forum, 2024, 34(5): 1-17.

一、引言

在全球数字化浪潮中,数据已然成为当今世界的基础性战略资源,尤其是大数据的作用日益凸显。

* 收稿日期:2024-04-22;修回日期:2024-07-30

基金项目:国家社会科学基金重大项目(21ZD079);教育部哲学社会科学重大专项项目(2023JZDZ022);教育部人文社会科学研究规划基金项目(21YJA790004)。

作者简介:王乾坤(1993),男,河北邯郸人;博士研究生,主要从事数字经济、收入分配研究;Tel:13315019285, Email:vigor009@163.com。

为促进大数据发展,2015年,国务院印发《促进大数据发展行动纲要》,党的十八届五中全会明确提出要实施“国家大数据战略”。2016年2月,贵州省率先启动国家大数据综合试验区建设;2016年10月,国家分别在京津冀、珠三角设立跨区域类国家大数据综合试验区,在上海市、重庆市、河南省、辽宁省设立区域示范类国家大数据综合试验区,在内蒙古自治区设立基础设施统筹发展类国家大数据综合试验区。自此,八大国家大数据综合试验区共同引领和驱动我国的大数据发展。

随着国家大数据综合试验区的建设和发展,其经济社会效益成为学界研究的热点问题。现有文献主要从城市(地区)和企业两个层面考察国家大数据综合试验区的政策效应,比如对城市高质量发展(苏锦旗等,2023;杨玉琪等,2023)^[1-2]、绿色发展(郭炳南等,2022;张莹莹等,2023;刘占芳等,2024)^[3-5]、创新发展(徐林等,2022;刘传明等,2023;胡婧玮等,2024)^[6-8]、产业转型(魏丽莉等,2022;钟昌标等,2023;陈启斐等,2024)^[9-11]、全要素生产率(邱子迅等,2021;刘军等,2024)^[12-13]以及共同富裕(彭国柱等,2024)^[14]和新质生产力(赵鹏等,2024)^[15]等的促进作用,对企业技术创新(曹平等,2021;戴艳娟等,2023)^[16-17]、数字化转型(孙伟增等,2023;石玉堂等,2024)^[18-19]以及全要素生产率(张益豪等,2023;孙鹏等,2023)^[20-21]等的积极影响。除了企业以外,家庭也是市场经济中重要的微观经济主体,国家大数据综合试验区建设会对家庭发展产生怎样的影响,也需要深入探究,然而,目前缺乏这方面的实证研究。

同时,中国式现代化是全体人民共同富裕的现代化,在扎实推动共同富裕的过程中,要通过完善收入分配制度来多渠道增加居民收入。与此相适应,大数据发展能否促进以及怎样促进居民收入增长就成为具有重要理论价值和现实意义的研究课题。然而,目前关于大数据发展与居民收入及收入分配之间关系的研究还不多,尤其缺乏相关经验证据。而且,少量的实证研究只是从企业层面来考察大数据发展对企业劳动收入份额的影响,而没有从家庭层面分析大数据发展对居民收入的影响。比如:卜寒等(2023)^[22]、戴艳娟和沈伟鹏(2024)^[23]、李麦收和李华(2024)^[24]研究发现,国家大数据综合试验区建设能够显著提升企业的劳动收入份额;但宋华盛和卢历祺(2024)^[25]的分析却表明,大数据发展显著降低了企业的劳动收入份额。

鉴于上述,本文在已有研究的基础上,探讨国家大数据综合试验区建设对家庭收入的影响及其机制,并将国家大数据综合试验区建设作为一项准自然实验,采用中国家庭金融调查(China Household Finance Survey, CHFS)2013年、2015年、2017年、2019年4期的面板数据进行实证检验。相比已有文献,本文的边际贡献主要在于:第一,在研究视角上,从家庭收入视角拓展和深化了设立国家大数据综合试验区的政策效应研究,并为通过大数据发展提高居民收入提供了经验证据,有助于深入认识数字经济与共同富裕之间的内在关系;第二,在研究方法上,将家庭层面数据与城市层面数据相匹配,为深入分析各类试点政策在家庭层面的各种效应提供了思路参考和方法借鉴;第三,在研究内容上,探究了国家大数据综合试验区建设通过促进数字普惠金融发展和创新发展来增加家庭收入的机制,并从城乡差异以及家庭社会资本和人力资本差异等方面考察了国家大数据综合试验区建设影响家庭收入的异质性,为充分发挥国家大数据综合试验区建设的积极作用,以大数据发展赋能共同富裕提供了路径启示。

二、理论分析与研究假说

1. 国家大数据综合试验区建设对家庭收入的影响

居民收入通常与地区经济发展水平具有正相关性,经济增长往往伴随着家庭收入增长。尽管从短

期来看,作为一种新型生产要素,大数据与其他生产要素的融合成本及转型成本可能抑制经济产出,但从长期来看,大数据可以通过乘数效应提升中间品质量和促进技术进步,持续推动经济增长(杨俊等,2022)^[26]。因此,国家大数据综合试验区的设立和建设能够通过大数据的领先发展使得试验区城市实现更快的经济增长(刘德林等,2021)^[27]。大数据发展对地区经济增长的促进作用是多方面的,在宏观层面上,可以通过数字技术创新促进整体技术进步(莫赞等,2021)^[28],通过信息环境改善优化资源配置,从而提高地区全要素生产率(邱子迅等,2021;刘军等,2024)^[12-13];在微观层面上,可以缓解企业融资约束(孙洁等,2022)^[29],促进企业技术创新(任英华等,2023)^[30],提升企业生产效率(张益豪等,2023)^[20];还可以提高政务透明度和社会信用水平(张红春等,2020)^[31],促进国家治理现代化(赵术高等,2021)^[32],有利于更好地发挥政府作用。

大数据发展带来的经济增长为家庭收入增长提供了基础和条件,而要在经济增长中获得收入增长红利,最主要的渠道就是参与生产活动。一方面,大数据发展能够显著增加劳动力就业和创业机会。大数据投入到生产过程中将产生就业创造效应,即通过技术创新和业务流程再造衍生出对具备不同技能层次劳动力的新增需求(Gardiner et al.,2018)^[33],使得更多的劳动者能够在更具比较优势的岗位上发挥作用,并获取相应报酬,从而带来家庭收入水平的整体提升(Acemoglu et al.,2020)^[34]。大数据技术的应用推动了平台经济、共享经济等新经济形态的孕育和发展(侯冠宇等,2023)^[35],大量新产业、新业态、新模式的涌现为更多家庭创造了更好的创业机会,有助于家庭收入增长。另一方面,大数据发展有利于提高劳动者就业和创业的收入水平。大数据发展缓解了教育资源的分配不均,降低了知识技能学习的门槛和成本,劳动者可以借助大数据平台学习掌握专业技能,提升自身人力资本水平和就业竞争力(陶长琪等,2022)^[36],从而获得更多的劳动报酬。大数据发展有利于集聚关键生产要素和营造良好的创新创业生态环境,能够通过拓展销售渠道、降低中间成本等提高生产经营利润(尹志超等,2024)^[37],从而提高家庭创业收入。此外,大数据发展还能改善收入分配,提高劳动收入份额(卜寒等,2023;戴艳娟等,2024;李麦收等,2024)^[22-24],而劳动收入通常是家庭收入的主要来源。因此,国家大数据综合试验区建设能够显著提高试验区内劳动者的生产性收入(主要是在就业和创业等社会生产活动中获得的工资性收入和经营性收入),从而促进家庭总收入增长。

基于上述分析,本文提出假说 H1:国家大数据综合试验区建设能够显著促进家庭收入增长,且该效应主要通过增加家庭的工资性收入和经营性收入来实现。

2. 国家大数据综合试验区建设、数字普惠金融发展与家庭收入增长

一方面,国家大数据综合试验区建设有利于试验区的数字普惠金融发展。大数据与信息技术的整合运用(包括大规模数据的收集、处理与分析等)显著缓解了信息不对称问题(祁怀锦等,2024)^[38],金融机构能够凭借技术手段对各类用户风险特征进行精细化识别与评估,从而使信贷资源得以更准确地配置至各个经济主体,不仅拓展了金融服务边界,也提高了金融服务的可达性和运行效率(刘典,2024)^[39]。同时,大数据技术的实时监控和风险预警功能替代或优化了人工审核流程,使金融机构能够在精准把控风险的前提下保持金融服务的连续性和稳定性,减少了风险管理不善导致的金融服务中断或收缩(张颖熙等,2022)^[40]。此外,金融机构还可以利用大数据技术创新和推广普惠金融产品,针对小微企业及消费者的个性化需求推出定制化金融服务,并降低金融服务准入门槛,使得以往难以触及传统金融服务的低收入群体和偏远地区居民能够方便快捷地享受低成本、高效率的金融服务,从而实质性地推动普惠金融理念在更广范围内落地与深化,促进普惠金融目标的实现(刘元维等,2023)^[41]。

另一方面,数字普惠金融发展能够促进家庭收入增长。金融的主要服务对象是企业和家庭,然而现实场景中,中小企业和低收入家庭常常面临金融排斥问题(任太增等,2022)^[42]。而大数据技术赋予数字普惠金融强大的功能,不仅能够有效地弥补传统金融体系的缺陷,还能为那些在传统金融体系中处于不利地位的弱势群体开辟更多获取金融服务的途径。随着大数据技术与数字普惠金融的深度融合,地域、阶层间的金融服务差距将被削弱,数字普惠金融服务不需要用户具备较高的数字素养、金融素养等,个人和家庭可以通过获取更优质的金融服务增强抵抗经济风险的能力,进而增加其收入来源(刘艳华等,2023;王小华等,2023)^[43-44]。同时,大数据赋能下的数字普惠金融发展能够促进产业结构升级和人力资本提升。数字普惠金融有助于破解生产经营过程中的资金瓶颈难题,扩大生产规模,并推动产业转型升级,从而创造更多更高质量的就业机会,促进家庭收入增长。而且,数字普惠金融为劳动者获取教育信贷、接受公平教育、提升人力资本提供了更多便利(孙俊娜等,2023)^[45],有助于提升人力资源的质量与效能,并驱动经济增长模式向人才导向型转变,从而在整体上带动家庭收入水平和社会财富的增长(陶长琪等,2022)^[36]。

基于上述分析,本文提出假说 H2:国家大数据综合试验区建设能够通过促进数字普惠金融发展的路径来增加家庭收入。

3. 国家大数据综合试验区建设、创业发展与家庭收入增长

一方面,国家大数据综合试验区建设有利于试验区的创业发展。大数据发展能够增加地区创业机会,并为创业活动聚集各类资源要素。国家大数据综合试验区不仅通过多种优惠政策激励企业增加研发投入以实现创新,还通过引入优质金融资源、高新技术产业、创新创业人才等为创业活动集聚金融、技术、劳动等关键要素(孙伟增等,2023)^[18]。大数据技术的创新和应用,打破了资源配置的时空约束,提高了知识共享与传播效率,为创业机会的发现和创业活动的兴起创造了有利条件(石玉堂等,2023)^[46],并通过积累创新知识激发创业思维,促进创业机会的发现和创业活动的产生。同时,大数据技术与传统产业深度融合,在产业数字化和数字产业化转型发展中创造出新的创业机会。

另一方面,创业发展能够促进家庭收入增长。熊彼特认为,创业活动可以产生新的产品、市场和技术,为经济注入活力并促进创新型经济增长。而且,创业活动有助于低收入家庭向更高收入群体流动。随着大数据技术与传统经济活动的深度融合,新产业、新业态、新模式不断涌现,更多家庭可以通过开展创业活动获得更多经营性收入。大数据技术赋能下的创业活动,能够将商品与消费者直接对接,减少中间环节,降低交易成本,提高商品利润率,进而增加创业者收入(陈明生等,2024)^[47]。创业活动还能为劳动力市场提供更多职位,吸收剩余劳动力就业,为家庭带来稳定、持续的收入增长。此外,大数据技术的应用促使传统工业时代的集中型劳动组织模式逐渐转变为当今信息时代的分散型劳动组织模式,这种模式下的用工要求对于人力资本较低群体极具吸引力(戚聿东等,2022)^[48],从而能够有效增加此类群体的就业概率与收入水平。

基于上述分析,本文提出假说 H3:国家大数据综合试验区建设能够通过促进创业发展的路径来增加家庭收入。

三、实证研究设计

1. 基准模型设定

为检验国家大数据综合试验区建设对家庭收入的影响,本文构建如下基准模型:

$$Inc_{it} = \alpha + \beta Data_{it} + \gamma X_{it} + \delta_i + \tau_t + \varepsilon_{it}$$

其中,下标 i 、 t 分别代表家庭和年份。被解释变量 (Inc_{it})“家庭收入”为 t 年 i 家庭总收入(包括工资性收入、经营性收入、财产性收入和转移性收入)的自然对数值。核心解释变量 ($Data_{it}$)“大数据试验区”为 t 年 i 家庭所在省份是否是国家大数据综合试验区的虚拟变量,若是赋值为 1,否则赋值为 0。 X_{it} 代表控制变量,借鉴肖伟等(2023)^[49]、尹志超等(2021)^[50]的研究,从户主、家庭和地区 3 个层面选取以下控制变量:户主层面包括“年龄”(观测年与户主出生年之差)、“年龄平方”(除以 100)、“婚姻状况”(已婚赋值为 1,否则赋值为 0)、“受教育程度”(根据学历赋值 0~19)^①、“住房情况”(拥有自有住房赋值为 1,否则为赋值为 0)、“户籍”(农村户口赋值为 1,城镇户口赋值为 0),家庭层面包括“家庭资产”(取自然对数)^②、“家庭负债”(取自然对数)^③、“家庭劳动力”(与受访户共享收入、共担支出的并在上一年有工作的人员总数),地区层面包括“经济发展水平”(城市 GDP 的自然对数值)、“人口密度”(城市年末总人口除以城市面积的自然对数值)、“财政收入”(城市地方财政一般预算内收入的自然对数值)。 δ_i 和 τ_t 分别代表家庭和年份固定效应, ε_{it} 为随机扰动项。模型回归采用家庭层面聚类稳健标准误下的 OLS 估计方法。

2. 中介效应模型设定

为检验国家大数据综合试验区建设能否通过促进数字普惠金融发展和创业发展的路径来增加家庭收入,本文在基准模型的基础上设定如下中介效应模型:

$$M_{it} = \beta_0 + \beta_1 Data_{it} + \beta_2 X_{it} + \delta_i + \tau_t + \varepsilon_{it}$$

$$Inc_{it} = \gamma_0 + \gamma_1 Data_{it} + \gamma_2 M_{it} + \gamma_3 X_{it} + \delta_i + \tau_t + \varepsilon_{it}$$

其中, M 为中介变量,其余变量与基准模型一致。根据前文理论分析,选取以下中介变量:一是“数字普惠金融”。采用北京大学数字普惠金融指数的自然对数值来衡量城市数字普惠金融发展水平,用以检验数字普惠金融的中介作用。二是“家庭创业”和“城市创新”。其中,“家庭创业”为家庭是否创业的虚拟变量,即从事工商业经营的家庭赋值为 1,否则赋值为 0(宋冬林等,2022;温永林等,2023)^[51-52]，“城市创业”为城市样本家庭中创业家庭的占比,分别用以反映家庭和城市层面的创业概率,进而检验创业的中介作用。

3. 样本选取与数据处理

本文以家庭为研究对象,家庭层面的数据来自中国家庭金融调查(CHFS)数据库。该数据库收集了家庭收入信息(包括工资性收入、经营性收入、财产性收入和转移性收入)以及其他个人和家庭层面的信息,为本文的研究提供了数据支持。该调查自 2011 年开始,每两年进行一次,目前已经完成 5 轮数据采集。借鉴尹志超和郭沛瑶(2022)^[53]的研究,并结合国家大数据综合试验区设立的时间,本文保留政策发生前两期与政策发生后两期的样本,即采用 2013 年、2015 年、2017 年、2019 年 4 期的面板数据来进行实证检验。仅保留户主年龄大于等于 18 岁且小于等于 75 岁的样本,根据价格指数将所有货币性指标平

① 没上过学、小学、初中、高中(或者中专、职高)、大专/高职、大学本科、硕士研究生及以上分别赋值为 0、6、9、12、15、16、19。

② 家庭资产包括非金融资产和金融资产,其中非金融资产包括农业经营资产、工商业经营资产、土地资产、房产、车辆资产等,金融资产包括社保账户余额、现金、存款、股票、基金、债券、衍生品、理财、外币资产、黄金等。

③ 家庭负债包括农业负债、工商业负债、房产负债、车辆负债、股票负债、教育负债、医疗负债等。

减为以 2013 年为基期的不变价格,并对所有连续型变量进行上下 1%的缩尾处理,最终形成 4 期面板数据。表 1 报告了主要变量的描述性统计结果。

表 1 主要变量描述性统计结果

变 量	观测值	平均值	中位数	标准差	最小值	最大值	
被解释变量	家庭收入	29 240	10. 21	10. 49	1. 49	12. 84	
	经营性收入	29 240	3. 96	0. 69	4. 42	15. 88	
	转移性收入	29 240	3. 29	0. 00	3. 99	14. 44	
	财产性收入	29 240	1. 34	0. 00	2. 81	14. 95	
	工资性收入	29 240	3. 83	0. 00	4. 94	13. 30	
解释变量	大数据试验区	29 240	0. 15	0. 00	0. 36	1	
控制变量	年龄	29 240	50. 70	51. 00	9. 17	65. 00	
	年龄平方	29 240	26. 55	26. 01	8. 94	42. 25	
	婚姻状况	29 240	0. 92	1. 00	0. 27	0	
	受教育程度	29 240	7. 96	9. 00	3. 11	15	
	住房情况	29 240	0. 76	1. 00	0. 43	0	
	户籍	29 240	0. 70	1. 00	0. 46	0	
	家庭特征	家庭资产	29 240	12. 20	12. 33	1. 41	15. 31
		家庭负债	29 240	4. 07	0. 00	5. 14	13. 12
		家庭劳动力	29 240	1. 92	2. 00	1. 33	5. 00
	城市特征	经济发展水平	29 240	17. 09	17. 01	1. 03	19. 34
人口密度		29 240	5. 95	6. 05	0. 76	7. 48	
财政收入		29 240	14. 56	14. 31	1. 22	17. 81	
中介变量	数字普惠金融	29 240	5. 25	5. 31	0. 26	5. 77	
	家庭创业	29 240	0. 50	0. 00	0. 50	1. 00	
	城市创业	29 240	0. 48	0. 06	1. 22	0. 83	

四、实证检验结果分析

1. 基准回归

表 2 为基准模型的回归结果。第(1)列未纳入控制变量,第(2)列至第(4)列依次纳入户主层面、家庭层面和城市层面的控制变量,“大数据试验区”的回归系数虽有所变化,但均在 1%水平上显著为正,表明国家大数据综合试验区建设显著提高了试验区内的家庭收入。进一步考察国家大数据综合试验区建设对家庭不同来源收入的影响,分别以“工资性收入”“经营性收入”“转移性收入”“财产性收入”为被解释变量的检验结果见表 3。“大数据试验区”对“工资性收入”和“经营性收入”的回归系数显著为正,而对“转移性收入”和“财产性收入”的回归系数为正但不显著,表明国家大数据综合试验区建设提高了试

验区内家庭的工资性收入和经营性收入,但对家庭转移性收入和财产性收入的影响不显著,即国家大数据综合试验区建设的家庭增收效应主要是通过提高工资性收入和经营性收入来实现的。由此,本文提出的假说 H1 得到验证。

表 2 基准回归结果

变 量	家庭收入			
	(1)	(2)	(3)	(4)
大数据试验区	0.137*** (0.035)	0.14*** (0.034)	0.096*** (0.033)	0.112*** (0.034)
年龄		0.094*** (0.017)	0.037** (0.016)	0.037** (0.016)
年龄平方		-0.107*** (0.018)	-0.045** (0.017)	-0.045*** (0.017)
婚姻状态		0.324*** (0.065)	0.198*** (0.060)	0.198*** (0.061)
受教育程度		0.027*** (0.006)	0.02*** (0.006)	0.02*** (0.006)
住房情况		0.109*** (0.040)	0.028 (0.038)	0.025 (0.038)
户籍		-0.126 (0.088)	-0.11 (0.085)	-0.109 (0.086)
家庭总资产			0.203*** (0.011)	0.203*** (0.011)
家庭劳动力			0.357*** (0.012)	0.358*** (0.012)
家庭总负债			0.001 (0.002)	0.001 (0.002)
经济发展水平				0.319*** (0.085)
人口密度				-0.005 (0.477)
财政收入				-0.144** (0.067)
常数项	10.194*** (0.005)	7.743*** (0.411)	6.078*** (0.400)	2.738 (2.917)
年份固定效应	控制	控制	控制	控制
家庭固定效应	控制	控制	控制	控制
样本量	29 240	29 240	29 240	29 240
R ²	0.533	0.536	0.577	0.577

注:***、**、*分别表示 1%、5%、10%的显著性水平,括号内数值为标准误,下表同。

表 3 国家大数据综合试验区建设对不同来源家庭收入的影响

变 量	工资性收入	经营性收入	转移性收入	财产性收入
大数据试验区	0.329*** (0.104)	0.193** (0.094)	0.103 (0.065)	0.023 (0.095)
样本量	29 240	29 240	29 240	29 240
R ²	0.629	0.507	0.549	0.623

注:所有模型均控制了控制变量以及年份和家庭固定效应,限于篇幅,控制变量和常数项估计结果略,下表同。

为进一步考察在不同的家庭收入水平下,国家大数据综合试验区建设对家庭收入的影响是否存在差异,本文采用面板分位数回归模型分别进行家庭收入 25%、50%、75%分位数检验,结果见表 4。可以发现,在不同分位数下“大数据试验区”的回归系数均在 1%水平上显著为正,表明国家大数据综合试验区建设的家庭增收效应普遍存在;同时,随着家庭收入水平的提高,“大数据试验区”的回归系数趋于减

小,表明国家大数据综合试验区建设对于收入水平较低家庭的增收效应更强,能够在一定程度上缩小家庭收入差距,有利于促进共同富裕。

表 4 分位数回归结果

变 量	25 分位数	50 分位数	75 分位数
大数据试验区	0.15 *** (0.037)	0.121 *** (0.023)	0.088 *** (0.021)
样本量	29 240	29 240	29 240
R ²	0.135	0.177	0.16

2. 稳健性检验

(1) 平行趋势检验。采用双重差分法进行政策效应分析需要满足平行趋势假设,借鉴尹志超和郭沛瑶(2022)^[53]的研究,设定模型: $Inc_{i,t} = \alpha + \beta_1 Data_{i,t-3} + \beta_2 Data_{i,t+1} + \beta_3 Data_{i,t+3} + \gamma \cdot X_{i,t} + \delta_i + \tau_t + \varepsilon_{i,t}$ 。其中, $Data_{i,t-3}$ 表示国家大数据综合试验区设立3年前的虚拟变量(样本数据为2013年), $Data_{i,t+1}$ 表示国家大数据综合试验区设立1年后的虚拟变量(样本数据为2017年), $Data_{i,t+3}$ 表示国家大数据综合试验区设立3年后的虚拟变量(样本数据为2019年),其余变量与基准模型一致。平行趋势检验结果见表5第(1)列,在国家大数据综合试验区设立前,回归系数不显著,实验组和对照组之间不存在显著差异,满足平行趋势假设;在国家大数据综合试验区设立后,回归系数显著为正,表明政策效应显著。

(2) 增加固定效应。考虑到地区特征可能对家庭收入产生影响,在基础模型中加入年份与省份的交互固定效应,重新进行检验,估计结果见表5第(2)列,“大数据试验区”的回归系数依然显著为正。

(3) 更换被解释变量。以“家庭人均收入”(家庭总收入/家庭人口规模)为被解释变量,重新进行模型检验,估计结果见表5第(3)列,“大数据试验区”回归系数仍然显著为正。

(4) 控制相关政策影响。本文结论可能受到其他政策的干扰,如“宽带中国”示范城市建设也可能对家庭收入产生影响。为此,在基准模型加入“宽带中国”示范城市建设的政策虚拟变量,检验结果见表5第(4)列,“大数据试验区”的回归系数还是显著为正。

(5) 反事实检验。采用虚构政策变量的方法进行证伪检验:一是将政策实施时间提前1年,与实验组构建伪政策变量(实验组×2015年虚拟变量);二是虚构实验组,与政策实施时间构建伪政策变量(虚构实验组×2016年虚拟变量)。分别检验的结果见表5第(5)(6)列,伪政策变量的回归系数均不显著,表明基准模型得到的试验区家庭增收效应具有非随机性,是国家大数据综合试验区建设的结果。

(6) 安慰剂检验。随机抽取政策时间和试验区构建伪政策变量,并进行模型检验,重复500次,绘制伪政策变量估计系数的核密度图(见图1),可以发现,估计系数均值接近于0,且基准模型的估计系数为显著异常值,表明试验区家庭具有相对较高的收入水平是由除国家大数据综合试验区建设以外其他因素导致的可能性极低,再次验证了本文核心结论的稳健性。

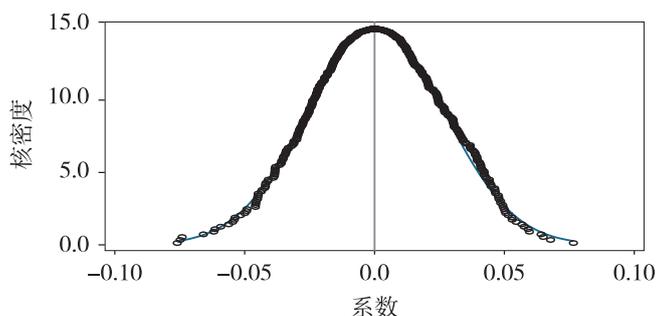


图 1 安慰剂检验结果

表 5 稳健性检验结果

变 量	家庭收入	家庭收入	家庭人均收入	家庭收入	家庭收入	家庭收入
	(1)	(2)	(3)	(4)	(5)	(6)
大数据试验区 _{<i>t-3</i>}	-0.103(0.122)					
大数据试验区 _{<i>t+1</i>}	0.204**(0.088)					
大数据试验区 _{<i>t+3</i>}	1.071***(0.120)					
大数据试验区		0.111***(0.034)	0.097***(0.037)	0.118***(0.034)		
“宽带中国”示范城市				0.054*(0.031)		
实验组×2015年虚拟变量					0.076(0.069)	
虚构实验组×2016年虚拟变量						-0.005(0.01)
年份×省份固定效应	未控制	控制	未控制	未控制	未控制	未控制
样本量	29 240	29 240	29 240	29 240	29 240	29 240
R ²	0.577	0.577	0.543	0.577	0.577	0.588

3. 内生性处理

为缓解可能存在的反向因果关系、样本自选择、遗漏变量等内生性问题对回归结果的干扰,本文使用样本匹配法和工具变量法进行内生性处理。

(1) 样本匹配法。首先,将户主控制变量、家庭控制变量以及经济发展水平作为协变量,使用 Logit 模型估计倾向得分;然后,依据 1:1 近邻匹配原则进行样本匹配,获取 7 252 个样本;最后,采用匹配后的样本进行模型检验,回归结果见表 6 第(1)列。“大数据试验区”回归系数在 5%的水平上显著为正。其次,使用熵平衡法进行样本匹配。熵平衡法不仅可以保证实验组和对照组在各协变量高阶矩上的分布相近,还可以避免倾向得分匹配主观选取协变量对匹配结果的干扰。使用熵平衡法的检验结果见表 6 第(2)列,“大数据试验区”的回归系数在 1%的水平上显著为正。

(2) 工具变量法。借鉴施炳展和李建桐(2020)^[54]的方法,使用 1996 年各省份人均函件数量作为国家大数据综合试验区设立的工具变量。一方面,地区函件数量反映了当地的通信需求,随着通信技术的发展,这种需求会延伸至大数据需求,并成为试验区设立的条件之一,满足工具变量的相关性要求;另一方面,地区函件数量的历史数据不会对当下家庭收入产生直接影响,满足工具变量的外生性要求。表 6 第(3)列和第(4)为工具变量法 2SLS 检验结果。第一阶段估计结果显示,“工具变量×2016 年虚拟变量”对“大数据试验区”的回归系数显著为正,且 F 值为 164.50(远大于经验值 10),表明不存在弱工具变量问题;Cragg-Donald Wald F 统计量为 1206.68,大于临界值,表明工具变量有效;第二阶段估计结果显示,工具变量拟合的“大数据试验区”对“家庭收入”的回归系数在 1%的水平上显著为正,表明在缓解模型内生性问题后,国家大数据综合试验区建设促进了试验区家庭收入增长的结论仍然成立。

表 6 内生性处理结果

变 量	家庭收入	家庭收入	大数据试验区	家庭收入
	(1)	(2)	(3)	(4)
大数据试验区	0.147 ^{**} (0.065)	0.112 ^{***} (0.034)		0.391 ^{***} (0.145)
工具变量×2016年虚拟变量			0.052 ^{***} (0.002)	
样本量	7 252	29 240	29 240	29 240
R ²	0.626	0.577	0.67	0.577

3. 机制检验

(1) 促进数字普惠金融发展路径。以“数字普惠金融”为中介变量的检验结果见表 7 第(1)(2)列。“大数据试验区”对“数字普惠金融”的回归系数在 1%的水平上显著为正,表明国家大数据综合试验区建设提高了试验区城市数字普惠金融发展水平;“数字普惠金融”对“家庭收入”的回归系数在 1%的水平上显著为正,表明城市数字普惠金融发展水平的提高对家庭收入具有显著的正向影响。Sobel 检验结果也显示中介效应显著。可见,数字普惠金融在国家大数据综合试验区建设影响家庭收入中确实发挥了显著的中介作用,即试验区建设可以通过促进数字普惠金融发展来增加家庭收入,假说 H2 得以验证。

(2) 促进创业发展路径。分别以“家庭创业”和“城市创业”为中介变量的检验结果见表 7 第(3)至(6)列。“大数据试验区”对“家庭创业”和“城市创业”的回归系数均显著为正,表明国家大数据综合试验区建设促进了试验区城市的创业发展;“家庭创业”和“城市创业”对“家庭收入”的回归系数均显著为正,表明城市创业发展对家庭收入具有显著的正向影响。Sobel 检验结果也显示中介效应显著。可见,创业在国家大数据综合试验区建设影响家庭收入中确实发挥了显著的中介作用,即试验区建设可以通过促进创业发展来增加家庭收入,假说 H3 得以验证。

表 7 机制检验结果

变 量	数字普惠金融发展路径		创业发展路径			
	数字普惠金融	家庭收入	家庭创业	家庭收入	城市创业	家庭收入
	(1)	(2)	(3)	(4)	(5)	(6)
大数据试验区	0.039 ^{***} (0.004)	0.169 ^{***} (0.034)	0.022 ^{***} (0.003)	0.230 ^{***} (0.031)	0.510 ^{***} (0.039)	0.114 ^{***} (0.035)
数字普惠金融		0.18 ^{***} (0.041)				
家庭创业				0.104 [*] (0.058)		
城市创业						0.053 ^{***} (0.012)
样本量	29 240	29 240	29 240	29 240	29 240	29 240
R ²	0.857	0.571	0.945	0.573	0.874	0.571
Sobel 检验	P=0.000 ^{***}		P=0.012 ^{**}		P=0.000 ^{***}	

五、进一步研究:异质性分析

1. 城乡差异

根据家庭位于农村还是城市将样本划分为“城市家庭”和“农村家庭”两组,分别检验国家大数据综合试验区建设对家庭总收入以及工资性收入和经营性收入的影响,回归结果见表8。“大数据试验区”对“家庭收入”的回归系数,“城市家庭”组在1%的水平上显著为正,而“农村家庭”组在10%的水平上显著为正;“大数据试验区”对“工资性收入”的回归系数,“城市家庭”组在5%的水平上显著为正,而“农村家庭”组为正但不显著;“大数据试验区”对“经营性收入”的回归系数,“城市家庭”组为正但不显著,而“农村家庭”组在1%的水平上显著为正。

上述结果表明:(1)相比农村地区,国家大数据综合试验区建设的家庭增收效应在城市地区更显著;(2)国家大数据综合试验区建设对家庭工资性收入增长的促进作用在城市地区更显著,且其对城市家庭的增收效应主要通过增加工资性收入来实现;(3)国家大数据综合试验区建设对家庭经营性收入增长的促进作用在农村地区更显著,且其对农村家庭的增收效应主要通过增加经营性收入来实现。原因可能在于:由于城市地区在经济发展、技术水平、数字基础设施以及人力资本等方面具有明显优势,国家大数据综合试验区建设通常首先推动城市地区的大数据发展,使得城市家庭比农村家庭能够更快更多地获取数字经济红利;相比之下,地理位置、基础设施、劳动力素质等的限制使得农村家庭能够获得的数字经济红利相对较少(王邵军,2023)^[55]。因此,国家大数据综合试验区建设对城市家庭收入增长的促进作用比农村家庭更显著。同时,大数据发展创造的就业机会更多的是在城市地区,而其对家庭创业约束(如家庭融资约束)的缓解作用在农村地区更明显(侯瑜等,2023)^[56],导致国家大数据综合试验区建设主要通过增加城市家庭的工资性收入和农村家庭的经营性收入来促进家庭收入增长。

表8 城乡异质性分析结果

变 量	家庭总收入		工资性收入		经营性收入	
	城市家庭	农村家庭	城市家庭	农村家庭	城市家庭	农村家庭
大数据试验区	0.103*** (0.030)	0.111* (0.065)	0.375** (0.167)	0.069 (0.115)	0.244 (0.186)	0.405*** (0.129)
N	8 698	20 542	8 698	20 542	8 698	20 542
R ²	0.593	0.567	0.535	0.501	0.657	0.606

2. 家庭社会资本差异

借鉴张要要(2022)^[57]的方法,采用人情费用支出来衡量家庭的社会资本水平,并根据试验区设立前受访家庭所在城市社会资本水平的中位数将样本划分为“社会资本高”和“社会资本低”两组,分别检验国家大数据综合试验区建设对家庭总收入以及工资性收入和经营性收入的影响,回归结果见表9。“大数据试验区”对“家庭收入”的回归系数,“社会资本高”组在5%的水平上显著为正,而“社会资本低”组在10%的水平上显著为正;“大数据试验区”对“工资性收入”的回归系数,“社会资本高”组在5%的水平上显著为正,而“社会资本低”组为正但不显著;“大数据试验区”对“经营性收入”的回归系数,“社会

资本高”组为正但不显著,而“社会资本低”组在1%的水平上显著为正。

上述结果表明:(1)相比社会资本水平较低的家庭,国家大数据综合试验区建设的家庭增收效应对社会资本水平较高的家庭更显著;(2)国家大数据综合试验区建设对家庭工资性收入增长的促进作用在社会资本水平较高的家庭中更显著,且其对社会资本水平较高家庭的增收效应主要通过增加工资性收入来实现;(3)国家大数据综合试验区建设对家庭经营性收入增长的促进作用在社会资本水平较低的家庭中更显著,且其对社会资本水平较低家庭的增收效应主要通过增加经营性收入来实现。其原因可能在于:数字经济本身具有显著的网络化特征,社会资本的增加有助于家庭获取更多的大数据发展红利,因而国家大数据综合试验区建设对社会资本较多家庭收入增长的促进作用比社会资本较少家庭更显著。同时,社会资本较多的家庭在大数据发展中更易实现就业,加上其对传统金融服务具有一定的依赖性,往往会选择通过就业或提高就业质量的方式来提高家庭收入(杨希雷等,2023)^[58];相对而言,社会资本较少的家庭虽然不易获得传统金融服务,但是数字普惠金融的发展使其能够获得新型金融服务,从而有更多的创业资金开展创业活动(蔡卫星等,2023)^[59]。因此,对于社会资本水平较高的家庭,国家大数据综合试验区建设主要通过促进其就业进而增加工资性收入来产生家庭增收效应;而对于社会资本水平较低的家庭,国家大数据综合试验区建设主要通过促进其创业进而增加经营性收入来产生家庭增收效应。

表9 社会资本异质性分析结果

变 量	家庭总收入		工资性收入		经营性收入	
	社会资本高	社会资本低	社会资本高	社会资本低	社会资本高	社会资本低
大数据试验区	0.136** (0.061)	0.096* (0.051)	0.364** (0.143)	0.247 (0.245)	0.085 (0.216)	0.1*** (0.033)
N	10 250	18 990	10 250	18 990	10 250	18 990
R ²	0.665	0.603	0.688	0.654	0.677	0.651

3. 家庭人力资本差异

采用户主受教育程度来衡量家庭人力资本水平,根据试验区设立前受访家庭所在城市人力资本水平的中位数将样本划分为“人力资本高”和“人力资本低”两组,分别检验国家大数据综合试验区建设对家庭总收入以及工资性收入和经营性收入的影响,回归结果见表10。在“人力资本高”组,“大数据试验区”对“家庭收入”“工资性收入”“经营性收入”的回归系数均显著为正;而在“人力资本低”组,“大数据试验区”对“家庭收入”“工资性收入”“经营性收入”的回归系数虽然为正,但均不显著。

以上结果表明,无论是总收入还是不同来源的收入,国家大数据综合试验区建设的家庭增收效应在人力资本水平较高的家庭中都更显著。其原因可能在于:一方面,试验区建设创造了更多更多元化的就业和创业机会,为家庭的生产性收入增长提供了更好条件,但是这些就业和创业的实现往往需要具备一定的知识储备,尤其是较高的数字素养(周立新等,2022)^[60],因而人力资本水平较高的家庭更有可能从中获益;另一方面,试验区建设还可以通过发展数字普惠金融来提高家庭的金融可获得性,从而有助于家庭收入增长,但是数字普惠金融本身对金融素养以及互联网使用方法和技巧有一定要求,因而人力资本水平较高的家庭在金融可获得性方面的改善也更显著(王修华等,2022)^[61]。

表 10 人力资本异质性分析结果

变量	家庭总收入		工资性收入		经营性收入	
	人力资本高	人力资本低	人力资本高	人力资本低	人力资本高	人力资本低
大数据试验区	0.110 ^{***}	0.074	0.483 ^{**}	0.045	0.321 ^{***}	0.289
	(0.034)	(0.045)	(0.193)	(0.071)	(0.116)	(0.290)
N	4 926	24 314	4926	24 314	4926	24 314
R ²	0.594	0.574	0.604	0.548	0.674	0.622

六、结论与启示

持续稳定地提高家庭收入是实现共同富裕的必要条件。大数据发展不仅为经济增长提供了新动能,而且其带来的大量新产业、新业态、新模式涌现为家庭通过参与生产活动增加收入提供了条件。以国家大数据综合试验区带动大数据发展是加快数字中国建设的重要举措,能够促进技术进步、优化要素配置,并通过促进数字普惠金融发展和创新创业发展等多渠道增加家庭收入。本文以国家大数据综合试验区建设为准自然实验,采用中国家庭金融调查 2013 年、2015 年、2017 年、2019 年 4 期面板数据,实证检验国家大数据综合试验区建设对家庭收入的影响及其机制,研究发现:(1)国家大数据综合试验区建设显著提高了试验区内的家庭收入,该结论在经过样本匹配法和工具变量法的内生性处理以及平行趋势检验、增加固定效应、更换被解释变量、控制相关政策影响、反事实检验、安慰剂检验等一系列稳健性检验后仍然成立;(2)国家大数据综合试验区建设对收入水平较低家庭的收入增长有更强的促进作用,有利于缩小家庭收入差距;(3)国家大数据综合试验区建设能够显著增加家庭的工资性收入和经营性收入,但对财产性收入和转移性收入的影响不显著,表明大数据发展主要通过提高生产性收入来促进家庭收入增长;(4)数字普惠金融和创业在国家大数据综合试验区建设影响家庭收入中发挥了显著的中介作用,即国家大数据综合试验区建设可以通过促进数字普惠金融发展和创业发展的路径来增加家庭收入;(5)国家大数据综合试验区建设对家庭总收入和工资性收入增长的促进作用在城市家庭、社会资本水平较高家庭、人力资本水平较高家庭中更显著,而对经营性收入增长的促进作用在农村家庭、社会资本水平较低家庭、人力资本水平较低家庭中更显著;(6)国家大数据综合试验区建设对城市家庭、社会资本水平较高家庭、人力资本水平较高家庭的增收效应主要通过增加工资性收入来实现,而对农村家庭、社会资本水平较低家庭、人力资本水平较低家庭的增收效应主要通过增加经营性收入来实现。

基于本文研究结论,可以得到以下启示:第一,筑牢数字基础设施底座,持续放大数字经济红利。政府应重视大数据基础设施建设与升级,特别是在农村和经济欠发达地区,要提高互联网覆盖范围和接入速度,确保数字服务的稳定性、持续性和安全性。同时,应通过公共教育、职业培训提高居民数字素养,引导数据要素与人力资本相结合,鼓励和帮助居民利用数字技术增强自身职业竞争力,进而提高家庭收入。第二,推动数字普惠金融发展,营造公平的创新创业环境。一方面,政府应当不断推进数字普惠金融发展,大力支持金融机构开展数字普惠金融业务,保证金融服务的连续性和稳定性。同时,要积极建设创业孵化基地,整合资本、技术、劳动等资源要素;搭建产学研一体化平台,促进知识共享和传播;简化企业注册和运营流程,降低创业门槛。另一方面,金融机构应利用大数据、云计算、物联网等数字技术提供更精准更便捷的金融服务,特别是对小微企业、个体工商户和农村家庭,应不断提高信贷可得性和易得性,全面降低金融服务门槛。第三,因地制宜推进大数据发展,全面增加家庭收入来源。政府通过发

展数字经济提升社会福利,应当以协调发展和共享发展是关键。针对城市地区家庭、社会资本和人力资本水平较高家庭,应当引导其参加数字技能培训,通过提升劳动者技能溢价、满足新兴产业发展需求等进一步提升其工资收入;对于农村家庭、社会资本和人力资本水平较低家庭,应助力其创业发展,促进数字技术在生产经营各个环节的应用,进一步提升其经营收入。

参考文献:

- [1] 苏锦旗,唐诗瑶,张营营. 国家级大数据综合试验区能否促进区域经济高质量发展——基于试点区政策的准自然实验[J]. 现代财经(天津财经大学学报),2023,43(10):56-73.
- [2] 杨玉琪,王小华. 数字化赋能与城市经济高质量发展——基于国家级大数据综合试验区的准自然实验[J]. 经济问题探索,2023,(12):105-123.
- [3] 郭炳南,王宇,张浩. 数字经济发展改善了城市空气质量吗——基于国家级大数据综合试验区的准自然实验[J]. 广东财经大学学报,2022,37(1):58-74.
- [4] 张营营,许钊,彭硕毅. 大数据发展对中国城市低碳转型的影响及其机制研究[J]. 城市问题,2023(9):34-43+65.
- [5] 刘占芳,杨彭宇. 数字经济、环境规制与减霾效应——基于国家级大数据综合试验区的实证研究[J]. 工程管理科技前沿,2024,43(2):69-75.
- [6] 徐林,侯林岐,程广斌. 国家级大数据综合试验区创新效应研究[J]. 科技进步与对策,2022,39(20):101-111.
- [7] 刘传明,陈梁,魏晓敏. 数据要素集聚对科技创新的影响研究——基于大数据综合试验区的准自然实验[J]. 上海财经大学学报,2023,25(5):107-121.
- [8] 胡婧玮,郭金花,陈鑫. 大数据试验区建设对城市创新效率的影响——基于国家大数据综合试验区试点的准自然实验[J]. 商业研究,2024(2):104-113.
- [9] 魏丽莉,修宏岩,侯宇琦. 数字经济对城市产业生态化的影响研究——基于国家级大数据综合试验区设立的准自然实验[J]. 城市问题,2022(11):34-42.
- [10] 钟昌标,卢建霖. 大数据试验区建设推动我国工业绿色转型了吗[J]. 江西社会科学,2023,43(1):122-133.
- [11] 陈启斐,田真真. 大数据与产业赋能——基于国家级大数据试验区的分析[J]. 南开经济研究,2023(7):90-107.
- [12] 邱子迅,周亚虹. 数字经济发展与地区全要素生产率——基于国家级大数据综合试验区的分析[J]. 财经研究,2021,47(7):4-17.
- [13] 刘军,朱可,钱宇. 数字经济对全要素生产率的影响研究——来自国家级大数据综合试验区的证据[J]. 南京审计大学学报,2024,21(1):101-111.
- [14] 彭国柱,周湘莲. 数据要素配置与共同富裕:基于国家级大数据综合试验区的准自然实验[J]. 经济问题探索,2024(4):87-102.
- [15] 赵鹏,朱叶楠,赵丽. 国家级大数据综合试验区与新质生产力——基于230个城市的经验证据[J]. 重庆大学学报(社会科学版),2024,30(4):62-78.
- [16] 曹平,陆松,梁明柳. 大数据战略、知识管理能力与中国企业创新[J]. 产经评论,2021,12(2):102-119.
- [17] 戴艳娟,沈伟鹏,谭伟杰. 大数据发展对企业数字技术创新的影响研究——基于国家大数据综合试验区的准自然实验[J]. 西部论坛,2023,33(2):16-28.
- [18] 孙伟增,毛宁,兰峰,等. 政策赋能、数字生态与企业数字化转型——基于国家大数据综合试验区的准自然实验[J]. 中国工业经济,2023(9):117-135.
- [19] 石玉堂,王晓丹. 大数据综合试验区设立能否驱动企业数字化转型?——基于准自然实验的实证研究[J]. 科学学研究,2024,42(7):1482-1492.
- [20] 张益豪,郭晓辉. 大数据发展与企业全要素生产率——基于国家级大数据综合试验区的实证分析[J]. 产业经济研究,2023(2):69-82.

- [21] 孙鹏,柳力群,周可懂.数字经济与企业全要素生产率——来自国家级大数据综合试验区的证据[J].海南大学学报(人文社会科学版),2023,41(5):114-123.
- [22] 卜寒,高远东,寻舟.大数据如何影响劳动收入份额?——来自国家级大数据综合试验区的证据[J].南方经济,2023(11):62-82.
- [23] 戴艳娟,沈伟鹏.大数据发展与企业劳动收入份额——来自大数据综合试验区的证据[J].南京财经大学学报,2024(3):100-110.
- [24] 李麦收,李华.大数据发展与企业劳动收入份额——基于国家大数据综合试验区的准自然实验[J].管理学报,2024,37(4):14-29.
- [25] 宋华盛,卢历祺.大数据发展和企业劳动收入份额——来自“国家级大数据综合试验区”的证据[J].经济学动态,2024(1):111-128.
- [26] 杨俊,李小明,黄守军.大数据、技术进步与经济增长——大数据作为生产要素的一个内生增长理论[J].经济研究,2022,57(4):103-119.
- [27] 刘德林,周冬.大数据产业发展与地方经济增长[J].统计与决策,2021,37(19):102-105.
- [28] 莫赞,赵琦智,唐华奕.大数据政策的技术创新效应研究[J].广西大学学报(哲学社会科学版),2021,43(5):142-149.
- [29] 孙洁,李杰.大数据应用、融资约束和企业创新效率[J].证券市场导报,2022(11):13-23.
- [30] 任英华,刘宇钊,胡宗义,等.大数据发展、知识产权保护对企业绿色技术创新的影响[J].中国人口·资源与环境,2023,33(7):157-167.
- [31] 张红春,邓剑伟,邱艳萍.大数据驱动的透明政府建设——媒介选择与政民互动重构[J].北京理工大学学报(社会科学版),2020,22(4):60-69.
- [32] 赵术高,李珍.信息不对称、交易费用与国家治理现代化——兼论大数据在国家治理现代化中的应用逻辑[J].财经问题研究,2021(4):28-36.
- [33] GARDINER A,AASHEIM C,RUTNER P,et al. Skill requirements in big data;a content analysis of job advertisements [J]. Journal of Computer Information Systems,2018,58(4):374-384.
- [34] ACEMOGLU D,RESTREPO P. Robots and jobs:evidence from US labor markets[J]. Journal of Political Economy,2020,128(6):2188-2244.
- [35] 侯冠宇,熊金武.数字经济对共同富裕的影响与提升路径研究——基于我国30个省份的计量与QCA分析[J].云南民族大学学报(哲学社会科学版),2023,40(3):89-99.
- [36] 陶长琪,丁煜.数据要素何以成为创新红利?——源于人力资本匹配的证据[J].中国软科学,2022(5):45-56.
- [37] 尹志超,张紫璇,岳鹏鹏.创业与中国家庭收入流动[J].数量经济技术经济研究,2024,41(2):68-88.
- [38] 祁怀锦,李若琳,刘斯琴.数字化转型的公司治理效应:基于管理层在职消费视角[J].改革,2024(4):108-125.
- [39] 刘典.金融-科技-产业:金融强国战略的三元结构和历史演进[J].金融经济研究,2024,39(1):50-59.
- [40] 张颖熙,夏杰长.科技向善赋能共同富裕:机理、模式与路径[J].河北学刊,2022,42(3):115-122.
- [41] 刘元维,华桂宏,庞思璐.数字金融发展、资本跨区流动与产业结构高级化[J].西部论坛,2023,33(6):1-16.
- [42] 任大增,殷志高.数字普惠金融与中国经济的包容性增长:理论分析和经验证据[J].管理学报,2022,35(1):23-35.
- [43] 刘艳华,余畅婉.数字鸿沟阻碍农村家庭金融投资了吗?——基于2018年CFPS微观数据的实证研究[J].财贸研究,2023,34(11):51-61.
- [44] 王小华,刘云,宋檬.数字能力与家庭风险金融资产配置[J].中国农村经济,2023(11):102-121.
- [45] 孙俊娜,胡文涛,汪三贵.数字技术赋能农民增收:作用机理、理论阐释与推进方略[J].改革,2023(6):73-82.
- [46] 石玉堂,王晓丹.大数据综合试验区设立能否驱动企业数字化转型?——基于准自然实验的实证研究[J].科学学研究,2024,42(7):1482-1492.

- [47] 陈明生,王乾坤. 数字基础设施的家庭收入效应——基于中国家庭金融调查(CHFS)的实证研究[J]. 湘潭大学学报(哲学社会科学版),2024,48(2):46-53.
- [48] 戚聿东,丁述磊,刘翠花. 数字经济时代互联网使用对灵活就业者工资收入的影响研究[J]. 社会科学辑刊,2022(1):125-138+2.
- [49] 肖伟,刘文华,谢婷. 就地城镇化的家庭收入效应——基于中国家庭金融调查(CHFS)的实证研究[J]. 金融研究,2023(2):152-170.
- [50] 尹志超,蒋佳伶,严雨. 数字鸿沟影响家庭收入吗[J]. 财贸经济,2021,42(9):66-82.
- [51] 宋冬林,田广辉,徐英东. 数字金融改善了收入不平等状况吗?——基于创业的收入与就业效应研究[J]. 兰州大学学报(社会科学版),2022,50(3):38-51.
- [52] 温永林,张阿城. 信息基础设施建设能促进创业吗?——基于“宽带中国”示范城市建设的准自然实验研究[J]. 外国经济与管理,2023,45(7):138-152.
- [53] 尹志超,郭沛瑶. 精准扶贫政策效果评估——家庭消费视角下的实证研究[J]. 管理世界,2021,37(4):64-83.
- [54] 施炳展,李建桐. 互联网是否促进了分工:来自中国制造业企业的证据[J]. 管理世界,2020,36(4):130-149.
- [55] 王邵军. 数字乡村建设促进农民农村共同富裕的作用机理、现实挑战与实现路径研究[J]. 南开经济研究,2023(11):21-36.
- [56] 侯瑜,袁鹏妞. “宽带中国”示范城市建设有助于缩小城乡收入差距吗? [J]. 西部论坛,2023,33(2):96-110.
- [57] 张要要. 数字鸿沟与农户家庭创业[J]. 山西财经大学学报,2022,44(2):103-114.
- [58] 杨希雷,黄杏子. 非正规就业抑制了居民家庭收入向上流动吗[J]. 经济学家,2023(8):43-53.
- [59] 蔡卫星,韦庆芳,林航宇. 数字金融发展的劳动力需求效应——来自2000万在线招聘岗位的经验证据[J]. 金融研究,2023(10):28-46.
- [60] 周立新,屈彩萍,王淑敏. 数字素养的农户收入增长效应研究[J]. 西部论坛,2024,34(2):40-54.
- [61] 王修华,赵亚雄. 数字金融发展与城乡家庭金融可得性差异[J]. 中国农村经济,2022(1):44-60.

Household Income Increasing Effect of the National Big Data Comprehensive Pilot Zone

WANG Qian-kun

(School of Business, China University of Political Science and Law, Beijing 100088, China)

Abstract: With the advent of the fourth industrial revolution, information collection, calculation, and statistics based on big data technology provide practical tools for solidly promoting common prosperity. The “14th Five-Year Plan” emphasizes the need to improve the income distribution system and increase residents’ income through multiple channels. This requires that the economic effects of big data should not only be considered from the perspective of “distributing the cake” but also from the perspective of “making the cake bigger”. However, there is a lack of research in the existing literature on understanding common prosperity from the perspective of “making the cake bigger”, especially the lack of in-depth research on how big data can practice its mission value in promoting common prosperity.

This paper takes the national-level big data comprehensive pilot zones as a quasi-natural experiment. It uses the four-round survey samples of the China Household Finance Survey (CHFS) from 2013 to 2019 to explore the causal relationship between big data development and household income. The empirical study finds

that: firstly, the construction of big data pilot zones significantly increases household income, and the income increase effect is more evident for low-income groups. The conclusion still holds after a series of robustness tests, such as the parallel trend test, exclusion of policy interference, and counterfactual test. Secondly, big data pilot zones primarily enhance wage income and operating income among household income sources. Specifically, for groups with high social capital and high human capital in urban areas, the big data pilot zones mainly increase their wage income; for groups with low social capital and high human capital in rural areas, the big data pilot zones mainly increase their operating income. Thirdly, the big data pilot zones mainly increase family income by promoting the development of digital inclusive finance and increasing regional entrepreneurial opportunities. In addition, as the distance from the pilot provinces gradually increases, the family income increasing effect of the big data pilot zones shows a trend of initially rising and then falling.

Compared with previous literature, this paper expands on three aspects. Firstly, in terms of research perspective, it enriches the study of the economic benefits of big data development and influencing factors of household income in the context of the digital economy by investigating the impact of big data pilot zones on household income as a quasi-natural experiment. Secondly, in terms of research methodology, it identifies the causal relationship between big data development and household income by treating the national big data comprehensive pilot zone as an exogenous policy event and employing methods such as the difference-in-differences model and instrumental variable approach to effectively address endogeneity issues. Thirdly, in terms of research content, the study not only verifies the mechanisms through which big data development promotes household income growth but also explores the differentiated economic effects and quantile effects of big data pilot zones through subsample regressions, providing references for enriching big data technology application scenarios.

The research reveals the inherent logic of how big data development promotes household income improvement, which may assist government departments in formulating more targeted development strategies in the practice of common prosperity, providing differentiated policy support for different regions and groups, and better promoting the sharing of development dividends of the digital age among all people, thus contributing to achieving common prosperity through high-quality development.

Key words: big data; household income; digital inclusive finance; entrepreneurship development; wage income; operating income

CLC number: F124. 7; F207

Document code: A

Article ID: 1674-8131(2024)05-0001-17

(编辑:刘仁芳)