

国家级大数据综合试验区设立与就业增长

沈坤荣, 乔刚, 谭睿鹏

[摘要] 就业是最基本的民生,要突出就业优先导向,推进多措并举稳就业促增收。本文基于2010—2021年上市公司数据,使用双重差分法,实证考察了国家级大数据综合试验区设立对企业就业的影响及作用机制。研究发现,大数据试验区设立能够有效释放数据要素价值,提升企业就业水平,这一效应在城市层面同样得到了验证。机制检验表明,大数据试验区设立通过新岗位创造、市场扩大和融资约束缓解等渠道提升企业就业水平。进一步分析发现,大数据试验区设立改善了员工福利报酬和企业声誉,吸引了更多劳动力流入。基于就业结构的分析表明,大数据试验区设立对不同经济地位、社会声望职业的就业均有显著的提升效应,也带动了各类教育背景员工就业的增加。本文还使用机器学习方法,比较了政策最优执行规则与实际执行情况,研究表明大数据试验区政策在就业促进方面仍存在较大的优化空间。本文的研究揭示了大数据试验区设立的就业促进效应,为进一步推动大数据技术应用与实现稳就业目标提供了重要的政策启示。

[关键词] 数字经济; 国家级大数据综合试验区; 就业增长

[中图分类号] F126 **[文献标识码]** A **[文章编号]** 1006-480X(2024)12-0005-19

一、引言

就业是最基本的民生,事关人民群众切身利益,事关经济社会健康发展,事关国家长治久安。当前,面对复杂严峻的国内外形势和多重超预期因素冲击,企业经营遭遇困难,导致就业形势严峻(崔小勇等,2023)。根据国家统计局数据,2023年中国城镇地区的平均失业率下降至5.2%,但青年劳动力的失业率仍居高不下。^①随着外部环境不确定性因素增加,企业发展面临诸多挑战,稳就业的基础尚不牢固,就业压力依然较大。党的二十大报告指出,实施就业优先战略,强化就业优先政策,健全就业促进机制,促进高质量充分就业。在数字经济与实体经济深度融合的背景下,数字产业的蓬勃发展为经济增长提供了强劲动力,并对就业产生了深远影响。数字技术作为支撑整个数字产业发展壮大的核心,其功效的发挥在很大程度上依赖于数据资源这一重要传导载体。然而,鲜有学者关注数据资源能否以及如何影响微观企业就业吸纳,针对这一问题的深入探讨还较为缺乏。

[收稿日期] 2024-05-31

[基金项目] 国家社会科学基金重大项目“推动经济实现质的有效提升和量的合理增长研究”(批准号24&ZD044)。

[作者简介] 沈坤荣,南京大学商学院教授,博士生导师,经济学博士;乔刚,南京大学商学院博士研究生;谭睿鹏,合肥工业大学经济学院教授,经济学博士。通讯作者:乔刚,电子邮箱:DG21020015@smail.nju.edu.cn。感谢匿名评审专家和编辑部的宝贵意见,文责自负。

^① 2023年12月,16—24岁、25—29岁的青年劳动力失业率分别为14.9%、6.1%,均高于当年平均失业率。

当前,数据资源已跃升为一种生产要素,成为建设数字中国和实现高质量发展的重要抓手。中共中央政治局第二次集体学习时对“实施国家大数据战略加快建设数字中国”进行了讨论,明确强调推动大数据战略的实施,完善数字基础设施建设,推进数据资源开放与共享,利用大数据技术赋能社会保障和民生改善。2016年,国家发展和改革委员会等部门正式确立贵州等八大地区为国家级大数据综合试验区(简称大数据试验区),旨在发展大数据技术,搭建大数据平台,加快数据要素的流通与共享。2024年的《政府工作报告》指出,适度超前建设数字基础设施,推动数据资源开放与流通,支持平台企业在促进创新、增加就业中大显身手。国家互联网信息办公室发布的《数字中国发展报告(2022年)》显示,2022年中国大数据产业规模为1.57万亿元,数据资源存储量达到724.5EB,同比增长21.1%,占全球数据总储量的14.4%。面对如此庞大的数据资源,如何推动数据要素价值充分释放,助力数字经济高质量发展显得至关重要。2023年2月,《数字中国建设整体布局规划》进一步强调,畅通数字资源大循环的关键在于打通数字基础设施“大动脉”,促进数字企业发展壮大。鉴于此,大数据基础设施建设已成为数据资源流通的重点,对于实现稳就业和高质量发展具有重大意义。

在经济下行和就业压力倍增的背景下,如何借助大数据技术激发数据要素价值,充分发挥数据要素的稳就业效应已得到广泛关注。已有文献就新兴技术应用对劳动力市场的影响进行了大量有意义的探索(Acemoglu and Restrepo, 2020; Acemoglu and Johnson, 2024; Adachi et al., 2024; Hui et al., 2024)。Autor(2015)指出,过去几十年间,自动化并未完全取代大部分工作岗位;相反,机器与人类的比较优势相互作用使得机器替代工人执行常规的、可编码的任务,同时体现了人类在解决问题、适应性和创造力等方面的比较优势。

多数研究认为当下技术进步对就业具有双重影响,即消极的破坏效应和积极的创造效应。具体而言,一部分文献,如Acemoglu and Restrepo(2020)、王永钦和董雯(2020)、Acemoglu et al.(2022)等,强调了机器人和人工智能对就业的替代作用,减少了低技能就业岗位,但同时刺激了智能化相关职位需求的快速增长,表现为技能偏向型技术进步;另一部分文献则探讨了信息基础设施建设(孙伟增和郭冬梅,2021)、数字经济(田鸽和张勋,2022)、人工智能(尹志锋等,2023)等对就业的影响,发现新兴数字技术对就业具有促进作用。Autor(2024)研究表明,人工智能等新一代信息技术扩展了人类专业知识的相关性和范围,使得具备互补知识的工人得以执行一些原本由精英专家承担的高风险决策任务,如律师的文件制作、计算机工程师的软件编码等,为劳动力市场提供了独特的就业机会。Autor(2024)进一步指出,如果人工智能使用得当,将有助于恢复被自动化挤出的工作岗位。尽管已有文献对新兴技术的就业效应进行了大量探讨,但未形成一致结论,即新兴技术对就业的影响具有不确定性。作为数字技术发挥作用的主要传导媒介,数据资源和数据要素能否以及如何影响劳动力市场,仍是一个有待深入探讨的问题。厘清这一问题对于制定和优化数据要素发挥促就业效应的产业政策至关重要。

本文基于国家级大数据综合试验区设立这一独特视角,利用2010—2021年上市公司和地级市数据,考察了大数据试验区设立对就业的影响。研究发现,大数据试验区设立显著促进了企业雇佣劳动增加。具体地,大数据试验区的设立使得试点地区的就业增加7.13%。机制检验表明,新岗位创造、市场扩大和融资约束缓解成为大数据试验区促进就业增长的主要影响渠道。进一步分析表明,大数据试验区设立不仅提高了员工福利报酬,还提升了企业声誉,有利于吸引员工流入。基于就业结构视角的分析显示,从职业特征看,大数据试验区设立增加了不同经济地位和社会声望的职业就业;从学历结构看,大数据试验区设立对各类教育背景员工的就业均有促进作用。最后,本文

通过机器学习方法,以促进就业为政策目标,计算了最优政策执行规则,发现大数据试验区政策的优化空间较大。

本文的研究创新点在于:①为数据要素就业增长效应的研究提供了微观证据。现有文献主要从理论和宏观层面探讨数据要素的创新与增长效应(杨俊等,2022;徐翔等,2023;刘涛雄等,2023;方锦程等,2023),缺乏对微观企业层面的关注。本文基于大数据试验区设立这一外生政策冲击的独特视角,利用双重差分法量化评估了大数据试验区设立对就业的影响,从微观视角丰富和拓展了大数据试验区经济效应的相关研究。②深化和补充了新兴技术对劳动力市场影响的相关研究。与本文主题较为相关是数字经济中的诸多重要因素对就业的影响,包括机器人、数字技术专利、网络基础设施以及人工智能等(王永钦和董雯,2020;Cirillo et al., 2021;孙伟增和郭冬梅,2021;尹志锋等,2023)。数据资源作为数字经济发展的关键要素,具有非排他性和非稀缺性,是推动技术发展的重要基础。与上述研究视角不同,本文直接评估了数据作为生产要素的经济价值。③揭示了大数据试验区影响就业的作用机制。本文利用丰富的微观数据,从多个维度系统考察了大数据试验区设立对就业影响的潜在机制,补充了数据资源发挥促就业效应的潜在机理。④通过分析政策实施效果,提出进一步优化政策的可能性。本文在市场均衡处理效应的框架中,使用机器学习方法,以最大化大数据试验区设立的就业效应为政策目标,比较了政策最优执行规则和实际执行情况,表明大数据试验区政策存在较大优化空间,为后续政策优化提供了重要启示。

二、制度背景与理论分析

1. 大数据试验区的发展历史及现状

2014年中国政府首次将大数据的重要性写入《政府工作报告》。2015年正式拉开中国大数据发展的帷幕,并在《促进大数据发展行动纲要的通知》中首次提出推进贵州等大数据综合试验区的设立,完善区域性大数据基础设施建设和数据资源的汇集应用,强调加快大数据部署,深化大数据应用,提升经济运行效率,使得大数据发展成为促增长和稳就业的重要保障。同年9月,贵州大数据综合试验区建设正式启动,力争通过构建“三大体系”、打造“七大平台”、实施“十大工程”将贵州大数据试验区建设成为全国数据资源汇集新高地和政策创新的先行区。“十三五”规划指出,加快推进数字基础设施建设,构建全国一体化大数据中心,促进大数据与实体经济深度融合。2016年,国家发展和改革委员会等部门正式批复同意贵州建设全国首个大数据试验区,明确提出构建大数据流通平台,健全数据资源流通机制,促进大数据跨行业、跨地区流通。同年10月,国家发展和改革委员会等部门批准启动了第二批大数据试验区建设试点项目,覆盖了京津冀地区、珠江三角洲、上海市、河南省、重庆市、沈阳市和内蒙古自治区七个地理区域。最终,设立的诸多大数据试验区在大数据开放与共享、数据要素流通等方面进行探索式试验,搭建起中国大数据发展的立体框架。

大数据试验区建设促进了大数据技术和数据平台的发展,有助于企业创造新岗位和改善生产经营活动。例如,贵州不仅建立大数据金融中心,促进大数据资源与资金等生产要素充分融合,还打造基于社会服务的大数据金融平台,有利于缓解企业融资约束、促进企业市场扩张和就业增长。截至2020年,贵州省拥有5000多家大数据企业和23个正在运营或建设中的数据中心,形成了千亿级的电子信息制造业、软件和信息技术服务业等产业集群,吸引了阿里、华为、腾讯、百度等国内大数据巨头入驻。毋庸置疑,新企业进入将创造更多的就业岗位,推动了当地的就业增长。根据贵州

省统计局数据,2015年,与大数据产业相关的信息传输、软件与信息技术服务业的从业人数为3.26万人,截至2022年增长到4.52万人,年均增长率为4.78%。此外,贵州利用大数据优势提升城市管理水平,创新社会服务,同时重视人才引进,促进产学研深度合作,从而吸引更多劳动力流入,进一步促进就业增长。河北省提出建设全国性“大数据谷”,集聚大数据龙头企业,打造全链条大数据产业,形成跨大数据生命周期的产业生态,吸引更多新企业进入。珠三角大数据试验区努力打造创业创新生态区,营造富有活力的大数据创业创新氛围,以吸引更多企业进入,提高就业水平。广东省推进社保大数据应用,构建人社大数据平台,提高就业精准扶持,对社保基金、就业专项基金以及劳动用工进行智能监管,并根据就业形式、劳动力市场供求、职业发展和薪资待遇进行研判,有利于提高就业水平和实现高质量就业。

可以预见,大数据试验区设立促进数据要素充分流通,释放数据要素价值,激发企业活力,进而提升企业就业水平。基于上述背景,本文将大数据试验区政策作为准自然实验,利用丰富的微观数据,考察了大数据试验区有效释放的数据要素价值对企业就业的影响,并探究其潜在影响机制。本文系统论证了大数据试验区设立对就业的影响,不仅是对现有文献的补充和拓展,还为后续政策配套提供重要参考。

2. 理论分析

人类历史上每一次技术进步都会对劳动力市场产生广泛而深远的影响,既带来了新的就业机会,也使部分传统岗位被替代。即技术进步对就业的影响存在不确定性,取决于负面的替代效应和正面的创造效应孰大孰小。其中,就业的替代效应是指新兴技术与劳动力相比更具优势,技术进步会以“机器换人”的方式将劳动力挤出岗位,减少企业对劳动力的需求;相反地,就业的创造效应指的是新技术催生出新经济业态,形成新产业、新模式和新职业,进而增加企业对劳动力的需求(尹志锋等,2023)。基于任务理论模型,Autor et al.(2003)在统一任务框架下考虑新技术与劳动力的比较优势,认为技术进步与常规任务的劳动投入减少以及非常规任务的劳动投入增加有关。因此,大数据技术的发展在更换生产方式、增强企业生产力以及催生新产业的同时,不可避免地会淘汰落后产业,可能导致劳动密集型岗位减少甚至消失。这种正负影响的交织,使得大数据技术对就业的影响充满复杂性和不确定性。

然而,一直以来,技术进步并未消灭大部分工作岗位,其根本原因在于新技术替代劳动力的同时也创造了新的就业机会,且新技术与劳动力之间存在强大互补性(Autor, 2015)。这一互补作用不仅提高了生产力,也为新产业和新岗位的创造奠定了基础。作为中国首个致力于推动数据基础设施建设、数据资源共享以及最大化数据要素价值的数字经济试验性政策,大数据试验区设立对就业的影响包括直接和间接两个方面。在直接影响方面,大数据试验区推动大数据技术和平台的发展,有利于形成新业态、新模式和新职业,创造了大量新岗位,尤其是增加与大数据技术相关领域非常规岗位的劳动力需求。不仅如此,大数据试验区还通过打造数字政府与数字平台,推动了数据资源开放与共享,促进当地新创企业数量的增加(Prantl, 2012),从而创造更多就业岗位(Hjort and Poulsen, 2019)。在间接影响方面,大数据试验区设立能够降低企业获取市场信息的成本,提高了生产效率,扩大了市场份额,从而增加企业对劳动力的需求。此外,大数据技术在降低融资双方的信息不对称和交易成本方面能够发挥重要作用,有效缓解企业面临的融资约束,促使企业追加投资,进行扩大再生产,进而增加就业。鉴于此,本文基于大数据试验区的特征,从新岗位创造效应、市场扩大效应以及融资约束缓解效应三个方面,探讨大数据试验区设立影响就业的潜在机制。

(1)新岗位创造效应。通常,新旧岗位更替与技术进步相伴而生。随着数字技术广泛应用,众

多新的就业场景不断涌现,新旧职业更替日趋频繁。以大数据和人工智能为代表的新一代信息技术,通过学习和模拟人类行为,不仅加深了对劳动的替代程度,还大幅减少了程序化、流程化的常规工作任务(王林辉等,2023)。例如,大数据和人工智能技术逐渐取代了客服、翻译、收银、驾驶等可编码任务。尽管如此,技术进步对职业属性的影响是动态的,许多曾经作为非常规型“新职业”,最终也变成被新技术替代的常规“旧职业”(陈琳等,2024)。回顾历次工业革命,新的技术进步破坏了旧工作岗位,但也创造了新职业和新岗位,这是以自动化为代表的技术进步并未造成大规模失业的重要原因(Autor, 2015)。

首先,大数据试验区以大数据技术催生新岗位。大数据试验区的设立推动了大数据技术发展,有助于促进专业化分工效率提升和分工细化,进而创造出新工作任务、新职业和新岗位(王林辉等,2023),推动部分劳动者从传统职业转向新职业岗位,如数据分析师、大数据工程师、数据标注员等。其次,大数据技术降低技能门槛催生新业态。大数据技术并非旨在取代劳动力,只是简化了工作任务,降低了工人执行工作任务中所需的技能要求,并拓展了工人掌握专业知识的相关性和应用范围(Autor, 2024),使得劳动者群体能够执行一些原本由专业人士从事的技术岗位。例如,在医疗领域,以大数据支撑的图像识别和计算机视觉技术能够辅助医技人员完成识别和诊断任务,从而减少医生的工作量,提高整个诊断流程的效率。由此可见,大数据技术作为互补性劳动工具是改善劳动生产率的有效途径,能够提升劳动力要素的边际生产力,促进劳动力需求的增加(尹志锋等,2023)。伴随着中国互联网技术的快速发展,大数据技术和数据资源的充分流动催生了电子商务、直播带货等诸多新业态,塑造了智能营销和智慧物流等劳动辅助工具,有效扩大了新岗位创造效应的作用范围。最后,大数据技术催生认知型和交互型任务创造新岗位。新的技术进步促使更多依赖人的理解、分析、决策的认知型任务,以及需要人际互动、沟通、决断的交互型任务不断涌现。这些非常规工作任务难以通过编程解决,需要人机协同的方式执行任务,从而催生出新的工作岗位。例如,金融科技在银行中的应用创造了数据分析师和风险管理师岗位。Acemoglu et al.(2022)指出,2010—2018年美国劳动力市场与人工智能相关的空缺职位和招聘人数快速增长,而受人工智能影响越大的企业,招聘人工智能领域的人才越多。

(2)市场扩大效应。大数据技术能够直接提高企业生产效率,并通过扩大生产规模来抢占市场份额。理论上,大数据通过提升生产率和拓展生产规模发挥市场扩大效应。孙伟增和郭冬梅(2021)研究发现企业生产率的提升增强了其市场竞争力,使企业市场份额进一步扩大,进而增加劳动力雇佣。因此,市场扩大能够促进企业就业增长。首先,以推动企业转型和优化要素配置促进生产率提升。大数据试验区驱动企业数字化转型,转变传统生产方式,利用新技术和数据信息制定更为科学合理的生产方案,提升生产效率(孙伟增和郭冬梅,2021;孙伟增等,2023);并且,数据要素具有非排他性和非稀缺性特征,在一定程度上能够替代相对稀缺的传统生产要素,进行要素合理配置并提升传统要素的使用效率。生产率提升能够提高企业市场竞争力,抢占市场份额,增加企业对劳动力的雇佣。其次,以更有效的数据平台获取关键信息。大数据平台降低了企业获取市场信息的搜集和验证成本(Goldfarb and Tucker, 2019),有助于企业从海量的数据信息中筛选出重要的商机信息,以便开拓新市场,扩大生产规模;大数据收集的海量信息能够提高市场透明度,降低企业投资和创新风险,增加新增投资和创新成功的概率,进而扩大生产规模。最后,以优质信号传递提升企业形象并扩大市场规模。根据信号传递理论,拥有高质量产品或服务的企业会通过大数据平台向公众传递优质企业信息,帮助企业树立可靠和诚信的形象(毛其淋和王玥清,2023)。消费者通常会青睐消费优质产品,增加这类企业产品或服务的消费需求,进一步促使企业扩大生产规模。市场规

模扩大促使企业增加对内部岗位的劳动力雇佣需求,最终表现为企业层面的就业增长(尹志锋等, 2023)。

(3)融资约束缓解效应。资金约束是影响企业雇佣劳动的重要原因。大数据平台减少融资双方的信息不对称和交易成本,改善企业融资水平(He et al., 2024)。一方面,大数据试验区降低信息不对称提高企业融资能力。企业和投资者之间信息不对称是融资限制的主要原因(Myers and Majluf, 1984),大数据技术能够提高企业生产经营过程的可视化程度,以便投资者实时访问企业内部管理流程、生产流程和财务管理等重要经营过程;企业还能通过大数据平台向金融机构和潜在投资者传递优质形象,降低投资者和企业家之间的信息不对称,提高信贷资源可获得性(Balakrishnan et al., 2014)。另一方面,大数据试验区以政府优惠政策和资金支持促进企业扩大生产。在设立大数据试验区的起始阶段,地方政府根据自身的具体情况和发展目标,出台了诸多政策来促进本地大数据的发展。例如,贵州省除了为大数据公司提供贷款补贴外,还对满足特定条件的企业实施了研发费用额外扣除的激励措施;珠三角地区引导社会设立大数据产业发展基金,开展股权投资和创新投资以满足大数据企业发展需求。政府补贴能够通过培养与引进数字人才、补充稀缺要素资源以及为企业生产经营兜底等方式提升企业数字化绩效(孙伟增等, 2023),进而影响企业市场规模和劳动力雇佣。融资约束的缓解直接影响了企业流动资金,改善了企业财务困境,促使企业追加投资,进行扩大再生产,增加了企业劳动力需求,进而提高就业水平(毛其淋和王玥清, 2023)。基于上述分析,本文认为,大数据试验区设立通过新岗位创造效应、市场扩大效应及融资约束缓解效应影响就业水平。

三、实证设计

1. 数据来源

(1)企业数据。本文使用2010—2021年中国沪深A股上市公司的数据,其中,企业层级的数据主要来源于CSMAR、Wind和Choice数据库。为了消除异常样本对结果的影响并提高估计的精确度,本文对原始样本做如下处理:①剔除金融行业上市公司;②剔除ST、*ST类企业;③剔除资不抵债的企业;④仅保留正常上市的企业,即剔除暂停上市、终止上市和退市整理期的企业;⑤删除核心变量缺失的企业样本。为了降低异常数据对结果的影响,本文对连续变量进行了缩尾处理。

(2)城市数据。城市层面的数据主要来自相关年份《中国城市统计年鉴》,省份层面数据来自相关年份《中国统计年鉴》。

2. 模型设定

本文从稳就业视角考察大数据试验区设立对就业的影响。为了评估政策影响的净效应,本文借助双重差分模型对就业在大数据试验区这一试点政策实施前后的个体差异和时间差异进行控制。基于此,本文将位于大数据试验区试点城市的企业作为处理组,其他企业作为控制组,基准回归模型设定如下:

$$Labor_{it} = \alpha + \beta Bigdata_{jt} + \gamma X_{it} + \theta Z_j \times f(t) + \delta_i + \lambda_t + \varepsilon_{ijt} \quad (1)$$

其中, i, j, t 分别代表企业、城市、年份。因变量 $Labor_{it}$ 表示企业劳动力雇佣人数,以对数化的企业员工人数表征。解释变量 $Bigdata_{jt}$ 为大数据试验区设立的虚拟变量,当城市 j 在2015年或2016年为大数据试验区的试点城市取值1,否则为0。系数 β 能够衡量相较于非试点地区,试点城市的大数据试验区设立对企业雇佣人数的影响,若 β 大于0,则表明大数据试验区能够增加就业。本文在

回归中增加了企业级控制变量(X_{it})和事前(2010年)城市级控制变量与时间趋势交乘项($Z_{it} \times f(t)$),以排除大数据试验区政策之外其他影响就业的因素。本文还加入了企业固定效应(δ_i)和年份固定效应(λ_t),用以控制企业固有特征和时间趋势等混淆因素对估计结果产生的影响,在一定程度上缓解了遗漏变量偏误问题。 ε_{ijt} 为误差项,本文在基准回归中采用城市级聚类稳健标准误。

3. 变量构建

(1)被解释变量。本文的被解释变量为企业雇佣人数,依据陈胜蓝等(2023)、周亚虹等(2023),本文使用上市公司员工总人数的自然对数作为衡量企业劳动力雇佣的代理变量。

(2)核心解释变量。本文通过手动方式从中国政府网站收集并整理了2015—2016年被批准为国家级大数据综合试验区试点的地区名单。最终,本文得到八大国家级大数据综合试验区:2015年第一批试点名单仅包括贵州省^①;2016年选定的第二批试点区域包括京津冀地区(北京、天津和河北)、珠江三角洲(位于广东)、上海、河南、重庆、沈阳以及内蒙古。在珠三角地区推进大数据试验区建设的过程中,该区域的发展和影响力不仅限于地区中主要城市,同时也扩展至广东省内的其他城市。因此,遵从邱子迅和周亚虹(2021)、孙伟增等(2023)的做法,本文将位于广东省所有地级市的上市公司作为处理组。

(3)控制变量。①企业层面控制变量。参照既有研究,本文加入了资产负债率(Lev)、企业年龄(Age)、资产回报率(Roa)、固定资产投资($Capital$)(以固定资产净额对数表示)、经营性现金流($Cash$)(以经营性净现金流与资产总计之比表示)、企业成长能力($Growth$)(以营业收入增长率衡量)、管理层持股比例($Mshrrat$)、最大股东持股份额($Shrholder1$)、独立董事会成员比重($Share_indep$)。②采用事前城市层面控制变量(Hoynes et al., 2016),包括:地级市经济发展水平(Gdp),以地级市地区生产总值的对数衡量;城市人口(Pop),以城市平均人口的对数衡量;二产结构($Second$),以第二产业占地区生产总值的比重表示;三产结构($Third$),以第三产业与地区生产总值的比值表示;城市劳动力成本($Wage$),以职工平均工资的对数表示;教育水平(Edu),以普通高等学校在校学生人数占总人口的比重表示;工业发展基础($Firm_num$),以规模以上工业企业数量的对数表示。

四、实证结果及分析

1. 基准结果^②

表1展示了大数据试验区设立对企业劳动力雇佣影响的估计结果。在基准回归模型中,本文采用了逐步回归的方法。在第(1)列中,模型仅包含政策变量,并加入企业固定效应。结果显示,大数据试验区($Bigdata$)的系数为0.37,并且在1%的水平上显著,意味着大数据试验区的设立显著推动了当地就业人数的增长。第(2)列在第(1)列基础上加入时间固定效应,估计系数($Bigdata$)仍在

① 值得注意的是,尽管2016年2月,国家发展和改革委员会等部门才批复贵州建设全国首个国家大数据综合试验区,但2015年出台的《促进大数据发展行动纲要》明确提出推进贵州大数据综合试验区的发展,同年9月,贵州大数据综合试验区已经正式成立。因此,本文将贵州作为第一批试点地区。此外,鉴于贵州省试点时间不同于其他地区,在稳健性检验中,本文剔除了贵州省样本重新进行回归,结果参见《中国工业经济》网站(ciejournal.ajcass.com)附件。

② 本文考虑了双重差分(DID)模型不满足个体处理稳定性假设(SUTVA)条件,并进行了一系列稳健性检验,相关检验结果参见《中国工业经济》网站(ciejournal.ajcass.com)附件。

1%的水平上显著为正,但回归系数的数值有较大幅度下降,说明时间趋势对本文估计结果的干扰较大。第(3)、(4)列进一步加入企业层面控制变量、事前城市层面控制变量与时间趋势交乘项,不难发现,不论是否引入企业或事前城市级控制变量,*Bigdata*的系数始终在1%的水平上显著为正,并且估计系数的数值变化较小,表明大数据试验区政策具有较强的外生性。从经济显著性看,以第(4)列为例,相较控制组企业,大数据试验区设立地区的就业增加7.13%。当大数据试点地区的企业雇佣劳动人数均值为7552.02时,试点地区的企业雇佣人数将增加538人,相当于样本均值(5786.75)的9.30%。由此可见,大数据试验区试点的促就业效应在经济意义上同样显著。综上所述,国家级大数据综合试验区政策对提高就业水平发挥了积极作用。

表1 基准结果

	(1)	(2)	(3)	(4)
	<i>Labor</i>	<i>Labor</i>	<i>Labor</i>	<i>Labor</i>
<i>Bigdata</i>	0.3654*** (0.0186)	0.0938*** (0.0296)	0.0748*** (0.0276)	0.0713*** (0.0236)
企业级控制变量	否	否	是	是
$Z_j \times f(t)$	否	否	否	是
企业固定效应	是	是	是	是
年份固定效应	否	是	是	是
观测值	29343	29343	29343	29343
Adj.R ²	0.8653	0.8771	0.9129	0.9135

注: Z_j 表示事前城市层面控制变量, $f(t)$ 为时间趋势项;*,**,***分别表示10%、5%和1%的显著性水平;括号内为聚类到城市层面的稳健标准误。以下各表同。

2. 平行趋势检验^①

符合平行趋势假设是使用双重差分方法进行因果推断的关键。本文采用事件研究方法验证处理组与对照组之间是否符合平行趋势假设。由于反事实结果的不可观测性,平行趋势假设实际上无法检验,一种间接的做法是比较政策事前的估计系数。如果政策事前的估计系数接近零,且在统计上无法拒绝与零的显著差异,则表明处理组与控制组在事前遵循相同的变化趋势,即本文的双重差分模型估计有效。为了验证平行趋势的有效性,基于式(1),本文构建了一种更为灵活的动态回归形式:

$$Labor_{it} = \alpha + \sum_{k=-5}^{k=5} \beta_k D_k \times Treated_j + \gamma X_{it} + \theta Z_j \cdot f(t) + \delta_i + \lambda_t + \varepsilon_{ijt} \quad (2)$$

上式采用样本内各独立年份 D_k 与处理组虚拟变量($Treated_j$)的交互项替换双重差分项进行回归。 k 为第 t 年与大数据试验区设立的年份之差,表示大数据试验区设立的第 k 年,本文选取(-5, 5)的时间窗口。为了避免多重共线性问题,将-1期作为基期。因此, k 的取值变为-5, -4, -3, -2, 0, ..., 5。系数 β_k 表示,与-1期相比,大数据试验区设立第 k 年实验组和对照组企业雇佣劳动的差异。图1报告了平行趋势检验结果,不难发现,在 $k < 0$ 时, β_k 系数并不显著,意味着在大数据试验区

① 除此之外,本文还进行了以下稳健性检验,包括排除混淆政策干扰、置换检验、双重差分的异质性处理效应检验、倾向得分匹配与工具变量的内生性检验、控制城市和行业固定效应、更换聚类标准误、剔除特殊样本等。相关检验结果参见《中国工业经济》网站(ciejournal.ajcass.com)附件。

成立之前,处理组与对照组企业在就业水平上并未展现出显著的系统性差异。换言之,本文的双重差分估计符合平行趋势假设。而 $\beta_k(k \geq 0)$ 系数基本上显著为正,说明在大数据试验区设立之后,试点地区企业的就业水平显著高于非试点地区企业,且估计系数随着 k 的增加,呈现先上升后下降的趋势,这一显著促进作用在样本期间持续存在,说明大数据试验区的设立促进企业就业增加的效果具有长期动态效应。

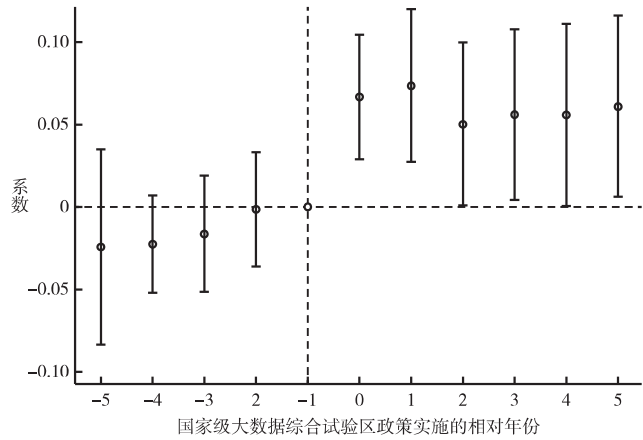


图1 平行趋势检验

五、机制检验

前文采用严谨的实证设计证实了大数据试验区设立的促就业效应。理论分析表明,大数据试验区通过新岗位创造效应、市场扩大效应以及融资约束缓解效应影响就业。本部分将从以上三个方面实证考察大数据试验区促进企业雇佣劳动增加的作用机制。

1. 新岗位创造效应

大数据试验区的设立有效推动了大数据技术的发展,提升了专业化分工效率与细化程度,从而催生了新工作任务、新职业以及新岗位(王林辉等,2023)。这一技术并非以替代劳动力为目标,而是在简化工作任务的同时,降低了工人执行任务所需的技能要求,并扩展了其掌握专业知识的相关性与广度(Autor, 2024)。此外,作为一种与劳动相互补充的工具,大数据技术通过改善劳动生产率来提升劳动力的边际生产力,从而增加企业对劳动力的整体需求(尹志锋等,2023)。本文从新注册企业视角剖析大数据试验区设立的新岗位创造效应。这一逻辑是,大数据试验区设立能够促进地区统一大数据信息平台的构建,消除“信息孤岛”,打通要素流通障碍,打破市场进入壁垒,极大降低创业风险,有利于新创企业成立。通常而言,新创企业成立能够带来较多的就业岗位(Hjort and Poulsen, 2019)。

本文从天眼查数据库获取了工商注册企业信息,手工整理出地级市层面每年新注册的企业数量,并取对数来衡量地区创业水平(*Newfirm*),结果呈现在表2第(1)列。可以发现,*Bigdata*的估计系数显著为正,表明大数据试验区设立能够提高地区创业活跃度,进而促进就业增长。其次,采用中国区域创新创业指数(IRIEC)中新注册企业得分作为城市创业活跃度的代理变量(*Newfirm_score*)^①,第(2)列结果显示,*Bigdata*的系数在1%的水平上显著为正,再次验证大数据试验区设立能够激发城市创业活力,提升了就业水平。

本文进一步采用城市层面新增上市公司的子公司数量(*New_affiliate*)来衡量上市公司新进入情况。具体而言,借鉴毛其淋和盛斌(2013)的做法,根据以下策略识别企业的进入:如果上市公司子公司*i*在第*t*期存在,而在第*t-1*期不存在,则将该公司*i*标记为第*t*期进入。相比上市公司本身而言,其子公司通常具备更高的市场进入与扩展弹性,是企业进行市场探索和业务扩展的主要手

① 中国区域创新创业指数(IRIEC)由北京大学企业大数据研究中心编制,该指数位于1—100之间,数值越大,说明城市创业活跃度越高。

表 2 新岗位创造效应

	(1)	(2)	(3)	(4)	(5)
	<i>Newfirm</i>	<i>Newfirm_score</i>	<i>New_affiliate</i>	<i>Bigdata_newfirm</i>	<i>Bigdata_emp</i>
<i>Bigdata</i>	0.0548* (0.0297)	1.4102*** (0.4721)	0.1762*** (0.0489)	0.0953*** (0.0332)	0.0760** (0.0336)
控制变量	是	是	是	是	是
企业固定效应	是	是	是	是	是
年份固定效应	是	是	是	是	是
观测值	3312	3036	3216	3312	2766
Adj.R ²	0.9615	0.9540	0.8410	0.9418	0.9477

段。从第(3)列可以发现,大数据试验区设立增加了城市层面上市公司新增子公司的数量,表明大数据试验区促进新企业进入的作用不仅体现在所有新注册企业中,也通过上市公司新增子公司得到了进一步验证。具体而言,上市公司通过设立子公司,在当地扩展了新业务,直接增加对各类岗位的需求,从而推动了就业的增长。此外,考虑到大数据相关行业的新注册企业和就业人数直接受大数据试验区设立的影响,本文还整理了信息传输计算机服务和软件业、科学研究技术服务业等与大数据相关行业的新创企业数量(*Bigdata_newfirm*)和就业人数(*Bigdata_emp*),并取对数进行回归。第(4)、(5)列结果表明,大数据试验区设立促进了大数据相关行业的新企业进入,由此引发新岗位创造效应,带动大数据相关行业的就业人数显著增加。^①以上分析证实了大数据试验区设立通过创造新岗位的直接作用影响企业雇佣水平,促进就业增长。

2. 市场扩大效应

首先,大数据试验区设立以大数据技术提升效率。大数据技术能够助力企业智能化改造和数字化转型,推动生产方式转型升级,使其生产技术更加合理、管理方式更加高效,进而促进生产效率提升(孙伟增和郭冬梅,2021);大数据有助于企业生产经营过程透明化,不仅能够提高管理效率,优化企业内部资源配置,还能搭建可视化的供应商—客户网络,降低外部信息不对称,提高企业经营效率(沈坤荣等,2024)。生产效率改善能够提高企业竞争力,扩大市场份额,进而促进就业增长。为验证大数据试验区设立对企业生产率的影响,本文测算了企业全要素生产率水平(*TFP*),并将其作为被解释变量代入基准回归方程。表3第(1)、(2)列报告了大数据试验区设立对生产率影响的回归结果,第(1)列中,*Bigdata*的系数在1%的水平上显著为正,即该试点政策实施提升了当地企业生产率;在第(2)列中同时加入政策变量与生产率指标,重新进行回归。结果表明,大数据试验区设立通过提升生产率,促进了就业水平提升。

其次,大数据试验区设立以大数据平台改善信息不对称。大数据试验区加快了一体化大数据平台的搭建,降低企业在设计、生产、研发、销售等环节中的信息不对称,及时掌握市场信息,扩大生产规模以抢占市场份额;另外,企业还能通过大数据平台向外界传递其高质量产品或服务的信号,有利于提高企业声誉价值,吸引消费者增加对产品或服务的需求,促使企业扩大生产规模。为检验市场规模渠道,参考尹志锋等(2023),本文使用企业所占市场份额(*Market_share*)衡量企业市场规模,并将其作为被解释变量加入基准回归。在第(3)列中,*Bigdata*的系数显著为正,说明大数据试

^① 本文还从新职业、常规岗位和非常规岗位视角提供了新岗位创造渠道的相关检验,参见《中国工业经济》网站(ciejournal.ajcass.com)附件。

表3 市场扩大效应

	(1)	(2)	(3)	(4)
	<i>TFP</i>	<i>Labor</i>	<i>Market_share</i>	<i>Labor</i>
<i>Bigdata</i>	0.0538*** (0.0194)	0.0512* (0.0262)	0.0020*** (0.0007)	0.0576*** (0.0221)
<i>TFP</i>		0.5119*** (0.0172)		
<i>Market_share</i>				6.9833*** (0.8085)
控制变量	是	是	是	是
企业固定效应	是	是	是	是
年份固定效应	是	是	是	是
观测值	24582	24582	29343	29343
Adj.R ²	0.9288	0.9328	0.8470	0.9153

验区设立提升了企业市场份额,扩大了生产规模。第(4)列同时加入政策变量与市场规模指标,结果表明大数据试验区设立通过提升市场规模、发挥市场扩大效应,增加了就业。

3. 融资约束缓解效应

大数据试验区设立以更有效的大数据技术缓解信贷约束。大数据平台不仅能帮助企业拓宽融资渠道,还能向外界披露企业生产经营状况,传递优质企业信号,有助于金融机构和潜在投资方及时了解和掌握企业经营信息(Balakrishnan et al., 2014; He et al., 2024)。为验证大数据试验区设立能否通过缓解企业融资约束进而促进企业雇佣劳动增加,本文使用WW指数衡量上市公司融资约束(*Fin_constraint*),并将其作为被解释变量代入基准回归。表4第(1)列结果表明,大数据试验区设立缓解了企业面临的融资约束,增加了试点地区的企业内部可用资金,促使企业扩大生产规模,从而增加对劳动力的雇佣。本文还使用企业从政府获取的补助总额(*Subsidy*)和创新补助(*Inno_subsidy*)作为衡量外部资金扶持的代理变量。第(2)、(3)列结果表明,大数据试验区中的企业确实得到了政府扶持,并获取了政府补助,企业面临的融资约束有所缓解。

进一步地,参考蔡庆丰等(2020),以上市公司周边的银行分支机构数量表征企业所面临的信贷资源可获得性。Degryse and Ongena(2005)指出,为了降低交易成本,银行通常更愿意为距离近的企业提供金融服务。本文获取了上市公司办公地点周边的银行网点数量^①,并根据事前银行网点数量的样本中位数,将样本分为信贷资源可获得性大、小两组进行回归。第(4)、(5)列显示,*Bigdata*的系数仅在信贷资源可获得性小的样本中显著为正,对信贷资源可获得性较大样本的影响不显著,表明大数据试验区设立提高了企业的信贷资源可获得性。最后,借鉴王贤彬和陈春秀(2023)的做法,本文利用企业贷款水平来验证信贷宽松优惠。一般而言,企业平均贷款水平越高,受到的融资约束越小,就越有可能获得信贷支持。本文使用企业流动负债(*Debt*)来验证信贷溢出,可以看出,大数据试验区设立确实能提高处理组企业融资水平,促进就业增长。综上所述,大数据试验区设立通过扩大市场和缓解融资约束的间接作用促进就业增长。

① 参照蔡庆丰等(2020)的做法,基于百度地图API的地理坐标匹配系统,根据企业办公地点和银行分支机构的经纬度计算二者之间的(5公里)直线距离。

表 4 融资约束缓解效应

	融资约束	政府补助	创新补助	信贷可得小	信贷可得大	信贷溢出
	(1)	(2)	(3)	(4)	(5)	(6)
	<i>Fin_constraint</i>	<i>Subsidy</i>	<i>Inno_subsi</i>	<i>Labor</i>	<i>Labor</i>	<i>Debt</i>
<i>Bigdata</i>	-0.0035** (0.0018)	0.1303** (0.0653)	0.3409** (0.1370)	0.0977*** (0.0355)	0.0582 (0.0359)	0.0534** (0.0240)
控制变量	是	是	是	是	是	是
企业固定效应	是	是	是	是	是	是
年份固定效应	是	是	是	是	是	是
观测值	24575	28508	7509	11361	12045	29343
Adj.R ²	0.7516	0.6959	0.3517	0.9020	0.9106	0.9490

六、进一步分析^①

1. 福利效应分析

在机制检验部分,本文研究表明大数据试验区设立通过新岗位创造、市场扩大以及融资约束缓解促进就业增长,而工资、福利水平和企业声誉也是吸引劳动力流入的重要因素(Jackman and Savouri, 1992;毛其淋和王玥清,2023)。本文进一步探讨大数据试验区设立对就业的影响是否还与员工薪酬福利和企业声誉等软实力有关。如果大数据试验区设立能够引导企业改善员工的薪酬福利、增强企业声誉,将进一步验证大数据试验区设立的就业促进效应是真实可靠的。表5第(1)一(3)列的被解释变量分别为工资总额(*Total_wage*)、五险一金(*Social_security*)、福利津贴总额(*Allowances*),可以看到,*Bigdata*的系数均显著为正,说明大数据试验区设立能够提升员工的工资总额、社会保障水平以及各种福利津贴,这对于求职者来说具有较强的吸引力,从而促进了更多劳动力流入。此外,本文还从企业声誉视角提供额外的证据。利用商道融绿 ESG(Environmental, Social, Governance)评级和彭博 ESG 指数衡量企业声誉,表5第(4)、(5)列结果显示,大数据试验区设立能够提升企业 ESG 表现,而 ESG 表现优异的企业能够通过向社会传递负责任的信号,打造高品质的企业形象(毛其淋和王玥清,2023),有利于吸引更多的求职者,进而促进就业增长。

表 5 福利效应

	(1)	(2)	(3)	(4)	(5)
	<i>Total_wage</i>	<i>Social_security</i>	<i>Allowances</i>	<i>ESG_SGF</i>	<i>ESG_Bloomberg</i>
<i>Bigdata</i>	0.0875* (0.0492)	0.1981** (0.0922)	0.1032*** (0.0355)	0.0377** (0.0166)	1.0612** (0.4196)
控制变量	是	是	是	是	是
企业固定效应	是	是	是	是	是
年份固定效应	是	是	是	是	是
观测值	28680	23709	13923	29343	9821
Adj.R ²	0.8619	0.7499	0.9108	0.5321	0.8137

2. 就业结构

前文的分析主要集中讨论大数据试验区设立对就业数量的影响,尚未深入探讨就业结构的变

^① 排他性检验结果参见《中国工业经济》网站(ciejournal.ajcass.com)附件。

化。就业是民生的基础,支持并规范新型就业形态的发展,强化灵活就业和新型就业形态劳动者的权益保障,有助于推动就业高质量发展。为了实现稳就业目标,不仅需要推动就业数量的增长,还需减轻结构性就业压力、实现劳动力资源的最优配置。鉴于此,本部分进一步分析大数据试验区设立对就业结构的影响,这对实现高质量充分就业具有现实指导意义。

首先,本文根据职业特征评估大数据试验区对就业结构的影响。国际社会经济地位指标(International Socioeconomic Index, ISEI)与职业社会声望的标准化指标(Standard International Occupational Prestige Scale, SIOPS)是常用来评价职业特征的相关指标(Ganzeboom et al., 1992)。其中,ISEI通过赋予每个职业平均收入和教育水平相应的权重,测量从事该职业个体的社会地位,该指标越大,表示职业的社会经济地位越高;SIOPS则是通过受访者对各种职业的受欢迎程度进行打分,并统计每个职业的好评比例,最终形成职业声望得分。本文统计了上市公司各职业岗位的ISEI和SIOPS指标,然后按照职业特征指标中位数将上市公司岗位类别划分为经济地位高(*High_status*)与经济地位低(*Low_status*)组、社会声望高(*High_kudos*)与社会声望低(*Low_kudos*)组,分别进行回归。

表6第(1)–(4)列表明,大数据试验区设立不仅显著促进了经济地位和社会声望高的岗位就业,还促进了经济地位和社会声望低的岗位就业。可能的原因在于:一是大数据试验区设立引致了传统职业向新职业转变,创造了大量诸如大数据分析师、大数据工程师等新岗位,这些新职业通常需要更高的技能水平和教育背景;同时,随着大数据和人工智能技术的发展,更多依赖人际互动、沟通、分析和决策的非常规工作任务不断涌现,而此类职业的经济地位和社会声望通常较高。二是当前大数据和人工智能等新一代信息技术仍处于技术发展的起步阶段,在技术应用层面主要表现为劳动辅助而非劳动替代(尹志锋等,2023)。在这一阶段,大数据技术不仅促进高技能岗位的增加,还有效地拓宽低技能工人的就业机会(孙伟增和郭冬梅,2021)。此外,新技术具有“去技能化”和“再技能化”的特征,能够降低职业技能的专业化程度,并拓展职业技能的适用范围(王林辉等,2023)。具体而言,大数据技术通过降低某些岗位的技能门槛,使得低技能劳动力能够进入以前较难涉及的岗位,尤其是一些支持性和辅助性岗位,如数据录入、客户支持等。这可能加速了职业变迁,推动更多低技能、低学历工人流入辅助性岗位,而这类职业的经济地位和社会声望通常较低。因此,大数据试验区设立促进了不同经济地位和社会声望职业的就业增长。

其次,本文还按照受教育水平将员工分为专科及以上学历(*Highedu_staff*)、高中及以下学历(*Lowedu_staff*)两组,在此基础上考察大数据试验区设立对不同学历结构员工就业的影响。第(5)、(6)列报告了基于员工学历结构的回归结果,可以发现*Bigdata*的系数均显著为正,表明大数据试验区设立提高了专科及以上学历、高中及以下(即各类教育背景)员工的就业水平。其背后的原因在于,一方面,大数据技术增加了新工作任务的多样性,从而催生了多种不同类型的工作任务。大数据技术的普及降低了员工执行某些工作任务的技能门槛,意味着低学历员工也可以参与到一些原本需要较高技能的岗位中(Autor, 2024),同时大数据技术作为互补性劳动工具成为改善劳动生产率的重要途径。例如,专科及以上学历的员工可能更多地从事数据分析、技术支持等涉及较高技能的岗位,高中及以下学历的员工则可能从事支持性、操作性等相对基础的任务。多样化的岗位需求促使大数据试验区同时吸纳高学历和低学历员工。另一方面,大数据试验区设立推动了技术进步,通过提高生产效率和扩大市场规模,发挥市场扩大效应,不仅促进了高附加值岗位的增加,还提供了大量支持性、基础性岗位。因此,大数据试验区设立能够同时显著提高专科及以上学历员工的就业水平。

表 6 就业结构

	职业特征				学历特征	
	(1)	(2)	(3)	(4)	(5)	(6)
	<i>High_status</i>	<i>Low_status</i>	<i>High_kudos</i>	<i>Low_kudos</i>	<i>Highedu_staff</i>	<i>Lowedu_staff</i>
<i>Bigdata</i>	0.0630** (0.0266)	0.0622* (0.0321)	0.0609** (0.0277)	0.0651** (0.0313)	0.0829*** (0.0283)	0.1015** (0.0450)
控制变量	是	是	是	是	是	是
企业固定效应	是	是	是	是	是	是
年份固定效应	是	是	是	是	是	是
观测值	21536	24730	21747	23789	21636	24683
Adj.R ²	0.9115	0.8789	0.9055	0.8862	0.9023	0.9124

3. 地区就业

劳动力具有动态流动和再平衡的特征,就本文而言,尽管大数据试验区设立促进企业增加对劳动力的雇佣,但是由于上市公司占地区所有企业的比重较小,从地区整体就业看,大数据试验区设立能否提升城市层面的总体就业水平有待进一步探究。本文进一步将视角转向城市层面,考察大数据试验区设立对地区劳动力市场的影响。^①结果表明,大数据试验区设立显著促进了地区层面就业增长,且与大数据相关行业中,这一促进作用更为明显。具体地,在政策执行后,试点城市的就业增长 4.59%。至此,本文发现大数据试验区设立不仅在企业层面发挥促就业效应,还能增加地区层面劳动力雇佣,这对于国家实现稳就业、保民生目标具有重要意义。

4. 政策优化

在详细探究大数据试验区设立对就业水平的影响及作用机制后,本文将进一步从政策制定者角度来考察这一政策实施的优化空间。大数据试验区设立能够增加企业和地区的劳动力雇佣水平,但与此同时,政府也对大数据试验区进行了补贴。因此,政策制定者可能更加关心的问题是应该在哪些城市建立大数据试验区,以最大化政策效果,从而使得促进就业水平提升的效应最大。为回答这一问题,本文基于 Munro et al.(2024)提出的市场均衡效应处理框架,使用广义随机森林算法(Athey et al., 2019),计算处理组和对照组样本的个体处理效应,并在一定约束条件下划定政策执行的最优规则,比较最优规则 and 实际执行情况。本文定义样本中某个体自身受到政策的影响效应如下:

$$\tau_{CADE}(x) = E[Y_{in}(W_i = 1; W_{-i}) - Y_{in}(W_i = 0; W_{-i}) | X_i = x] \quad (3)$$

其中, W 为一个 n 维向量,其第 i 个元素为 W_i 。用 X_i 表示个体 i 的相关特征。 Y_i 为第 i 个个体的潜在结果函数,依赖于所有个体的处理情况,记为 $Y_i(W)$ 。将样本中个体之间由于相互影响或策略调整使得自身受到其他个体的间接影响定义为条件平均间接处理效应($\tau_{CAIE}(x)$),如下所示:

$$\tau_{CAIE}(x) = (n - 1)E[Y_{jm}(W_i = 1; W_{-i}) - Y_{jm}(W_i = 0; W_{-i}) | X_i = x](i \neq j) \quad (4)$$

在本文中,条件平均间接处理效应产生的原因包括:处理组样本可能通过服务外包给对照组样本,从而增加对照组样本就业;处理组样本因政府补贴经济可能快速增长,资本和劳动力可能从对照组向处理组样本流动,增加处理组样本就业等。最优化问题可进一步表示为:

^① 有关地区层面的回归模型设定及结果参见《中国工业经济》网站(iejournal.ajcass.com)附件。

$$\operatorname{argmax} E \left\{ \pi(x_i) E [Y_i(1, P_q^*) - Y_i(0, P_q^*) | X_i = x] \right\} \quad (5)$$

$$\text{s. t.} \quad E \left\{ \pi(x_i) E [Z_i(1, P_q^*) - Z_i(0, P_q^*) | X_i = x] \right\} = b \quad (6)$$

其中, $\pi_i = \pi(x)$ 表示某个体实际受到政策处理的概率, $b = qE [Z_i(1, P_q^*) - Z_i(0, P_q^*)]$ 。当且仅当 $\pi(x)$ 具有以下形式时, $\pi(x)$ 既符合式(6)的约束条件, 又是最优化问题式(5)的解。

$$\pi(x) = \begin{cases} 1, & \text{如果 } \tau_{CADE}^Y(x) > C^T \tau_{CADE}^Z(x) \\ \alpha, & \text{如果 } \tau_{CADE}^Y(x) = C^T \tau_{CADE}^Z(x) \\ 0, & \text{其他} \end{cases} \quad (7)$$

其中, C 是一个常数向量, α 是一个常数且 $0 < \alpha < 1$ 。

根据前文对政策背景的梳理, 在 2015 年 9 月和 2016 年 10 月两批共建立八大综合试验区, 本文以政策实施后 1 年 (即 2017 年) 为例, 对政策优化结果进行说明。在对上述最优化问题进行求解后, 本文将政策实施最优规则以及处理组和对照组样本展示在图 2 中。其中, 虚线为政策最优实施规则, 虚线以上的散点为应该纳入政策实施的样本, 以下的散点为不应纳入政策实施的样本。

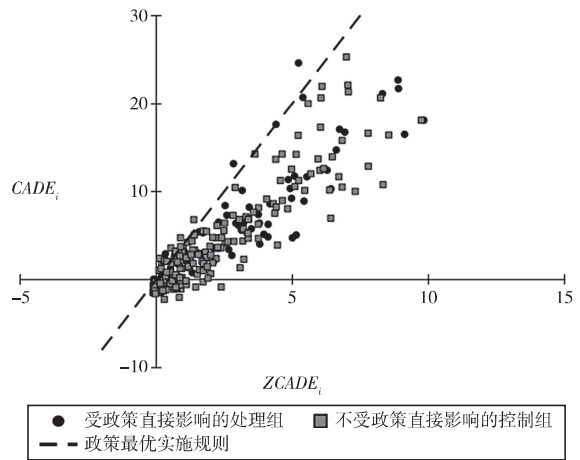


图 2 2017 年各城市的就业水平 (CADE) 和政府补贴 (ZCADE) 估计值的散点

表 7 呈现了政策试点与政策最优执行规则下的城市数量, 由于数据缺失, 2017 年共有城市样本 236 个, 其中, 有 58 个城市属于大数据试验区建设试点。在本文通过机器学习计算出的政策最优执行规则下, 只有 5 个城市属于大数据试验区建设试点, 其他 14 个城市尚未真实进行大数据试验区建设。在真实进行大数据试验区建设的 58 个城市中, 有 53 个城市不满足政策最优执行规则。由此可见, 大数据试验区政策的优化空间较大, 这为未来政策优化方向提供了重要参考。

表 7 政策试点与政策最优执行规则下的城市数量 单位: 个

属于政策试点的城市数量	符合政策最优执行规则的城市数量		不符合政策最优执行规则的城市数量	
	属于政策试点	不属于政策试点	属于政策试点	不属于政策试点
58	5	14	53	164

七、结论与启示

在全球经济增长放缓和外部环境不确定性加剧的背景下, 中国面临严峻的就业形势。劳动力充分就业既是经济增长的重要前提, 也是高质量发展的关键所在。企业作为推动创新、促进就业和改善民生的主体, 如何提升企业劳动力雇佣水平对于稳就业、保民生至关重要。推进大数据试验区

建设是优化企业外部环境的一项重要举措,旨在构建完善的大数据基础设施并整合区域数据资源,其能否以及如何影响就业,成为一个尚待研究的重要课题。

为了回答上述问题,本文采用2010—2021年中国上市公司数据,实证检验了大数据试验区设立对就业的影响及作用机制。研究发现,大数据试验区促进了企业雇佣员工人数增加,相比控制组企业,大数据试验区设立地区的就业增加7.13%,这一结论在一系列稳健性检验后仍然成立。机制检验表明,大数据试验区设立通过新岗位创造效应、市场扩大效应与融资约束缓解效应,促进企业劳动力雇佣规模增加。拓展性分析显示,大数据试验区设立不仅为企业员工提供丰厚报酬和社会保障,还提升了企业声誉,进而吸引劳动力流入。本文还考察了大数据试验区设立对就业结构的影响,基于职业特征视角,发现大数据试验区设立对不同经济地位和社会声望的职业岗位均存在显著促进作用;基于学历结构角度,发现大数据试验区设立对各类教育背景的员工就业均有促进作用。在地区层面,本文同样发现大数据试验区设立促进了就业增长,尤其是与大数据相关的行业中,这一促进效应更为显著。最后,基于市场均衡效应分析框架,采用机器学习方法比较了政策最优执行规则 and 实际执行情况,结果表明大数据试验区政策优化空间较大。结合研究结论,本文提出如下政策启示:

(1)扩展大数据技术覆盖面,促进就业增长。大数据技术是促进企业和地区就业吸纳的重要推动力。政府应在符合政策最佳执行条件的地区逐步增加大数据试验区的试点城市,优化政策实施条件,以释放政策红利,推动地区数字基础设施建设和大数据平台发展;应进一步利用健全的数字基础设施来改善地区的数字环境,吸引更多企业融入大数据网络,充分发挥数据要素的增长潜力;促进大数据技术在各行业的广泛应用,创造更多与大数据相关的就业机会。通过多层次的政策扶持,加速大数据在地方经济中的渗透,促进不同层次的劳动力群体进入数字经济领域,减少技术进步带来的就业结构性不平衡问题,从而更广泛地推动就业增长。

(2)加大对大数据基础设施的投资力度,提升就业吸纳能力。完善的大数据基础设施是促进就业增长的核心要素。由此,应加强基础设施的建设与维护,确保数据要素的流通与共享,最大化其在“稳就业”中发挥的积极作用;继续加大对数据中心、宽带网络等基础设施的投入,以提升其在技术和规模上的领先地位,缩小地区间数字经济发展的差距,进而提高劳动力市场的流动性和包容性。此外,应推动建立企业间的数据共享平台,助力企业更好嵌入大数据网络,支持更多企业通过大数据技术提升生产效率和管理能力,扩大企业招聘规模和提供更多就业机会。

(3)整合大数据平台资源,提高企业信贷融资及就业水平。大数据平台在缓解企业融资约束方面具有独特的优势,能够为企业提供信贷评估和风险管理,尤其是在缓解中小企业融资约束方面具有显著优势。为此,应推动金融机构与大数据平台合作,为企业提供更为精准的融资支持,尤其是为那些面临资金困境的中小企业提供信贷渠道。具体地,应通过建立信用信息共享机制和数据交换平台,降低金融机构的信息成本,提高信贷服务的覆盖面和精准性,使更多中小企业从大数据平台中获益,进而吸纳更多劳动力就业。

(4)推动大数据技术普及,加强劳动者技能与岗位匹配度。大数据技术的广泛应用不仅有助于促进新企业成立,还创造了大量新岗位,同时大数据对就业的促进作用仅体现在非常规工作岗位中,对常规岗位的作用并不明显。为了让更多劳动力从技术变革中受益,应加强劳动力技能培训,尤其是对低技能和低学历人群的培训。通过制定针对性的培训项目,提供大数据技术、数据分析等方面的技能提升课程,帮助劳动力掌握必要的技能,满足新兴岗位(非常规岗位)的需求。此外,政府可以通过补贴政策 and 培训激励,鼓励企业参与员工技能提升,减少技能不匹配,从而促进就业机

会的平等化。

本文研究仍有待深入探索之处:一是研究范围的扩大可以增强政策启示的普适性。当前研究主要聚焦于上市公司与大数据技术的应用,尚未对中小企业和非上市公司在大数据应用中的表现进行深入探讨。未来研究可以根据企业的类型进行差异化分析,进一步揭示大数据技术对不同类型企业就业吸纳的影响。二是加强对技术应用层面的深度探讨。当前的研究缺少对大数据如何改变劳动市场中的工作性质、工人角色以及职业结构的系统分析。未来,可以通过技术层面的深入分析,探索大数据技术如何重新定义劳动者技能要求、岗位属性以及工作流程。三是进一步细化政策设计也是未来研究的核心议题。通过更多的实地数据和多维度的政策评估,能够更精准地识别大数据试验区设立在不同区域和行业的政策效应,进而为政策制定者提供更加具体的建议。

〔参考文献〕

- [1]蔡庆丰,陈熠辉,林焜.信贷资源可得性与企业创新:激励还是抑制?——基于银行网点数据和金融地理结构的微观证据[J].经济研究,2020,(10):124-140.
- [2]陈琳,高悦蓬,余林徽.人工智能如何改变企业对劳动力的需求?——来自招聘平台大数据的分析[J].管理世界,2024,(6):74-93.
- [3]陈胜蓝,王鹏程,马慧,刘晓玲.《中小企业促进法》的稳就业效应——基于政府信用体系建设视角[J].管理世界,2023,(9):52-68.
- [4]崔小勇,蔡响珊,卢国军.增值税留抵退税能否促进企业吸纳就业?——来自2019年试行留抵退税制度的证据[J].管理世界,2023,(9):15-38.
- [5]方锦程,刘颖,高昊宇,董纪昌,吕本富.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.
- [6]刘涛雄,戎珂,张亚迪.数据资本估算及对中国经济增长的贡献——基于数据价值链的视角[J].中国社会科学,2023,(10):44-64.
- [7]毛其淋,盛斌.中国制造业企业的进入退出与生产率动态演化[J].经济研究,2013,(4):16-29.
- [8]毛其淋,王玥清.ESG的就业效应研究:来自中国上市公司的证据[J].经济研究,2023,(7):86-103.
- [9]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,(7):4-17.
- [10]沈坤荣,乔刚,林剑威.智能制造政策与中国企业高质量发展[J].数量经济技术经济研究,2024,(2):5-25.
- [11]孙伟增,郭冬梅.信息基础设施建设对企业劳动力需求的影响:需求规模、结构变化及影响路径[J].中国工业经济,2021,(11):78-96.
- [12]孙伟增,毛宁,兰峰,王立.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.
- [13]田鹤,张勋.数字经济、非农就业与社会分工[J].管理世界,2022,(5):72-84.
- [14]王林辉,钱圆圆,周慧琳,董直庆.人工智能技术冲击和中国职业变迁方向[J].管理世界,2023,(11):74-95.
- [15]王贤彬,陈春秀.重点产业政策与制造业就业[J].经济研究,2023,(10):34-54.
- [16]王永钦,董雯.机器人的兴起如何影响中国劳动力市场?——来自制造业上市公司的证据[J].经济研究,2020,(10):159-175.
- [17]徐翔,赵墨非,李涛,李帅臻.数据要素与企业创新:基于研发竞争的视角[J].经济研究,2023,(2):39-56.
- [18]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,(4):103-119.
- [19]尹志锋,曹爱家,郭家宝,郭冬梅.基于专利数据的人工智能就业效应研究——来自中关村企业的微观证据[J].

- 中国工业经济, 2023, (5):137-154.
- [20]周亚虹, 杨岚, 姜帅帅. 约束性碳减排与就业——基于企业和地区劳动力变化的考察[J]. 经济研究, 2023, (7): 104-120.
- [21]Acemoglu, D., and P. Restrepo. Robots and Jobs: Evidence from US Labor Markets[J]. *Journal of Political Economy*, 2020, 128(6):2188-2244.
- [22]Acemoglu, D., D. Autor, J. Hazell, and P. Restrepo. Artificial Intelligence and Jobs: Evidence from Online Vacancies[J]. *Journal of Labor Economics*, 2022, 40(S1):293-340.
- [23]Acemoglu, D., and S. Johnson. Learning from Ricardo and Thompson: Machinery and Labor in the Early Industrial Revolution and in the Age of Artificial Intelligence[J]. *Annual Review of Economics*, 2024, 16(1):597-621.
- [24]Adachi, D., D. Kawaguchi, and Y. U. Saito. Robots and Employment: Evidence from Japan, 1978—2017[J]. *Journal of Labor Economics*, 2024, 42(2):591-634.
- [25]Athey, S., J. Tibshirani, and S. Wager. Generalized Random Forests[J]. *Annals of Statistics*, 2019, 47(2):1148-1178.
- [26]Autor, D. H., F. Levy, and R. J. Murnane. The Skill Content of Recent Technological Change: An Empirical Exploration[J]. *Quarterly Journal of Economics*, 2003, 118(4):1279-1333.
- [27]Autor, D. H. Why Are There Still So Many Jobs? The History and Future of Workplace Automation[J]. *Journal of Economic Perspectives*, 2015, 29(3):3-30.
- [28]Autor, D. H. Applying AI to Rebuild Middle Class Jobs[R]. NBER Working Paper, 2024.
- [29]Balakrishnan, K., M. B. Billings, B. Kelly, and A. Ljungqvist. Shaping Liquidity: On the Causal Effects of Voluntary Disclosure[J]. *Journal of Finance*, 2014, 69(5):2237-2278.
- [30]Cirillo, V., R. Evangelista, D. Guarascio, and M. Sostero. Digitalization, Routineness and Employment: An Exploration on Italian Task-Based Data[J]. *Research Policy*, <https://doi.org/10.1016/j.respol.2020.104079>, 2021.
- [31]Degryse, H., and S. Ongena. Distance, Lending Relationships, and Competition[J]. *Journal of Finance*, 2005, 60(1):231-266.
- [32]Ganzeboom, H. B. G., P. M. De Graaf, and D. J. Treiman. A Standard International Socio-Economic Index of Occupational Status[J]. *Social Science Research*, 1992, 21(1):1-56.
- [33]Goldfarb, A., and C. Tucker. Digital Economics[J]. *Journal of Economic Literature*, 2019, 57(1):3-43.
- [34]He, J., X. Du, and W. Tu. Can Corporate Digital Transformation Alleviate Financing Constraints[J]. *Applied Economics*, 2024, 56(20):2434-2450.
- [35]Hjort, J., and J. Poulsen. The Arrival of Fast Internet and Employment in Africa[J]. *American Economic Review*, 2019, 109(3):1032-1079.
- [36]Hoynes, H., D. W. Schanzenbach, and D. Almond. Long-Run Impacts of Childhood Access to the Safety Net[J]. *American Economic Review*, 2016, 106(4):903-934.
- [37]Hui, X., O. Reshef, and L. Zhou. The Short-Term Effects of Generative Artificial Intelligence on Employment: Evidence from an Online Labor Market[J]. *Organization Science*, 2024, 35(6):1977-1989.
- [38]Jackman, R., and S. Savouri. Regional Migration in Britain: An Analysis of Gross Flows Using NHS Central Register Data[J]. *Economic Journal*, 1992, 102(415):1433-1450.
- [39]Munro, E., K. Xu, and S. Wager. Treatment Effects in Market Equilibrium[R]. arXiv Working Paper, 2024.
- [40]Myers, S. C., and N. S. Majluf. Corporate Financing and Investment Decisions When Firms Have Information that Investors Do Not Have[J]. *Journal of Financial Economics*, 1984, 13(2):187-221.
- [41]Prantl, S. The Impact of Firm Entry Regulation on Long-Living Entrants[J]. *Small Business Economics*, 2012, 39: 61-76.

The Construction of National-Level Comprehensive Big Data Pilot Zone and Employment Growth

SHEN Kun-rong¹, QIAO Gang¹, TAN Rui-peng²

(1. The School of Business, Nanjing University;

2. School of Economics, Hefei University of Technology)

Abstract: Employment is a fundamental aspect of people's livelihoods, impacting individual well-being and the healthy development of the economy, with significant implications for national stability. Currently, the complex domestic and international environment, along with unforeseen shocks, presents severe challenges for businesses and worsens the employment outlook. As the digital and real economies become increasingly integrated, the growing digital sector has demonstrated its positive effects on economic growth and employment stability. While existing literature explores the employment impact of emerging technologies, consensus remains lacking. Amid economic downturns and rising employment pressures, it remains urgent to understand how big data technology can unlock data's value and stabilize employment.

This study investigates the impact and mechanisms of national-level big data pilot zones on corporate employment, using panel data from Chinese listed companies from 2010 to 2021, and employing a Difference-in-Differences (DID) approach. The results show that big data pilot zones significantly enhance corporate employment levels, a finding confirmed at the city level. Mechanism analysis reveals that these zones promote employment growth through new job creation, market expansion, and easing financing constraints. Further analysis indicates that they improve employee welfare and corporate reputation, attracting more labor. Employment structure analysis demonstrates that big data pilot zones improve employment across a variety of economic and social status occupations and boost employment for individuals with diverse educational backgrounds. Lastly, using machine learning techniques, the study compares optimal policy execution with actual outcomes, highlighting the significant potential to improve big data policies' effectiveness in promoting employment.

This study emphasizes the importance of fully leveraging the employment-promoting effects of big data pilot zones. To achieve this, it is essential to expand the coverage of big data technology, improve infrastructure, integrate platform resources, and enhance corporate financing capabilities. Additionally, strengthening workforce skill training is critical to improving the alignment between skills and job requirements. The study contributes by addressing a gap in the literature, which has largely overlooked the micro-level perspective of enterprises. Using a DID approach, this research directly quantifies the impact of data as a production factor on employment, thereby enriching and expanding the understanding of the economic effects of big data pilot zones from a microeconomic perspective. Moreover, this study systematically examines the potential pathways through which big data pilot zones affect employment, using comprehensive microdata. This approach provides new insights into the underlying mechanisms of how data resources promote employment. Lastly, by applying machine learning methods to assess policy implementation, the study identifies opportunities for further optimizing these policies.

Keywords: digital economy; national-level big data comprehensive pilot zone; employment growth

JEL Classification: E24 J21 M51

[责任编辑:李鹏]