

国家级大数据综合试验区的就业效应研究

王智勇 杨金娇*

内容提要 就业是民生之本。在当前经济增长放缓形势下，稳就业具有重要现实意义。随着数字经济的发展，如何依托大数据发展，为企业扩大就业规模赋能，是实现稳就业目标的重要议题。本文将国家级大数据综合试验区的设立视为一项准自然实验，基于2009—2021年沪深A股上市公司的面板数据，采用双重差分法探讨了大数据试验区建设的就业效应。研究发现，大数据试验区建设显著增加了位于试点城市企业的劳动力需求，在进行一系列稳健性检验后，结论仍然成立。机制分析表明，大数据试验区建设主要通过规模扩大效应和产品创新效应，增加企业的劳动力需求。异质性分析显示，大数据试验区建设产生的就业扩张效应在国有企业、劳动密集型行业、资本密集型行业以及数据开放水平高的城市中更加明显。研究结论丰富了大数据综合试验区的政策评估研究，为推动大数据技术运用和稳定居民就业提供了政策启示。

关键词 大数据综合试验区 就业 数字经济

一 引言

作为民生之本，就业是社会稳定的重要基石，是经济繁荣的核心支柱，是民生福祉的最大支撑。在新冠疫情疤痕效应持续、全球经济增长乏力、地缘政治错综复杂的背景下，国内就业市场遭到严重冲击，就业形势不容乐观（毛其淋、王玥清，2023）。

* 王智勇，中国社会科学院人口与劳动经济研究所，电子邮箱：wangzy_iwep@cass.org.cn；杨金娇，中国社会科学院大学应用经济学院，电子邮箱：yangjinjiao1004@qq.com。作者感谢国家自然科学基金应急管理项目（项目批准号：71641034）、中国社会科学院创新工程项目（项目编号：RKSCX2017014）的资助。

党的二十大报告提出要“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“稳就业”目标之下，亟须寻找新动能以拓宽就业渠道、扩大就业规模、改善就业质量。2023 年中央经济工作会议强调，“要以科技创新推动产业创新，特别是以颠覆性技术和前沿技术催生新产业、新模式、新动能，发展新质生产力”，同时“更加突出就业优先导向，确保重点群体就业稳定”。

在数字经济蓬勃发展的浪潮中，数字技术正在加速融入经济社会各领域，中国迎来了新一轮技术变革和产业重构，逐渐迈入数据驱动的经济发展新阶段。数据作为数字经济时代的关键生产要素（杨俊等，2022），不仅承载着海量的信息，还能创造出巨大的价值。大数据在各行各业的运用重塑了现有生产方式和消费模式，为就业市场创造了新的机遇和需求（王春超、聂雅丰，2023）。一方面，数据要素的深度融合创造了大量新行业和新岗位，催生出新型就业模式和就业形态（李成明等，2023），产生就业创造效应。另一方面，数字技术的广泛运用提高了生产效率，降低了生产成本，促进企业生产规模扩张，从而增加劳动力需求，产生就业补偿效应（王春超、聂雅丰，2023）。

数据要素的低成本特征（Goldfarb & Tucker，2019）和非竞争属性（Jones & Tonetti，2020），将会对各行各业的发展方向、经营理念、生产方式、营销手段等方面带来系统性变革和革命性冲击，造成就业岗位和工作内容的更新迭代。此外，数字技术改造传统产业的过程中，机器人和人工智能技术的应用日益广泛，会对相关岗位造成冲击，导致岗位减少。例如，许多制造业工厂中工业机器人逐渐取代了产业工人；网络交易平台和支付手段的发展，使得银行的柜员岗位大量削减。那么，一个亟须回答的问题是，数字技术的发展到底对就业造成了何种影响？

近年来，随着物联网、云计算、区块链等信息技术蓬勃发展，海量数据的管理储存、高速传输和开放共享成为可能。全球主要经济体纷纷出台政策支持大数据应用以抢占战略高地，推动数字经济发展。美国于 2012 年发布《大数据研究与发展计划》，澳大利亚在 2013 年出台《公共服务大数据战略》，欧盟在 2020 年制定了《欧盟数据发展战略》。中国政府同样重视推动大数据发展，自从在 2014 年的《政府工作报告》中首次提及大数据以来，2015 年印发的《促进大数据发展行动纲要》鼓励大数据和传统产业深度融合。为了进一步总结可在全国推广复制的大数据产业发展经验，国家发展改革委、工信部和网信办于 2016 年 2 月批复贵州省开展全国首个国家大数据综合试验区建设；同年 10 月，批复京津冀、珠三角、上海市、河南省、重庆市、沈阳市、内蒙古自治区继续设立国家级大数据综合试验区（下称大数据试验区）。其中，贵州省运用

大数据推进了产业转型升级，推动了工业转型发展，取得了相当显著的成效（陈加友，2017）。

相应地，中国持续多年的大数据试验区试点在释放数据要素潜能的同时对企业就业产生了何种影响？其内在作用机制是什么？大数据试验区建设对就业产生的作用，在不同特征的企业、行业和城市中是否存在明显差异？科学回答上述问题，对于稳定社会就业形势、培育企业新增长点、促进经济高质量发展具有重要意义。

本文基于 2009 - 2021 年沪深 A 股上市公司的面板数据，利用双重差分法，考察了国家级大数据综合试验区建设对企业就业规模的作用效果、影响机制和异质性。本文可能的边际贡献有三方面。第一，从研究视角来看，本文较早从企业层面评估了国家级大数据综合试验区建设的就业效应，为全面认识大数据试验区试点政策的实施效果提供了证据，同时丰富了企业就业的相关研究。第二，从研究方法来看，将国家级大数据综合试验区的设立视为准自然实验，采用双重差分法评估大数据试验区建设对就业规模的净效应，能够缓解识别过程中的内生性问题。第三，从现实意义来看，基于企业微观特征、行业中观特征和区域宏观特征进一步区分企业所有制、行业要素密集度和城市数据开放水平的差异，全面分析大数据试验区试点政策对企业就业规模影响的异质性，可以为政府精准施策推动大数据试验区建设、促进“稳就业”目标实现提供理论依据。

二 文献综述

已有相关文献包括两个部分，一是围绕国家级大数据综合试验区试点的政策效果评估，二是针对数字经济对就业的影响研究。本部分对两方面的研究进展进行评述。

（一）大数据试验区试点政策评估的研究进展

自 2016 年国家级大数据综合试验区批复以来，诸多学者评估了该政策所产生的影响，现有文献主要集中于评估大数据试验区政策的经济效应和环境效应。对经济效应的评估聚焦于全要素生产率、创新水平和企业转型等方面。大数据综合试验区能够显著促进城市创新水平（张慧等，2023），增加外商直接投资的规模并提高其质量（耿伟等，2023；Brynjolfsson & McElheran，2016；Corcoran & Gillanders，2015），还通过促增外商直接投资和改善营商环境增强城市创业活跃度（支宇鹏、卢潇潇，2023），促进了要素流动，减少了要素错配，从而提高了资源配置效率（蔡跃洲、马文君，2021），对提高企业创新绩效、全要素生产率和绿色全要素生产率都具有积极作用（侯世英、宋

良荣, 2021; 黄漫宇、王孝行, 2022; 李言、朱棋娜, 2023; 刘军等, 2024; 邱子迅、周亚虹, 2021; 张益豪、郭晓辉, 2023)。设立大数据试验区将通过改善信息不对称、促进要素流动等方式提升数字技术创新水平(戴艳娟等, 2023), 推动企业进行数字化转型, 促使制造业企业“脱虚向实”(孙哲远, 2022)。

对环境效应的评估主要集中在绿色技术创新、空气质量和碳排放等方面。陈文和常琦(2022)、任英华等(2023)揭示了大数据试验区建设对企业绿色创新的影响, 大数据成为推动经济发展绿色转型的“新蓝海”。郭炳南等(2022)在环境库兹涅茨理论的框架下, 验证了以数据要素为核心的数字经济, 通过推动产业升级、促进技术创新以及优化资源配置使空气质量得到有效改善。常皓亮等(2023)聚焦电力消费产生的碳排放, 发现大数据试验区建设通过促进技术创新和数字普惠金融发展, 降低了碳排放水平。

(二) 数字经济影响就业的研究进展

越来越多的文献关注到数字经济带来的经济影响和社会影响, 其中部分文献关注到数字经济对就业的作用, 但现有文献并没有得出一致的结论。有些文献认为数字经济能够推动就业扩容。Ndubuisi et al. (2021) 发现发展中国家的服务业繁荣与数字经济崛起有密切联系, 数字基础设施对服务业就业有正向促进作用。孙伟增和郭冬梅(2021) 得出类似结论, 信息基础设施建设通过企业规模扩张和生产效率提升对企业劳动力需求总量产生积极贡献。还有文献从居民的就业行为出发展开研究, 如陈贵富等(2022) 认为数字经济有利于提升就业供给的配置效率, 提升居民就业水平。周晓光和肖宇(2023) 则以中国 30 个省为研究对象, 论证了数字经济发展有利于提升居民就业率。

但有些研究认为数字经济导致就业规模缩减。Autor et al. (2006)、Autor & Dorn (2013) 较早论证了信息技术的应用会替代重复性、常规性工作岗位, 导致此类就业规模缩减。David (2017) 进一步评估了计算机技术对日本就业的破坏风险, 估计出大约 55% 的工作将会被替代。Arntz et al. (2016) 的研究显示数字化对就业的影响没有那么消极, 他们以经济合作与发展组织国家作为研究对象, 测算出约有 9% 的工作岗位被替代。余明桂等(2022) 研究了银行的数字化转型, 发现数字化技术的应用会对银行就业产生破坏效应, 原因在于数字化应用大幅缩减了银行的分行和网点, 降低了就业需求。

Chatterjee et al. (2022) 发现大数据能够推动创新并维持中小企业的供应链稳定运作, 大数据还通过加速人力资本积累和推动绿色技术创新促进城市高质量发展 (Guo et al.,

2023)。此外，还有研究认为数字经济发展过程中可能存在“数字鸿沟”，导致城乡收入差距扩大（陈文、吴赢，2021），阻碍部分企业创新（董香书等，2022）。数字经济的发展及机器人技术的使用并未带来生产率的提高，甚至降低了全要素生产率（Shackleton，2018）。数字经济带来的过度信息化导致劳动力错配与资源浪费，抑制了全要素生产率增长（Acemoglu & Restrepo，2019）。

另有少量文献得出其他结论。如孟祺（2021）基于中国省际面板数据，发现数字经济对整体就业规模的影响并不显著，但是分行业结果显示，数字经济会降低制造业就业规模，提升服务业就业规模。袁冬梅等（2023）聚焦服务业，发现数字经济对就业的影响呈现出先促进后降低的“倒U形”效果。

综上所述，学术界关于大数据试验区试点政策的效果评估主要集中于经济效应和环境效应，缺乏对该试点政策社会效应的关注，尤其是缺少对就业效应的探讨，但研究这一问题对中国的社会稳定、经济增长和可持续发展具有重要意义。而关于数字经济对就业影响的研究，由于研究视角和研究内容差异，以及不同国家发展阶段的差异，现有文献并未达成共识，因此有必要就这一问题继续展开深入研究。

三 理论分析与研究假说

（一）大数据试验区建设对就业的直接影响

大数据试验区建设至少可以从两个方面直接促进就业规模增长。一是大数据试验区建设有助于带动传统就业增长。大数据试验区建设的任务之一是统筹规划大数据基础设施建设，与之相关的企业将抢占先机，带动就业增长（孙伟增、郭冬梅，2021）。一方面，大量的工程师、技术人员和运维人员需投身到大数据基础设施的建设、维护和管理中；另一方面，电子设备和通信行业、软件开发和信息服务行业、数据中心和云计算行业等大数据相关产业链快速发展，创造大量就业机会。二是大数据试验区建设有助于催生新型就业岗位。一方面，随着大数据技术的应用，企业对于数据工程师、系统工程师、数据架构师等数据管理岗位，以及数据分析师、数据科学家等数据分析岗位的需求不断增长；另一方面，随着数据要素与传统行业深度融合，将催生出平台经济、共享经济等新的商业模式，进而孵化出众多基于数字经济的新职业（李成明等，2023）。综上分析，提出假说1。

假说1：大数据试验区建设能够直接促进就业增长。

（二）大数据试验区建设通过规模扩张效应影响就业

国家级大数据综合试验区设立的核心目标是“全面推进我国大数据发展和应用，加快建设数据强国”。该政策将会有效促进试点地区企业广泛运用大数据技术，促进数据要素效能的发挥。大数据技术的广泛运用既会影响企业外部环境，也会影响企业内部环境。

从企业外部环境来看，大数据技术运用将会增进企业与消费者、供应商和投资者三个外部主体的联系，促进生产规模扩张。首先，大数据技术运用有利于企业赢得消费者的信任青睐。企业通过收集和分析消费者的数据，可以根据消费者偏好制定个性化营销策略，改善客户消费体验，还能建立客户数据库提供全方位服务支持，增强客户购物满意度，从而与更多的客户建立长期稳定的联系，提高销售效率。消费者需求增加会激励企业扩大生产规模（毛其淋、王玥清，2023）。其次，大数据技术运用有利于企业获得供应商的及时反馈。大数据发展使得供应商和企业能够实时共享数据，供应商通过实时监测企业的销售、库存和生产数据，保障原材料稳定供应，企业通过及时追踪供应商的交货时间、库存状况数据，确保生产计划有序推进。大数据技术的运用有助于解决供应链中信息不透明、交易链冗长等问题（李健等，2022），有利于降低生产成本，提高生产效率。生产效率提升会激励企业扩大市场份额（孙伟增、郭冬梅，2021）。最后，大数据技术运用有利于企业获得投资者的资金支持。大数据运用突破了传统金融体系的限制，在移动支付、区块链、云计算等金融科技的支持下，减少资本供需双方的信息不对称，优化信贷资源配置（宋敏等，2021）。一方面，借助大数据技术，企业可以更广泛地接触到投资者，从而降低企业的搜寻成本，提高匹配过程中的信息有效性（张益豪、郭晓辉，2023），拓宽融资渠道。另一方面，借助大数据技术，企业的信息披露更加充分，与投资者的交流更加便捷，促进企业的融资效率提升、融资成本降低（任英华等，2023）。融资约束的缓解可为企业扩大生产规模提供资金支持（刘睿雯等，2020）。

根据经济学理论，当企业决定扩大生产规模时，为了满足增加的生产需求，通常需要增加生产要素的投入，其中包括劳动力。因此，随着生产规模扩张，企业对劳动力的需求也会增加，从而导致企业的就业规模扩大。根据上述分析，提出假说 2。

假说 2：大数据试验区建设通过规模扩张效应间接促进就业增长。

（三）大数据试验区建设通过产品创新效应影响就业

从企业内部环境来看，大数据技术运用改善了企业的创新环境，有助于明确产品创新方向、激发员工创新潜力、提高部门协作效率，从而促进产品创新。第一，大数据

技术运用有助于企业精准获取产品反馈，明确产品创新方向。企业借助大数据技术可以收集大量历史消费数据、社交媒体数据、市场调研数据等（任英华等，2023），从海量数据中分析用户偏好、市场趋势和需求变化，及时在产品的设计、功能改进、品牌维护等方面动态调整，针对性地进行产品的迭代创新（郭炳南等，2022），从而更加贴合消费者偏好、满足市场需求、引领潮流变化。第二，大数据技术运用有助于企业借助知识溢出效应，激发员工创新潜力。随着大数据技术的发展，丰富的数据资源促进知识生产过程，畅通的传输渠道提升知识流动速度（张慧等，2023），便捷的数字平台激发知识共享效应。积极的知识溢出效应有利于打破信息孤岛，促进不同领域知识的交叉融合，从而激发员工的创新思维，产生创新性的解决方案（李健等，2022）。第三，大数据技术运用有助于企业建立数字化协作平台，提高部门协作效率。大数据技术的运用推动企业组织结构平台化、扁平化、去中心化。数字化办公平台具有协同效应，有利于信息快速、准确、及时地在企业内传递（李健等，2022），避免由于信息不对称和沟通不充分导致的创新失误（陈文、常琦，2022），还有利于加强部门协作和资源共享，优化决策流程，确保创新项目按计划有序推进，从而快速响应市场变化。

大数据技术能够实现产业结构与市场结构、能源结构之间的最优匹配，促进产业数字化转型（史丹，2022）。利用大数据的优势降低交易费用，提高供求对接精度，使得生产与销售决策更具针对性（戴艳娟等，2023），减少技术研究开发的不确定性，鼓励创新活动（许芳等，2020）。创新方向是舵，创新人才是帆，部门协作是桨，它们共同构成了推动产品创新的关键要素。新产品开发设计完成之后，需要生产制造、推广销售、提供技术支持和售后服务等，增加了对劳动力的额外需求。Harrison et al. (2014) 的研究表明，新产品的生产销售对企业雇佣员工数量的增长产生了显著的正向影响。根据以上分析，提出假说3。

假说3：大数据试验区建设通过产品创新效应间接促进就业增长。

四 研究设计

（一）模型设定

大数据试验区的建设，一方面可能对试点区域的企业造成冲击，使得该企业的就业量在政策实施前后出现差异；另一方面也可能使得试点区域和非试点区域的就业情况在同一时点上出现差异。这一试点政策的出台为本文提供了良好的准自然实验，可使用双重差分法评估政策实施效果，在有效控制试点区域和非试点区域的事前差异基

基础上,识别出大数据试验区政策的出台是否对就业造成冲击。具体而言,将位于试点区域的企业视为处理组,位于非试点区域的企业视为控制组,建立如下基准模型:

$$labor_{it} = \beta_0 + \beta_1 (treat \times post)_{it} + \beta_2 control_{it} + \varphi_i + u_t + \varepsilon_{it} \quad (1)$$

其中,下标 i 、 t 分别表示企业和年份;被解释变量 $labor_{it}$ 表示企业 i 在第 t 年的就业规模;核心解释变量 $(treat \times post)_{it}$ 反映企业 i 所在城市于 t 年是否实施了大数据试验区建设,若是则取值为 1,否则为 0,其估计系数 β_1 反映了大数据试验区的就业效应,系数大于 0 则表明大数据试验区建设能够对试点区域企业的就业产生正向冲击; $control_{it}$ 表示一系列控制变量; φ_i 和 u_t 分别为企业固定效应和年份固定效应; ε_{it} 是随机误差项。

(二) 变量选取

1. 被解释变量

被解释变量为就业水平 ($labor$)。借鉴相关文献(毛其淋、王玥清, 2023; 王锋、葛星, 2022),采用企业员工数的对数衡量企业的就业水平。

2. 核心解释变量

核心解释变量为大数据试验区建设 ($treat \times post$)。根据公布的国家级大数据综合试验区名单,若某省或某市获批建设国家级大数据综合试验区,则位于该省或该市的所有企业为处理组, $treat$ 取值为 1, 否则为 0; 进一步根据试点政策发布时间(2016 年),政策实施当年及以后的年份 $post$ 取值为 1, 政策实施之前的年份 $post$ 取值为 0。核心解释变量为 $treat$ 和 $post$ 的交乘项。

3. 控制变量

为了控制其他因素对就业的影响,参照既有研究,加入以下变量作为控制变量。
工资水平 ($wage$): 工资是企业的劳动力成本,当工资越高时,企业可能会减少雇佣员工数量以控制成本,采用职工平均工资(应付职工薪酬/员工人数)的对数衡量。
企业规模 ($size$): 一般来说,在大部分行业中,规模较大的企业,提供的就业岗位更丰富,能够吸纳更多劳动力,采用企业总资产的对数衡量。
资产负债率 (lev): 资产负债率反映企业的财务风险和偿债能力,间接影响企业的劳动力需求(王锋、葛星, 2022),采用负债总计与资产总计的比值衡量。
销售费用率 (ser): 销售费用率反映企业的营销效率,该数值低说明企业的营销效率高,经营业绩更突出,企业有动力雇佣更多员工(Ren et al., 2020),采用销售费用与营业收入的比值衡量。
所得税率 (tax): 一方面,企业缴纳的所得税越高会增加资金成本,降低利润,降低企业的劳动力投入;另一方面,所得税高可能意味着企业生产规模大,增加劳动力需求(王锋、葛星, 2022);采

用所得税费用与利润总额的比值表示。成长能力 (*grow*)：企业成长能力强，往往意味着市场份额的持续增加、经营业务的多元化、企业规模的不断扩大，企业需要投入更多的人力资源来支持其持续发展目标（王跃堂等，2012），采用托宾 *q* 值衡量。总资产净利润率 (*roa*)：企业总资产净利润率反映了企业以全部资产获得利润的水平；高净利润率意味着企业能够获得更多的利润，这为企业提供了更多的资金来源，可以用于扩大生产规模、增加技术投资，进而带动更多的雇佣需求；采用净利润与总资产的比值衡量。

（三）数据来源

本文选用 2009 - 2021 年沪深 A 股上市公司的面板数据作为观测样本，企业数据来自国泰安数据库。为了避免异常样本影响实证结果，对数据进行如下处理：第一，剔除股票名称中含有“ST”“*ST”的公司；第二，剔除上市状态为“ST”“*ST”“暂停上市”和“终止上市”的公司；第三，剔除变量有缺失的公司；第四，对数据进行 1% 的双边缩尾处理。最终得到 3679 家上市公司的面板数据，观测值为 31679 个。主要变量的描述性统计如表 1 所示。

表 1 主要变量描述性统计

变量	变量定义	观测值	平均值	标准差	最小值	最大值
<i>labor</i>	企业员工数（对数）	31 679	7.700	1.291	4.673	11.401
<i>treat × post</i>	大数据试验区建设	31 679	0.261	0.439	0.000	1.000
<i>wage</i>	职工平均工资（对数）	31 679	9.540	1.175	5.611	12.597
<i>size</i>	企业总资产（对数）	31 679	22.274	1.452	19.923	27.463
<i>lev</i>	企业资产负债率	31 679	0.428	0.214	0.050	0.927
<i>ser</i>	企业销售费用率	31 679	0.071	0.088	0.000	0.472
<i>tax</i>	企业所得税率	31 679	0.160	0.160	-0.523	0.785
<i>grow</i>	托宾 <i>q</i> 值	31 679	2.032	1.292	0.862	8.430
<i>roa</i>	总资产净利润率	31 679	0.040	0.058	-0.224	0.195

资料来源：根据国泰安数据库 A 股上市公司数据计算得到。

五 实证结果

（一）基准回归

表 2 汇报了大数据试验区建设对企业就业规模的影响结果。列（1）只加入大数据

试验区建设这一核心解释变量，列（2）进一步控制时间固定效应和企业固定效应，列（3）在此基础上控制了企业层面其他影响就业水平的变量，列（4）采用聚类稳健标准误进行估计以获得更稳健的估计结果。从第（1）~（3）列不难发现，核心解释变量的估计系数均为正数，且在 1% 的水平上显著为正；在加入控制变量以后，系数有所下降，表明企业层面存在其他影响就业的因素，所选取的控制变量是有效的。列（4）的结果显示，大数据试验区建设对企业就业水平影响的估计系数为 0.051，通过 5% 的显著性水平检验，说明与非试点区域的企业相比，大数据试验区试点政策的施行显著增加了位于试点区域企业的员工人数，大数据试验区建设存在就业扩大效应。这一结果初步支持了假说 1。

表 2 基准回归结果

	就业水平			
	(1)	(2)	(3)	(4)
大数据试验区建设	0.366 *** (0.009)	0.070 *** (0.011)	0.051 *** (0.008)	0.051 ** (0.020)
工资水平			-0.200 *** (0.003)	-0.200 *** (0.013)
企业规模			0.661 *** (0.005)	0.661 *** (0.020)
资产负债率			0.284 *** (0.022)	0.284 *** (0.056)
销售费用率			0.813 *** (0.058)	0.813 *** (0.135)
所得税率			-0.011 (0.014)	-0.011 (0.019)
成长能力			0.028 *** (0.002)	0.028 *** (0.004)
总资产净利润率			0.110 ** (0.049)	0.110 (0.092)
企业固定效应	否	是	是	是
时间固定效应	否	是	是	是
观测值	31 679	31 679	31 679	31 679
R ²	0.063	0.897	0.941	0.941

注：***、**、* 分别表示 1%、5% 和 10% 的显著性水平；括号内数值为标准误，其中列（4）括号内数值为聚类到企业层面的稳健标准误。

资料来源：根据国泰安数据库 A 股上市公司数据计算得到。

（二）平行趋势检验

使用双重差分法的一个重要前提是需要满足平行趋势假设，即处理组和对照组在政策实施前的变化趋势是一致的，只有通过平行趋势检验，基准回归的估计结果才是可信的。借鉴 Hering & Poncet (2014) 的做法，采用样本期第一年 2009 年为基期，设定模型 (2) 检验平行趋势，其中 k 表示年份，取值从 2010 年至 2021 年， $year$ 在对应年份取值为 1，其他年份取值为 0。主要关注估计系数 β_k 的方向和显著性，根据该系数可以考察第 k 年分别位于试点城市和非试点城市企业的平均就业水平是否存在差异。

$$labor_{it} = \beta_0 + \beta_k \sum_{k=2010}^{2021} (treat \times year)_it^k + \beta_2 control_{it} + \varphi_i + u_t + \varepsilon_{it} \quad (2)$$

平行趋势检验结果如图 1 所示。在国家级大数据综合试验区批复之前， β_k 均未通过显著性检验，表明处理组企业和控制组企业在政策实施之前并无显著差异，满足平行趋势假设的要求。2016 年开始实施大数据试验区政策，实施当年 $treat \times year$ 的估计系数显著为正，且随着时间推移，系数有增大的趋势，说明大数据试验区建设对就业的提升效应不存在时滞且具有持续性，扩大就业效应随时间推移逐渐增强。

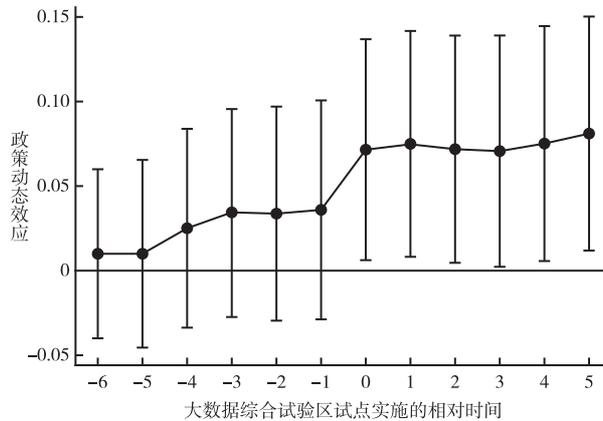


图 1 平行趋势检验

注：图中实心圆点表示估计系数 β_k ，竖线线段表示 90% 的置信区间。

资料来源：根据国泰安数据库 A 股上市公司数据计算得到。

（三）安慰剂检验

由于实证过程中难以控制所有影响就业的因素，仍可能存在一些未观测的因素对回归结果造成影响。这里借鉴 Chetty et al. (2009)、Liu & Lu (2015) 的做法，通过安

安慰剂检验增强回归结果的可信度。具体来说，首先按照大数据试点政策实施情况，在样本中随机抽取与真实处理组等数量的企业作为虚拟处理组，未抽到的则为虚拟控制组；然后基于模型（1）进行回归，得到虚拟的估计系数 β_1^{random} 和对应的 p 值；再将上述过程重复 1000 次，得到 1000 个虚拟的 β_1^{random} 及对应 p 值。绘制这 1000 个 β_1^{random} 的核密度分布和 p 值，如图 2 所示，虚拟政策冲击的系数大多分布在 0 附近，且大致服从正态分布。基准回归得到的大数据试验区建设的真实估计系数 β_1 为 0.051，位于虚假回归系数分布的高尾位置，如图 2 的虚线所示，说明真实估计值属于小概率事件。这一结果进一步支持了大数据试点政策能够显著影响就业水平。

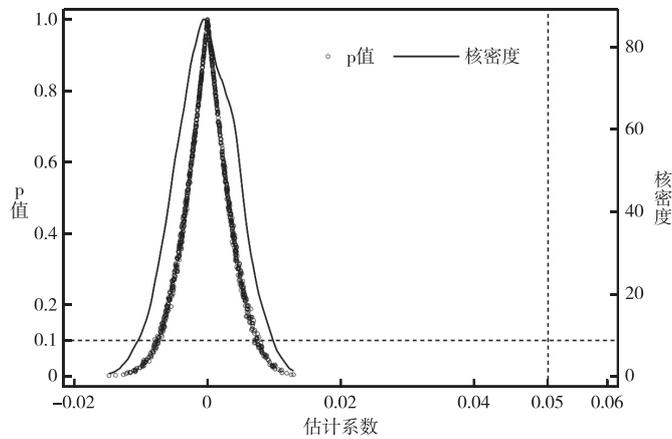


图 2 安慰剂检验

资料来源：根据国泰安数据库 A 股上市公司数据计算得到。

（四）PSM-DID 分析

大数据试验区政策并非严格的自然实验，试点区域的选择可能并不是随机的。虽然双重差分法估计出了政策的平均处理效应，仍可能存在选择性偏差问题造成的估计偏误，为此进一步采用 PSM-DID 方法检验基础回归估计结果的稳健性。

倾向得分匹配（PSM）适用于截面数据，而双重差分（DID）适用于面板数据。现有文献一般采用以下两种方式解决这一问题：第一，将面板数据看作截面数据直接进行倾向得分匹配；第二，对每年的截面数据逐年匹配然后合并成面板数据（白俊红等，2022）。构造截面 PSM 数据进行回归的具体步骤如下：第一，将基准模型中的控制变量工资水平、企业规模、资产负债率、销售费用率、所得税率、企业成长能力、总资产净利润率设定为匹配变量；第二，将所有样本视为截面数据，采用 1:1 近邻匹配为位

于大数据试验区的企业样本匹配企业层面特征相似的对照组，保留匹配成功的样本得到新数据集；第三，对匹配后生成的数据集进行平衡性检验；第四，基于新数据集重新使用 DID 方法估计大数据试验区建设对企业就业的影响效应。构造逐年面板 PSM 数据进行回归的步骤与上面类似，只有第二步不同，借鉴 Böckerman & Ilmakunnas (2009)，分别对每一年的企业样本采用 1:1 近邻匹配得到新数据集，再将所有年份的截面数据进行纵向合并得到新的面板数据。

截面 PSM 数据和逐年面板 PSM 数据均通过了平衡性检验^①。表 3 的列 (1) 和列 (2) 分别报告了基于截面 PSM 数据和逐年面板 PSM 数据进行 DID 回归的结果，核心解释变量的系数仍然为正，且均通过 5% 的显著性水平检验，与基准结果无本质差异，说明大数据试验区建设对企业就业的扩大效应是稳健的。

表 3 PSM-DID 回归结果

	截面 PSM	逐年 PSM
	(1)	(2)
大数据试验区建设	0.052 ** (0.023)	0.035 *** (0.012)
控制变量	是	是
企业固定效应	是	是
时间固定效应	是	是
观测值	16 546	16 552
R ²	0.943	0.946

注：***、**、* 分别表示 1%、5% 和 10% 的显著性水平；括号内数值为聚类到企业层面的稳健标准误。
资料来源：根据国泰安数据库 A 股上市公司数据计算得到。

(五) 稳健性检验

1. 排除其他政策干扰

大数据试验区试点政策实施的同时，还有其他相关政策实施，可能会对估计结果造成干扰。例如，工信部于 2013 年 12 月开始逐步启动国家信息消费试点政策，旨在通过实施信息基础设施提升工程，提升信息消费供给水平，扩大信息消费覆盖面，与大数据试验区的设立目标和建设任务存在部分重合（张益豪、郭晓辉，2023）。又如，住

^① 限于篇幅，平衡性检验结果未在文中汇报，备索。

建部于 2012 年 12 月开始启动国家智慧城市试点，旨在依托数字技术和数据要素，加强城市规划、建设和管理。李成明等（2023）论证了智慧城市建设具有稳就业的效果，城市智能化驱动下就业更具包容性。此外，王锋和葛星（2022）验证了低碳城市试点政策会推动试点城市全方位绿色低碳转型，进而对劳动力市场产生冲击。为避免此类政策交叉混淆大数据试验区政策的效果，在基准模型中同时控制这三项政策的虚拟变量，以剔除其他政策的干扰。结果如表 4 的列（1）所示，核心解释变量的估计系数与基准回归基本一致，表明上述政策并未对基准结果产生影响。

表 4 其他稳健性检验

	控制其他政策	预期影响	变量滞后	调整样本
	(1)	(2)	(3)	(4)
大数据试验区建设	0.052 *** (0.020)	0.038 ** (0.015)		0.049 ** (0.019)
大数据试验区建设 (前一年)		0.015 (0.019)		
大数据试验区建设 (滞后一年)			0.037 * (0.019)	
其他政策	是	否	否	否
控制变量	是	是	是	是
企业固定效应	是	是	是	是
时间固定效应	是	是	是	是
观测值	31679	31679	26868	24665
R ²	0.941	0.941	0.936	0.942

注：***、**、* 分别表示 1%、5% 和 10% 的显著性水平；括号内数值为聚类到企业层面的稳健标准误。
资料来源：根据国泰安数据库 A 股上市公司数据计算得到。

2. 考虑预期效应

在大数据试验区政策正式颁布前，不同区域可能会因为预期提前采取相关措施，从而使回归结果出现偏差。为了排除预期效应的影响，借鉴江小涓（2017）的研究，在基准模型中加入开展大数据试验区试点前一年的虚拟变量以控制预期效应。表 4 中列（2）显示，大数据试验区建设仍能显著扩大企业的劳动力需求，预期效应不存在。

3. 解释变量滞后一期

考虑到大数据试验区建设对就业的影响可能不会在政策颁布当年就体现出来，同时也为了缓解内生性问题，将核心解释变量和控制变量均滞后一期（毛其淋、王玥清，2023），重新进行回归。新的估计结果见表4的列（3）。大数据试验区建设的系数有所下降，但仍然在10%的显著性水平上为正，验证了基本结果的稳健性。

4. 筛选样本

考虑到新冠疫情在2020年初席卷全球，对企业的生产经营活动造成重大冲击，剔除2020年和2021年数据，消除疫情可能带来的干扰，结果如表4的列（4）所示。在对样本进行调整后，核心解释变量的估计系数在5%的水平上显著为正，大数据试验区建设仍然对就业具有显著的提升作用，这进一步强化了基础回归的结果。

六 机制检验

上文研究发现大数据试验区建设具有显著的就业扩大效应，那么试点政策通过何种机制影响企业的就业规模？理论分析与研究假说部分指出，大数据试验区建设可能通过规模扩大效应和产品创新效应来吸纳更多劳动力。本文借鉴薛飞等（2023）的做法，建立以下模型检验大数据试验区建设影响就业的机制：

$$M_{it} = \beta_0 + \beta_1 (treat \times post)_{it} + \beta_2 control_{it} + \varphi_i + u_t + \varepsilon_{it} \quad (3)$$

式（3）中， M_{it} 为机制变量，即大数据试验区建设可能影响就业规模的机制，其他变量的定义与式（1）一致。

为了验证规模扩大效应是否成立，借鉴毛其淋和王玥清（2023）、孙伟增和郭冬梅（2021），选择与企业生产规模直接相关的营业收入表征生产规模，用于验证大数据试验区建设所产生的规模扩大效应。表5的列（1）展示了大数据试验区建设对企业生产规模影响的估计结果。结果显示，核心解释变量的估计系数为0.695，且在1%的水平上通过了显著性检验。这说明，大数据技术运用在挖掘消费者需求、优化供应商管理、获取投资者支持方面发挥了重要作用，从而有效提高了试点地区企业的生产规模，规模扩大效应是大数据试验区建设影响就业的重要机制，证实了假说2的成立。

表 5 机制检验

	生产规模	创新数量	创新质量
	(1)	(2)	(3)
大数据试验区建设	0.695*** (0.178)	0.113* (0.061)	0.089** (0.044)
控制变量	是	是	是
企业固定效应	是	是	是
时间固定效应	是	是	是
观测值	31 679	25 397	25 397
R ²	0.919	0.785	0.819

注：***、**、* 分别表示 1%、5% 和 10% 的显著性水平；括号内数值为聚类到企业层面的稳健标准误。
资料来源：根据国泰安数据库 A 股上市公司数据计算得到。

为了验证产品创新效应是否成立，借鉴张慧等（2023）的研究，分别从数量和质量两个方面刻画企业创新水平。其中，创新数量采用实用新型专利、外观设计专利和发明专利之和衡量；创新质量采用发明专利申请数衡量。表 5 的列（2）报告了大数据试验区建设对创新数量的估计结果，系数估计值为 0.113，且在 10% 的水平上通过了显著性检验。大数据试验区建设对创新质量的影响如表 5 的列（3）所示，核心解释变量的估计系数为 0.089，且在 5% 的水平上通过了显著性检验。这说明大数据技术运用有利于企业收集产品反馈明确创新方向、通过知识溢出提升创新思维、建立数字协作平台加快创新执行，从而促进企业创新数量和创新质量的双提升，产品创新效应是大数据试验区建设影响就业的重要机制，假说 3 成立。

七 异质性分析

前文研究结果表明，大数据试验区建设具有促进就业的作用，且规模扩大效应和产品创新效应是重要机制。接下来进一步从企业、行业和城市三个维度的异质性视角考察大数据试验区建设与企业劳动力需求的关系。

（一）企业异质性

根据企业实际控制人的性质，将样本划分为国有企业和非国有企业，考察大数据试验区建设对就业的影响在国有企业和非国有企业中是否存在差异。回归结果如表 6 的列（1）和列（2）所示。可以看出，在国有企业分组中，核心解释变量的系数显著为正，说明大数据试验区试点政策的实行能够显著增加国有企业的劳动力需求；而在

非国有企业分组中，系数不显著，说明该政策对非国有企业的就业提升效应十分有限。一方面，大数据试验区试点是自上而下的顶层设计，以各级政府为推行主体，而国有企业是保障相关政策贯彻落实的主要阵地和重要抓手，在资源倾斜和政策扶持方面具备优势，更有动力迎合大数据发展要求（任英华等，2023）。另一方面，国有企业资金更雄厚，技术更先进，人才储备更丰富，更容易抓住大数据发展的机遇来促进大数据与传统业务的融合，更有实力践行大数据发展要求。因此，在推行大数据试验区建设过程中，国有企业受益更多，促进了其就业规模的增长。

表 6 异质性分析结果

	国有	非国有	技术密集	资本密集	劳动密集	数据开放高	数据开放低
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
大数据试验区建设	0.124 *** (0.031)	0.007 (0.025)	0.037 (0.025)	0.065 * (0.036)	0.085 ** (0.038)	0.054 ** (0.026)	-0.002 (0.033)
控制变量	是	是	是	是	是	是	是
企业固定效应	是	是	是	是	是	是	是
时间固定效应	是	是	是	是	是	是	是
观测值	11 679	19 625	13 956	7 954	9 670	18 292	13 371
R ²	0.944	0.937	0.950	0.955	0.947	0.947	0.934

注：***、**、* 分别表示 1%、5% 和 10% 的显著性水平；括号内数值为聚类到企业层面的稳健标准误。

资料来源：根据国泰安数据库 A 股上市公司数据计算得到。

（二）行业异质性

根据鲁桐和党印（2014）的研究，将样本划分为技术密集型行业、资本密集型行业和劳动密集型行业，考察大数据试验区建设对就业的影响在不同要素密集度企业中是否存在差异。表 6 的列（3）~（5）分别汇报了技术密集型、资本密集型、劳动密集型样本的回归结果。结果显示，大数据要素能够赋能资本密集型和劳动密集型企业，扩大就业规模，但对技术密集型企业的就业提升没有显著作用。这可能是因为，技术密集型企业天然的高科技属性，使其在大数据试验区试点政策颁布之前，就已经积累了一定的大数据资源，应用了先进的大数据技术，具备较强的信息挖掘和数据能力分析能力，大数据试点政策的边际效应递减（张慧等，2023）。而资本密集型和劳动密集型企业的数据基础较为薄弱，大数据试验区试点的推行，使得这些企业可以借助政策东风，推进传统优势业务与数据要素融合发展，释放发展潜力和增长动能。因此在

资本密集和劳动密集型的传统行业中，企业就业规模受到大数据试验区政策的冲击更强。

（三）城市异质性

根据复旦大学发布的《2021 年度中国地方政府数据开放报告》，将开放数林指数排名前 50 的城市划分为数据开放水平高的区域，其余城市为数据开放水平低的区域。分别对位于这两个区域的企业进行分组回归，讨论地方政府数据开放水平的不同，是否会影响大数据试验区试点的作用。结果如表 6 的列（6）和列（7）所示，在数据开放水平高的城市，大数据试验区建设对就业规模的作用显著为正，但在数据开放水平较低的城市，该试点政策对企业就业无显著影响。这或许是因为，城市的数据开放程度越高，当地企业可以利用更优质的数据资源，进行更深入的信息挖掘，获取更准确的决策支持（任英华等，2023）。对于当地企业来说，他们更容易发现新的商业机会、探索新的商业模式、提出新的解决方案，大幅提升当地企业的竞争力，从而促进就业规模扩大。

八 结论与政策建议

数字技术和大数据的发展对社会经济的各个领域都带来了深远的影响，国家级大数据综合试验区建设旨在进一步推动数字经济深入发展。本文将国家级大数据综合试验区试点政策视为一项准自然实验，基于 2009 - 2021 年沪深 A 股上市公司的面板数据构建双重差分模型，系统评估了大数据试验区建设对企业层面就业规模的影响。

基准回归显示，样本期内大数据试验区试点政策的推行使得试点区域企业的就业规模平均增加了约 5.1%，这说明该政策具有显著的扩大就业效应，此结论在经过平行趋势检验、安慰剂检验、PSM-DID 以及其他一系列稳健性检验后仍然成立。机制分析发现，规模扩大效应和产品创新效应是大数据试验区建设发挥就业扩大作用的影响机制。异质性分析表明，大数据试验区建设对就业的影响在不同企业、不同行业 and 不同城市中存在显著差异。在不同所有制企业中，大数据试验区建设对国有企业的劳动力雇佣数量具有正向促进作用，对非国有企业的影响不显著。在不同要素密集度行业中，大数据试验区建设对劳动密集型行业的就业促进作用最强，资本密集型行业次之，对技术密集型行业的就业没有影响。在不同城市中，大数据试验区建设显著促进数据开放水平较高城市中的企业就业，对于数据开放水平较低城市中的企业没有显著影响。上述结论对于充分利用大数据红利和实现稳就业目标具有重要的政策指导意义。

第一，鉴于大数据试验区的就业促进效应，要稳步扩大试点范围，发挥数据要素对就业的政策赋能效应。一方面，加强对试验区经验的归纳总结，系统评估试验区政策的经验教训，形成可推广复制的一般性规律，逐步扩大试点范围，充分发挥大数据试验区建设带来的就业扩张效应。另一方面，政府应制定相关政策，通过减税补贴等措施，鼓励企业在传统产业中运用大数据技术，推动新的商业模式、产业链条和职业需求的形成，激发市场活力，促进就业增长。

第二，结合机制检验结果，要重视大数据在扩大生产规模、促进产品创新方面的重要作用，疏通大数据试验区建设扩大就业规模的传导渠道。从扩大生产规模视角出发，政府要大力支持电商平台发展，加大消费者权益保护，激发线上消费活力，为生产规模扩张提供需求支撑；要鼓励企业建立数字化供应链管理系统，做好关键原材料的保供稳价，为生产规模扩张提供材料保障；要引导金融科技创新，推动金融服务变革，加强金融风险防控，为生产规模扩张提供资金支持。从促进产品创新视角出发，政府应提供财税政策优惠，加大金融机构支持，鼓励企业数字化转型，提升协作效率，保障创新项目有序推进；同时，加强知识产权保护，营造良好创新环境，提供法律保障，推动创新成果的市场转化。

第三，结合异质性分析结果，要因地制宜、因企制宜制定差异化战略，释放数据要素促进就业的潜力。具体来看，对于非国有企业，政府要设立专项资金，支持开展数字化技术的科技研发，同时建立产学研合作平台，着力培养适应数字化转型的人才队伍。对于资本密集和劳动密集型行业，政府要出台相关政策，鼓励倡导引入大数据技术，简化生产流程、提高生产效率、优化资源配置，发挥后发优势，推动行业的高质量发展，缓解就业压力。对于数据开放水平较低城市，政府要稳步推进公共数据的有序开放，促进数据资源的流通利用，释放数据要素的经济价值。

参考文献：

- 白俊红、张艺璇、卞元超（2022），《创新驱动政策是否提升城市创业活跃度——来自国家创新型城市试点政策的经验证据》，《中国工业经济》第6期，第61-78页。
- 蔡跃洲、马文君（2021），《数据要素对高质量发展影响与数据流动制约》，《数量经济技术经济研究》第3期，第64-83页。
- 常皓亮、金磊、薛飞（2023），《大数据战略对电力消费碳排放的影响——基于国家级

- 大数据综合试验区的准自然实验》，《经济与管理研究》第5期，第93-109页。
- 陈贵富、韩静、韩恺明（2022），《城市数字经济发展、技能偏向型技术进步与劳动力不充分就业》，《中国工业经济》第8期，第118-136页。
- 陈加友（2017），《国家大数据（贵州）综合试验区发展研究》，《贵州社会科学》第12期，第149-155页。
- 陈文、常琦（2022），《大数据赋能了企业绿色创新吗——基于国家级大数据综合试验区的准自然实验》，《财经科学》第8期，第76-92页。
- 陈文、吴赢（2021），《数字经济发展、数字鸿沟与城乡居民收入差距》，《南方经济》第11期，第1-17页。
- 戴艳娟、沈伟鹏、谭伟杰（2023），《大数据发展对企业数字技术创新的影响研究——基于大数据试验区的准自然实验》，《西部论坛》第2期，第16-28页。
- 董香书、王晋梅、肖翔（2022），《数字经济如何影响制造业企业技术创新——基于“数字鸿沟”的视角》，《经济学家》第11期，第62-73页。
- 耿伟、王筱依、王鑫源（2023），《国家级大数据综合试验区与FDI流入——基于数量和质量视角》，《国际经贸探索》，第1期，第19-35页。
- 郭炳南、王宇、张浩（2022），《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》，《广东财经大学学报》第1期，第58-74页。
- 侯世英、宋良荣（2021），《数字经济、市场整合与企业创新绩效》，《当代财经》第6期，第78-88页。
- 黄漫宇、王孝行（2022），《数字经济、资源错配与企业全要素生产率》，《宏观经济研究》第12期，第43-53页。
- 江小涓（2017），《高度联通社会中的资源重组与服务业增长》，《经济研究》第3期，第4-17页。
- 李成明、王霄、李博（2023），《城市智能化、居民劳动供给与包容性就业——来自准自然实验的证据》，《经济与管理研究》第3期，第41-59页。
- 李健、张金林、董小凡（2022），《数字经济如何影响企业创新能力：内在机制与经验证据》，《经济管理》第8期，第5-22页。
- 李言、朱棋娜（2023），《大数据技术如何影响绿色全要素生产率——来自国家大数据综合试验区试点的经验分析》，《环境经济研究》第3期，第27-45页。
- 刘军、朱可、钱宇（2024），《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》，《南京审计大学学报》第1期，第101-111页。

- 刘睿雯、徐舒、张川川 (2020), 《贸易开放、就业结构变迁与生产率增长》, 《中国工业经济》第6期, 第24-42页。
- 鲁桐、党印 (2014), 《公司治理与技术创新: 分行业比较》, 《经济研究》第6期, 第115-128页。
- 毛其淋、王玥清 (2023), 《ESG的就业效应研究: 来自中国上市公司的证据》, 《经济研究》第7期, 第86-103页。
- 孟祺 (2021), 《数字经济与高质量就业: 理论与实证》, 《社会科学》第2期, 第47-58页。
- 邱子迅、周亚虹 (2021), 《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》, 《财经研究》第7期, 第4-17页。
- 任英华、刘宇钊、胡宗义、李海彤 (2023), 《大数据发展、知识产权保护对企业绿色技术创新的影响》, 《中国人口·资源与环境》第7期, 第157-167页。
- 史丹 (2022), 《数字经济条件下产业发展趋势的演变》, 《中国工业经济》第11期, 第26-42页。
- 宋敏、周鹏、司海涛 (2021), 《金融科技与企业全要素生产率——“赋能”和信贷配给的视角》, 《中国工业经济》第4期, 第138-155页。
- 孙伟增、郭冬梅 (2021), 《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》, 《中国工业经济》第11期, 第78-96页。
- 孙哲远 (2022), 《数字经济发展如何影响制造业企业“脱虚向实”? ——来自国家级大数据综合试验区的证据》, 《现代经济探讨》第7期, 第90-100页。
- 王春超、聂雅丰 (2023), 《数字经济对就业影响研究进展》, 《经济学动态》第4期, 第134-149页。
- 王锋、葛星 (2022), 《低碳转型冲击就业吗——来自低碳城市试点的经验证据》, 《中国工业经济》第5期, 第81-99页。
- 王跃堂、王国俊、彭洋 (2012), 《控制权性质影响税收敏感性吗? ——基于企业劳动力需求的检验》, 《经济研究》第4期, 第52-63页。
- 许芳、田萌、徐国虎 (2020), 《大数据应用能力对企业创新绩效的影响研究——供应链协同的中介效应与战略匹配的调节效应》, 《宏观经济研究》第3期, 第101-119页。
- 薛飞、周民良、刘家旗 (2023), 《产业转型升级能否降低碳排放? ——来自国家产业转型升级示范区的证据》, 《产业经济研究》第2期, 第1-13页。

- 杨俊、李小明、黄守军 (2022), 《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》, 《经济研究》第 4 期, 第 103 - 119 页。
- 余明桂、马林、王空 (2022), 《商业银行数字化转型与劳动力需求: 创造还是破坏?》, 《管理世界》第 10 期, 第 212 - 230 页。
- 袁冬梅、吕书杰、龙瑞 (2023), 《数字经济发展如何影响了服务业就业? ——来自微观层面的经验证据》, 《劳动经济研究》第 1 期, 第 81 - 102 页。
- 张慧、易金彪、徐建新 (2023), 《数字化变革如何影响城市创新——基于大数据试验区建设的经验证据》, 《科学学研究》第 8 期, 第 1484 - 1494 页。
- 张益豪、郭晓辉 (2023), 《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》, 《产业经济研究》第 2 期, 第 69 - 82 页。
- 支宇鹏、卢潇潇 (2023), 《大数据试验区设立与城市创业活跃度——基于 284 个城市的经验证据》, 《中国流通经济》第 3 期, 第 84 - 96 页。
- 周晓光、肖宇 (2023), 《数字经济发展对居民就业的影响效应研究》, 《中国软科学》第 5 期, 第 158 - 170 页。
- Acemoglu, Daron & Pascual Restrepo (2019). Artificial Intelligence, Automation, and Work. In Ajay Agrawal, Joshua Gans & Avi Goldfarb (eds.), *The Economics of Artificial Intelligence: An Agenda*. Chicago: University of Chicago Press, pp. 197 - 236.
- Arntz, Melanie, Terry Gregory & Ulrich Zierahn (2016). The Risk of Automation for Jobs in OECD Countries: A Comparative Analysis. *OECD Social, Employment and Migration Working Papers*, No. 189.
- Autor, David & David Dorn (2013). The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market. *American Economic Review*, 103 (5), 1553 - 1597.
- Autor, David, Lawrence Katz & Melissa Kearney (2006). The Polarization of the U. S. Labor Market. *American Economic Review*, 96 (2), 189 - 194.
- Böckerman, Petri & Pekka Ilmakunnas (2009). Unemployment and Self-Assessed Health: Evidence from Panel Data. *Health Economics*, 18 (2), 161 - 179.
- Brynjolfsson, Erik & Kristina McElheran (2016). The Rapid Adoption of Data-Driven Decision-Making. *American Economic Review*, 106 (5), 133 - 139.
- Chatterjee, Sheshadri, Ranjan Chaudhuri, Mahmood Shah & Pratik Maheshwari (2022). Big Data Driven Innovation for Sustaining SME Supply Chain Operation in Post COVID - 19 Scenario: Moderating Role of SME Technology Leadership. *Computers & Industrial*

Engineering, 168, 108058.

- Chetty, Raj, Adam Looney & Kory Kroft (2009). Saliency and Taxation: Theory and Evidence. *American Economic Review*, 99 (4), 1145 – 1177.
- Corcoran, Adrian & Robert Gillanders (2015). Foreign Direct Investment and the Ease of Doing Business. *Review of World Economics*, 151 (1), 103 – 126.
- David, Benjamin (2017). Computer Technology and Probable Job Destructions in Japan: An Evaluation. *Journal of the Japanese and International Economies*, 43, 77 – 87.
- Goldfarb, Avi & Catherine Tucker (2019). Digital Economics. *Journal of Economic Literature*, 57 (1), 3 – 43.
- Guo, Bingnan, Yu Wang, Hao Zhang, Chunyan Liang, Yu Feng & Feng Hu (2023). Impact of the Digital Economy on High-Quality Urban Economic Development: Evidence from Chinese Cities. *Economic Modelling*, 120, 106194.
- Harrison, Rupert, Jordi Jaumandreu, Jacques Mairesse & Bettina Peters (2014). Does Innovation Stimulate Employment? A Firm-Level Analysis Using Comparable Micro-Data from Four European Countries. *International Journal of Industrial Organization*, 35, 29 – 43.
- Hering, Laura & Sandra Poncet (2014). Environmental Policy and Exports: Evidence from Chinese Cities. *Journal of Environmental Economics and Management*, 68 (2), 296 – 318.
- Jones, Charles & Christopher Tonetti (2020). Nonrivalry and the Economics of Data. *American Economic Review*, 110 (9), 2819 – 2858.
- Liu, Qing & Yi Lu (2015). Firm Investment and Exporting: Evidence from China's Value-Added Tax Reform. *Journal of International Economics*, 97 (2), 392 – 403.
- Ndubuisi, Gideon, Chuks Otioma & Godsway Tetteh (2021). Digital Infrastructure and Employment in Services: Evidence from Sub-Saharan African Countries. *Telecommunications Policy*, 45 (8), 102153.
- Ren, Shenggang, Donghua Liu, Bo Li, Yangjie Wang & Xiaohong Chen (2020). Does Emissions Trading Affect Labor Demand? Evidence from the Mining and Manufacturing Industries in China. *Journal of Environmental Management*, 254, 109789.
- Shackleton, Len (2018). Robocalypse Now? Why We Shouldn't Panic about Automation, Algorithms and Artificial Intelligence. *Institute of Economic Affairs Current Controversies*, No. 61.

The Employment Effect of National Big Data Comprehensive Pilot Zones

Wang Zhiyong¹ & Yang Jinjiao²

(Institute of Population and Labor Economics, Chinese Academy of Social Sciences¹;

Faculty of Applied Economics, University of Chinese Academy of Social Sciences²)

Abstract: Employment is a crucial aspect of people's livelihood. In the current economic downturn, it is important to stabilize employment. With the rise of the digital economy, empowering enterprises with big data to create more job opportunities has become a priority. This study examines the establishment of national big data comprehensive pilot zones as a quasi-natural experiment. Using panel data from 2009 to 2021 of companies listed on the Shanghai and Shenzhen stock exchanges, we have employed the difference-in-differences method to analyze the employment effect of the national big data pilot zones. Our results indicate that establishing the national big data comprehensive pilot zones increases the labor demand of local enterprises. This finding is robust across a range of tests. Mechanism analysis shows that the policy primarily drives production expansion and innovation improvement, thereby increasing the demand for labor. Heterogeneity analysis demonstrates that the employment expansion effect of the policy is more pronounced in state-owned enterprises, labor-intensive and capital-intensive industries, and cities with high levels of data openness. These conclusions provide valuable insights for promoting the application of big data technology and stabilizing employment.

Keywords: big data comprehensive pilot zone, employment, digital economy

JEL Classification: H54, J21, M51

(责任编辑: 合 羽)