

# 公共数据开放与就业机会

——来自 6800 万条在线招聘信息的证据

韦庆芳 魏下海 刘 缘 余玲铮\*

**内容提要** 有序开放公共数据、激发数字红利是促进劳动力市场发展的关键举措。本文利用全国 6800 万条在线招聘岗位信息构建在线招聘需求指标，将各地方政府上线公共数据开放平台作为准自然实验，考察公共数据开放对当地就业机会的影响。研究发现，公共数据开放显著增加了当地在线招聘需求，且这一效果在开放数据量级较大的地区更明显。作用机制在于，公共数据开放可通过激发创业活力、推动数实融合发展、促进人口集聚发挥就业机会创造效应。公共数据开放的效果充分体现了包容性和均等性，即企业愿意提供更多面向低学历、无工作经验和低收入群体的岗位需求。此外，公共数据开放主要增加服务业、数字行业、劳动和资本密集型行业的岗位需求。在传统要素流通程度较低、市场化程度较低、信息可达性较高的城市中，公共数据开放的就业机会创造效果更佳。本研究有助于深化对公共数据开放与劳动力市场发展的理解，为发挥数据要素的就业机会创造效应提供有益的政策启示。

**关键词** 数据要素 公共数据开放 就业机会 在线招聘

## 一 引言

就业是最大的民生，也是经济发展最基本的支撑。面对当前复杂严峻的外部环境，

\* 韦庆芳，广东财经大学金融学院，电子邮箱：weiqf1995@126.com；魏下海，华侨大学经济与金融学院，电子邮箱：xiahaiwei2005@126.com；刘缘，广东财经大学金融学院，电子邮箱：cooly@student.gdufe.edu.cn；余玲铮（通讯作者），华侨大学经济与金融学院，电子邮箱：lingzhengyu2005@126.com。作者感谢国家社会科学基金一般项目（项目编号：22BJL075）的资助。文责自负。

以及新一轮科技革命和产业变革,就业领域出现了诸多新变化新趋势。自 2019 年起,政府工作报告连续多年强调“就业优先”的政策导向,凸显党和政府对就业问题的高度重视。如何稳住就业基本盘、扩大就业容量,是当前亟待研究的重大课题。近年来,互联网和数字技术的迅速发展使得数据逐渐成为生产要素,融入人们的生产生活之中,其在经济活动过程中的作用越来越重要。党和政府高度重视数据要素的价值实现,通过一系列政策举措和制度建设,积极推动数据资源的深度挖掘和有效利用,其对劳动力市场高质量发展的作用也被寄予厚望。在此背景下,数据要素如何影响就业机会是值得深入研究的问题。

随着数字经济时代的到来,学者们围绕数字经济与就业展开了较为丰富的研究。这些研究按照就业数据来源大致可以分为两类:一是家庭或个人就业数据,二是企业员工数据。这些数据在不同年份间的变化相对较小,从而影响估计结果的可靠性(蔡卫星等,2023)。同时,由于数据要素的识别存在困难,且缺乏统一的测度方法,目前关于数据要素的实证研究相对不足。近年地方政府相继上线公共数据平台,这为识别数据要素对就业机会的影响提供了良好的契机,使得本文能够借助事件冲击识别数据要素的因果效应。理论上讲,公共数据开放为市场参与者提供极具价值的数据和信息资源,对企业投资、数实融合发展、劳动力流动等一系列经济活动产生积极影响,并进一步影响当地就业机会。公共数据开放是否能够稳就业和促就业?作用机制是什么?这些问题都有待进一步探讨。

基于上述背景,本文借助爬虫工具搜集 6800 万条在线招聘岗位信息构建在线招聘需求指标,并以各地方政府上线公共数据开放平台作为准自然实验,考察公共数据开放对就业机会的影响。本文的边际贡献主要体现在三个方面。首先,拓展了关于数据要素经济后果的文献边界。现有关于数据要素的文献大多从理论上分析数据要素的经济效应(陈晓红等,2022),经验研究主要关注数据要素对宏观层面的区域协调发展(方锦程等,2023)、城投债定价(欧阳伊玲等,2024),以及微观层面的企业创新(Ghasemaghaei & Calic, 2019)、企业生产率(彭远怀,2023)、企业市场价值(张叶青等,2021)等的影响。不同的是,本文从就业机会视角出发考察数据要素对在线招聘需求的影响及机制。

其次,本文通过搜集海量在线招聘信息,构建了在线招聘需求指标。随着大数据方法的逐渐普及,在线招聘数据在近年受到越来越多的关注,在数据量和颗粒度上具有明显优势(蔡卫星等,2023; Hershbein & Kahn, 2018)。从现有文献看,囿于数据限制,从需求侧出发测度就业机会的研究还比较少,大多是基于家户或个人就业状态

(宁光杰等, 2023)、微观企业员工人数(李磊等, 2021)等数据构建的就业状态指标。相较而言, 在线招聘需求指标能够及时地反映劳动力市场新增就业机会, 并且可以进一步细分为不同类型的就业机会, 具有较好的拓展价值。

最后, 本文充实了就业领域的文献研究。就业作为最大的民生, 一直备受学术界关注。现有文献对就业的影响因素展开了广泛讨论, 主要涵盖全球化(Feenstra et al., 2019)、技术进步(余玲铮等, 2021)、绿色发展(Liu et al., 2021)、经济政策手段(王贤彬、陈春秀, 2023)以及企业、家庭和个体特征(Ring, 2023)等方面。相比之下, 本文基于各地方政府上线公共数据开放平台的准自然实验, 探究数据这种新型生产要素如何影响就业机会。

本文余下内容的安排如下: 第二部分是制度背景, 简要回顾公共数据开放的发展历程和基本现状; 第三部分是理论分析和研究假设; 第四部分是研究设计, 介绍数据来源、计量模型、变量设定和描述性统计; 第五部分和第六部分报告实证结果, 包括基准回归结果、稳健性检验、机制检验和异质性讨论; 最后总结全文。

## 二 制度背景

在全球范围内, 公共数据开放已成为推动政务透明化、社会创新和经济增长的重要趋势。美国在2009年通过推出data.gov网站, 引领了全球公共数据开放的潮流。2011年, 美国携手英国、挪威、巴西、墨西哥、印度尼西亚、菲律宾和南非等八国共同签署了“开放数据声明”, 并发起了开放政府合作伙伴计划(OGP)。截至2023年底, 全球已有75个国家和地区加入该计划, 表明公共数据开放已成为国际共识, 各国正通过协作和标准化实践加速这一进程。

在此国际背景下, 中国积极顺应全球趋势, 近年来高度重视公共数据开放工作, 相继出台了相关行动纲领和政策措施。2015年8月, 国务院印发的《促进大数据发展行动纲要》明确指出, 要稳步推动公共数据资源开放。此后, 《中华人民共和国国民经济和社会发展第十四个五年规划和2035年远景目标纲要》强调扩大基础公共信息数据的安全有序开放。2020年3月, 《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》提出要加快培育数据要素市场, 特别强调推进政府数据开放共享、提升社会数据资源价值以及加强数据资源整合和安全保护, 为中国数据要素市场的发展指明了方向。在此基础上, 2021年12月, 国务院印发的《“十四五”数字经济发展规划》进一步明确, 要加快构建数据要素市场规则, 培育市场主体、完善治理体系,

目标是到 2025 年初步建立数据要素市场体系。2022 年 12 月，中共中央、国务院印发“数据二十条”，即《关于构建数据基础制度更好发挥数据要素作用的意见》，从数据产权、流通交易、收益分配、安全治理等方面提出了二十条政策措施，为加快培育发展数据要素市场提供了明确指引。总体而言，中国在公共数据开放方面已经迈出了坚实的步伐，并且随着相关政策的不断推进和实施，有望在这一领域取得更大的进展。

从现实数据看，自 2012 年以来，上海市和北京市率先上线公共数据开放平台，引领着各地方政府公共数据开放平台的发展。根据复旦大学数字与移动治理实验室发布的《2023 年中国地方公共数据开放利用报告（城市）》，各地方政府的公共数据开放平台数量从 2017 年的 20 个增长到 2023 年下半年的 226 个（见图 1）。在中国 337 个城市（含直辖市、副省级与地级行政区）中，有 60.53% 的城市已经上线了公共数据开放平台。这一趋势表明，中国政府在推进公共数据开放方面成效显著，正逐步实现公共数据的全面开放。

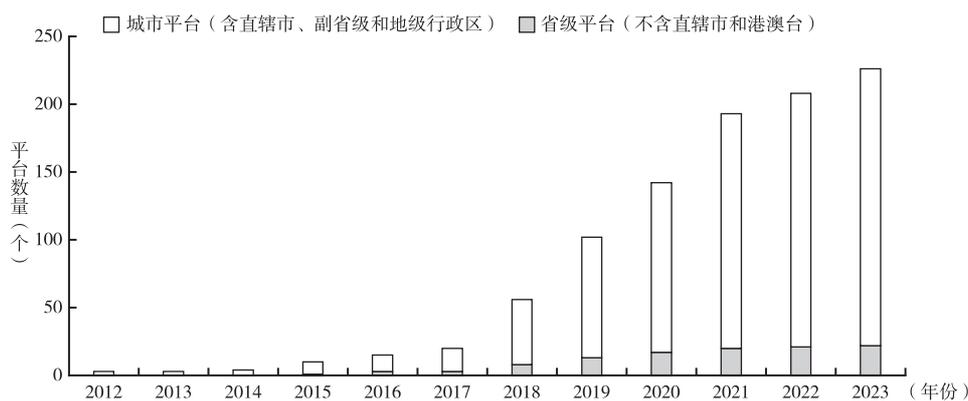


图 1 历年地级市及以上数据开放平台数量的增长情况

资料来源：根据复旦大学数字与移动治理实验室发布的《2023 年中国地方公共数据开放利用报告（城市）》计算得到。

与此同时，政府开放的公共数据集也呈现多元化的特点。政府开放的公共数据集中，社会民生、社保就业和经贸工商三个领域的数据集数量位居前列，分别有 5.04 万、4.52 万和 2.35 万个（见图 2），表明社会民生、就业和经济领域是公共数据开放的重点领域。值得一提的是，这几个重点领域与企业投资、数实融合和劳动力流动密切相关。此外，信用服务领域的数据集相对较少，这主要是由该领域的特殊性质决定的。信用服务数据涉及个人隐私和商业机密等敏感信息，因此在开放过程中需要更加谨慎和规范，以确保数据的安全性和合规性。

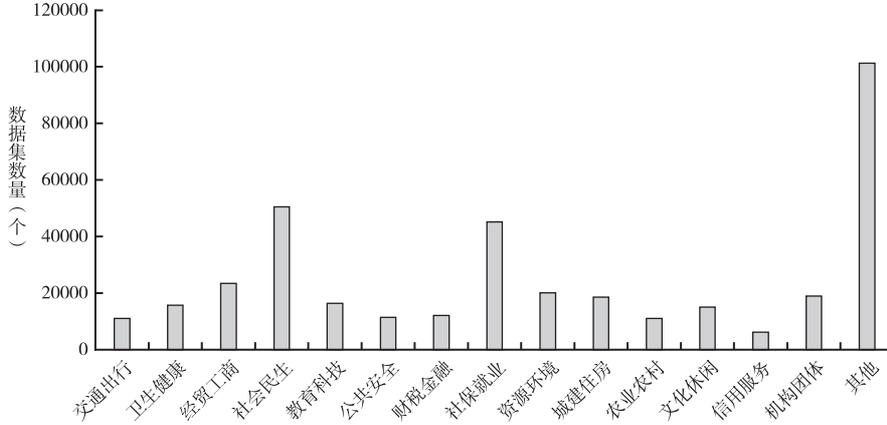


图 2 公共数据开放数据集数量的主题分布

资料来源：根据复旦大学数字与移动治理实验室发布的《2023 年中国地方公共数据开放利用报告（城市）》计算得到。

### 三 理论分析与研究假设

公共数据作为一种生产要素，具有边际成本为零、非竞争性、非排他性和规模报酬递增等特性，使其成为高效、有价值且具有公共物品属性的资源（陈晓红等，2022）。信息经济学相关理论认为，信息是一种重要的生产要素，其非竞争性和非排他性使得信息可以被广泛共享和使用，从而提高整个社会的生产效率和福利水平（Stiglitz，2000）。公共数据作为信息的一种形式，其开放和共享有助于减少信息不对称和市场失灵，从而提高市场的运行效率。公共物品相关理论强调，公共物品具有非排他性和非竞争性的特性，其供给和消费不依赖于市场机制（Samuelson，1954）。公共数据作为一种公共品，其边际成本为零，意味着数据的提供和维护成本与数据使用者的数量无关，这使得政府成为公共数据的主要供给者。

#### （一）创业活跃度机制

公共数据开放能够显著促进创业活动增加。一方面，通过开放政府公共数据，创业者能够减少信息不对称，更容易获得更为丰富、准确的信息和数据资源，如公共交通数据、人口分布数据、市场调查数据等。这些数据资源能够帮助创业者更加深入地了解市场动态、消费者需求、行业趋势和竞争格局。在此基础上，创业者能够更准确地识别潜在的市场需求，把握商机，从而制定出更为精准、有效的商业策略（Brynjolfsson &

McElheran, 2016)。另一方面，公共数据开放也有效地降低了创业的门槛和成本。公共数据开放为创业者提供了更低成本、便捷的数据获取渠道，从而减少了数据收集的成本和时间。在传统的模式下，创业者需要投入大量的人力、物力和财力采集和处理数据。而公共数据开放可以为创业者提供更为便捷的数据获取渠道和更为高效的数据处理方式，减少在数据收集和处理方面的投入，从而提高企业的运营效率（Farboodi et al., 2019）。上述优势使得更多的个体能够参与到创业活动中，从而增加了创业的活跃度。

创业是就业之源，对在线招聘需求有深远影响。一方面，新企业创建对就业有直接的推动作用，即直接催生在线招聘需求岗位增加（Bernal et al., 2024）。另一方面，新企业创建还会间接引发一系列的产业链活动，即推动供应商和分销商等企业的发展，从而进一步增加在线招聘需求。基于上述分析，提出本文的第一个研究假设。

H1：公共数据开放有助于提升创业活跃度，进而增加就业机会。

### （二）数实融合机制

数实融合，即数字经济与实体经济的深度融合，正深刻改变着经济结构与就业市场，显著影响就业机会（夏杰长、苏敏，2024）。首先，数实融合过程催生了大量新兴业态与岗位，如数据分析师、数字营销专员、智能制造工程师等。其次，传统行业在数字化转型过程中，也需要吸纳更多具备数字化技能的人才，以支撑其业务流程的智能化改造与升级，带动了相关就业机会的增长，如信息技术支持、数字化产品经理等。此外，数实融合还促进了产业链的延伸与跨界融合，创造出更多跨领域的就业机会。例如，农业与电商的结合催生了农村电商专员，制造业与服务业的融合则带动了服务型制造等新兴岗位的出现。

公共数据开放作为数实融合的关键催化剂，能够有效打破政府与市场的信息隔阂，释放海量政府数据资源，为数字经济注入强劲动能（方锦程等，2023）。这些数据的广泛流通，不仅丰富了数字经济的数据生态，还成为城市治理、产业升级及创新创业的重要支撑（张晨等，2023）。企业借助开放数据，加速产品创新与服务优化，开辟市场新蓝海。而传统产业则通过数字化转型，利用数据洞察市场需求，提升竞争力。因此，公共数据开放不仅是数据要素市场发展的重要里程碑，更是推动数实融合发展的关键路径。基于上述分析，提出本文的第二个研究假设。

H2：公共数据开放有助于推动数实融合发展，进而增加就业机会。

### （三）人口集聚机制

公共数据开放通过提供全面、准确、及时的信息，帮助流动人口缓解信息不对称，以更好地选择工作和生活的地点，从而进行更为合理流动和集聚（Huang et al.,

2022)。其影响主要体现在以下三个方面。首先，公共数据开放涵盖政府政策和服务信息，使流动人口能全面评估各地区的吸引力，做出更明智的选择。其次，通过公开的交通数据和城市规划信息，流动人口能更清晰地了解城市基础设施状况，从而选择更便利的工作和生活地点。最后，政府开放的全面社会保障和公共服务信息为流动人口提供了重要的安全保障。了解当地医疗、教育和社保等公共服务的质量和覆盖范围，可以增强流动人口安全感，使其更加安心地选择工作和生活地点。

人口集聚对就业机会具有重要影响。首先，集聚效应是推动城市产业发展的重要因素。随着人口和资源的集聚，城市内各产业之间的联系更加紧密，形成具有竞争力的产业链（蔡庆丰等，2023），不仅促进产业的快速发展，还提供更多的就业机会。其次，人口集聚对服务行业的发展具有重要影响。随着人口数量的增加，人们对服务的需求也随之增加。在餐饮、零售、医疗、教育等领域，人口集聚为这些行业提供了更多的消费者和市场机会，因此产生大量的就业机会。最后，人口集聚使得市场匹配效率显著提升。随着人口的流入，技术市场、风险投资市场、劳动力市场变得更“厚”，这提高了市场匹配效率，进而增加就业机会（Di Addario，2011）。基于上述分析，提出本文的第三个研究假设。

H3：公共数据开放有助于促进人口集聚，进而增加就业机会。

## 四 研究设计

### （一）数据来源

本文的数据来源主要包括以下三类。一是在线招聘需求数据。本文运用大数据技术搜集了2018年至2021年国内五大在线招聘网站（前程无忧、BOSS直聘、智联招聘、猎聘网、拉勾网）发布的所有招聘信息。在搜集数据时，以日度为单位进行持续获取，以尽可能地保证数据的完整性。对于在招聘网站上存在超过一天的同一条招聘信息，本文仅将其统计至发布当天，以避免重复计算。当同一条招聘信息在不同网站发布时，本文根据招聘单位名称、发布时间、岗位名称、工作地点和职位描述进行甄别，确保数据的唯一性。同时，本文仅保留已进行工商登记注册的单位招聘信息，以排除可能存在的虚假信息问题。最终，本文得到近6800万条有效的在线招聘需求数据。

二是各地方政府公共数据开放平台数据。该数据来源于复旦大学数字与移动治理实验室发布的《中国地方政府数据开放报告（城市）》（方锦程等，2023；欧阳伊玲等，2024；彭远怀，2023）。本文进一步通过在百度搜索相关关键词的方式核对平台网

站是否存在，并根据相关公开报道核对平台上线时间。

三是城市层面数据。城市层面的数据主要来源于多个权威数据库和资料。人工智能数据是基于国家知识产权局专利信息和《战略性新兴产业分类与国际专利分类参照关系表（2021）（试行）》中的智能制造专利分类一一比对得到的。政府就业关注度是基于全部地级市政府工作报告采用文本分析方法计算得到的。余下指标来源于《中国城市统计年鉴》。对于年鉴中缺失的部分数据，本文通过省统计局的数据进行补充。

## （二）模型设定

为评估公共数据开放对就业机会的影响，本文将整理的近 6800 万条在线招聘信息统计到城市 - 年份层面，并构建城市 - 年份面板数据作为基准样本。值得一提的是，本文进一步剔除了样本期之前已经上线公共数据平台的城市样本。参考方锦程等（2023）的做法，使用交叠双重差分（DID）模型进行实证检验，具体的设定如下：

$$Jobs_{c,t} = \alpha_0 + \alpha_1 \times Dataopen_{c,t} + \beta' \times Z + v_c + \mu_t + \varepsilon_{c,t} \quad (1)$$

其中，被解释变量  $Jobs_{c,t}$  是城市  $c$  在  $t$  年对数化的在线招聘需求数量。核心解释变量  $Dataopen_{c,t}$  是公共数据平台上线虚拟变量：若城市  $c$  在  $t$  年上线公共数据平台，则该城市在当年及之后取值为 1，否则为 0。 $Z$  是一系列控制变量，包括工资水平、产业结构、人口规模、人力资本、金融发展水平、对外开放程度、人工智能和政府就业关注度（方锦程等，2023）。 $v_c$  和  $\mu_t$  依次表示城市和年份固定效应， $\varepsilon_{c,t}$  是随机扰动项。 $\alpha_1$  是本文关注的核心估计系数，表示公共数据开放对就业机会的影响。表 1 是本文主要变量的具体定义。

表 1 主要变量定义

变量名称	变量符号	变量定义
就业机会	<i>Jobs</i>	城市在线招聘岗位数量（个）的自然对数
公共数据开放	<i>Dataopen</i>	城市是否已经上线公共数据平台（是 = 1）
地区生产总值	<i>GDP</i>	地区生产总值（万元）的自然对数
工资水平	<i>Wage</i>	城市职工平均工资（元/人）的自然对数
产业结构	<i>Industry</i>	城市第三产业产值占比（%）
人口规模	<i>Population</i>	城市人口规模（万人）的自然对数
人力资本	<i>Edu</i>	城市高校学生在校人数（人）的自然对数
金融发展水平	<i>Loan</i>	城市贷款余额占 GDP 的比重（%）
对外开放程度	<i>FDI</i>	城市对外直接投资占 GDP 的比重（%）
人工智能	<i>Robot</i>	城市人工智能专利（个）的自然对数
政府就业关注度	<i>JIE</i>	城市政府工作报告中就业词频（个）的自然对数

### （三）描述性统计

表2报告了主要变量的描述性统计结果。根据原始数据，就业机会变量的原值即城市在线招聘岗位数量的均值为41200，表明在本文的样本内，平均每个城市每年发布了4万多个在线招聘需求，反映了在线招聘在现代劳动力市场中的重要性和代表性。就业机会变量的标准差较大，意味着样本期间不同城市之间的就业机会存在较大差异。公共数据开放变量的均值为0.319，表明在样本中大约有32%的样本上线了公共数据平台。这一比例相对较高，说明公共数据开放正受到越来越多的重视。其他控制变量的统计结果与相关研究相近，不再赘述。

表2 描述性统计

变量符号	平均值	标准差	中位数	25分位数	75分位数
<i>Jobs</i>	8.666	1.848	8.510	7.501	9.640
<i>Dataopen</i>	0.319	0.466	0	0	1
<i>GDP</i>	16.700	0.832	16.654	16.123	17.243
<i>Wage</i>	11.173	0.191	11.166	11.045	11.296
<i>Industry</i>	46.750	7.676	46.255	41.455	50.915
<i>Population</i>	5.859	0.672	5.920	5.455	6.346
<i>Edu</i>	10.699	1.178	10.660	9.945	11.436
<i>Loan</i>	2.764	1.163	2.451	1.968	3.238
<i>FDI</i>	1.466	2.708	0.691	0.007	2.043
<i>Robot</i>	1.158	1.464	0.693	0.000	1.946
<i>JIE</i>	2.324	0.404	2.303	2.079	2.565

资料来源：根据前程无忧、BOSS直聘、智联招聘、猎聘网、拉勾网五大在线招聘网站、复旦大学数字与移动治理实验室、国家知识产权局专利信息、地级市政府工作报告和《中国城市统计年鉴》数据计算得到。

## 五 实证结果分析

### （一）主要估计结果

#### 1. 基准估计结果

表3报告了公共数据开放对就业机会的估计结果。本文在估计基准模型时，采用逐步加入控制变量的方式进行。第（1）列为仅考虑公共数据开放单一变量和城市、年份双向固定效应的估计结果。从估计结果上看，公共数据开放的估计系数为正，且在1%的水平下显著，表明公共数据开放显著增加在线招聘需求。第（2）列进一步加入

一系列影响在线招聘需求的经济社会因素。估计结果显示，公共数据开放的估计系数仍然在 1% 的水平下显著为正，表明在考虑一系列潜在的影响因素之后，公共数据开放对在线招聘需求仍然有显著的正向影响。从经济意义上看，公共数据开放带来了 32% 的在线招聘需求增长，是一个非常可观的数字。从控制变量的估计系数上看，经济发展水平较高、工资水平较高、第三产业较发达、外商投资较多、政府对就业的关注度越高的城市，在线招聘需求越多。

表 3 公共数据开放影响就业机会的基准估计

	(1)	(2)
<i>Dataopen</i>	0.373 *** (0.086)	0.324 *** (0.087)
<i>GDP</i>		0.901 ** (0.378)
<i>Wage</i>		0.965 *** (0.353)
<i>Industry</i>		0.031 *** (0.008)
<i>Population</i>		0.143 (0.180)
<i>Edu</i>		0.225 (0.145)
<i>Loan</i>		0.149 (0.134)
<i>FDI</i>		0.044 ** (0.020)
<i>Robot</i>		-0.028 (0.028)
<i>JIE</i>		0.082 * (0.049)
常数项	8.547 *** (0.028)	-22.601 *** (7.619)
城市固定效应	控制	控制
年份固定效应	控制	控制
R <sup>2</sup>	0.941	0.945
观测值	978	978

注：括号内为城市层面聚类稳健标准误；\*、\*\*和\*\*\*分别代表10%、5%和1%的显著性水平。

资料来源：根据前程无忧、BOSS直聘、智联招聘、猎聘网、拉勾网五大在线招聘网站、复旦大学数字与移动治理实验室、国家知识产权局专利信息、地级市政府工作报告和《中国城市统计年鉴》数据计算得到。

## 2. 开放数据量级与政策效果

公共数据开放的就业机会创造效应与开放数据量级密切相关。从理论上讲，政府公开的数据量越大、质量越高时，公共数据开放的就业机会创造效应会更加显著。复旦大学数字与移动治理实验室构建了数据层指数，该指数包含数据数量、开放范围、关键数据集质量、关键数据集规范五个一级指标，能够较好地反映政府开放的数据数量和质量。本文按照数据层指数高低将基准样本分为两组进行重新估计，估计结果见表4。从表中的估计结果看，在开放数据量级较大的组中，公共数据开放的估计系数无论大小和显著性均高于开放数据量级较小的组，并且两组之间的估计系数差异通过了组间估计系数差异性检验，p 值为 0.020<sup>①</sup>。这一结果表明，公共数据开放增加就业机会的效果在开放数据量级更高的组中更加显著，进一步强化了本文的基准结论。

表 4 开放数据量级差异

	开放数据量级较大	开放数据量级较小
	(1)	(2)
<i>Dataopen</i>	0.520 *** (0.146)	0.173 * (0.102)
控制变量	控制	控制
城市固定效应	控制	控制
年份固定效应	控制	控制
R <sup>2</sup>	0.940	0.952
观测值	467	511

注：括号内为城市层面聚类稳健标准误；\*、\*\* 和 \*\*\* 分别代表 10%、5% 和 1% 的显著性水平。

资料来源：根据前程无忧、BOSS 直聘、智联招聘、猎聘网、拉勾网五大在线招聘网站、复旦大学数字与移动治理实验室、国家知识产权局专利信息、地级市政府工作报告和《中国城市统计年鉴》数据计算得到。

## (二) 稳健性检验

### 1. 考虑异质性处理效应的平行趋势检验

不少文献指出 (Goodman-Bacon, 2021)，交叠 DID 面临异质性处理效应的挑战。在本文的情境下，由于各城市上线公共数据开放平台的时点不同，使用传统双向固定效应模型得到的估计量可能会产生偏误。针对这一问题，Callaway & Sant’Anna (2021) 建议使用交叠 DID 稳健估计方法 (CSDID) 进行动态趋势检验，以获取更为可靠的动

① 限于篇幅未汇报，后文同。备索。

态趋势检验结果。图 3 报告了 CSDID 动态趋势检验的估计结果。可以发现，在考虑一系列影响因素之后，在样本期内上线公共数据开放平台的城市，在上线年份之前与未上线城市的在线招聘需求并没有表现出显著差异，满足平行趋势假设的要求。同时，在上线年份之后，上线公共数据开放平台的城市的在线招聘需求开始逐渐显著高于未上线的城市，验证了本文基准结论的稳健性。

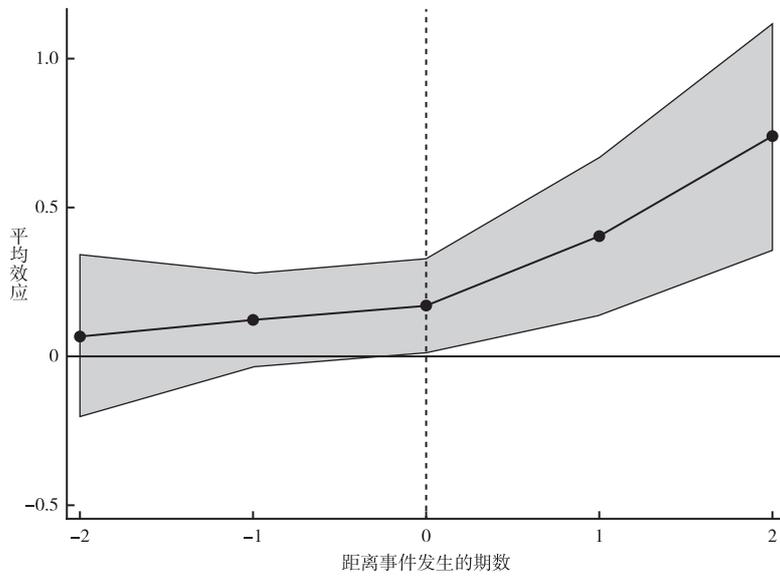


图 3 平行趋势检验

资料来源：根据前程无忧、BOSS 直聘、智联招聘、猎聘网、拉勾网五大在线招聘网站、复旦大学数字与移动治理实验室、国家知识产权局专利信息、地级以上政府工作报告和《中国城市统计年鉴》数据计算得到。

## 2. 安慰剂检验

在考察公共数据开放对就业机会的影响时，除一系列可观测城市特征外，还可能存在一些其他不可观测因素对估计结果产生影响。为缓解这一担忧，本文采用随机化的思路进行安慰剂检验。具体操作如下：本文重复 500 次实验，每次实验随机抽取与真实的公共数据开放平台上线城市等数量的城市作为处理组，并且随机给出处理时间，由此得到 500 个虚拟的核心解释变量  $Dataopen^{random}$ ；在此基础上，采用基准模型对这些虚拟的核心解释变量进行重新估计，得到 500 个  $Dataopen^{random}$  的估计系数，相应的核密度曲线和 p 值分布见图 4。

理论上，由于公共数据开放平台上线城市和上线时间都是随机赋予的，核心解释

变量  $Dataopen^{random}$  的估计系数应该和零没有差异。图 4 的结果显示，随机生成的核心解释变量  $Dataopen^{random}$  的估计系数确实主要集中在 0 附近，与真实的核心解释变量的估计系数 0.324 相比有明显差距。这一结果表明，在考虑一些其他不可观测因素的影响之后，公共数据开放增加就业机会的效果仍然保持不变。

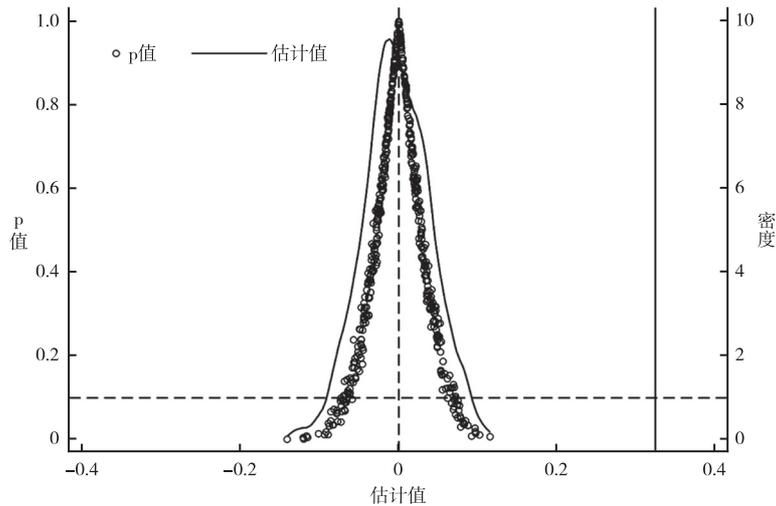


图 4 安慰剂检验

资料来源：根据前程无忧、BOSS 直聘、智联招聘、猎聘网、拉勾网五大在线招聘网站、复旦大学数字与移动治理实验室、国家知识产权局专利信息、地级市政府工作报告和《中国城市统计年鉴》数据计算得到。

### 3. 工具变量回归

公共数据开放与就业机会之间可能存在潜在的内生性问题。一方面，一些城市基于特定发展需求或政策导向可能会更早或更多地上线公共数据开放平台，而这些因素可能同时对城市就业机会产生影响。另一方面，就业机会更多的城市对公共数据开放的需求也可能更大。前述因素均可能导致估计结果出现偏误。

为了更准确地评估公共数据开放对就业机会的影响，本文采用工具变量法进行检验。具体来说，本文借鉴赵涛等（2020）的研究思路，引入地形起伏度作为工具变量。通常而言，地形起伏度越小，网络基础设施铺设成本越低，进而公共数据开放程度越高，从而满足相关性要求。同时，地形起伏度完全由经济系统外部因素决定，从而满足外生性要求。考虑到地形起伏度属于截面数据，在双向固定效应设定下无法作为公共数据开放的工具变量，本文引入了人均光缆长度变量。一般来说，人均光缆长度越

长，网络基础设施质量越高，从而公共数据开放程度越高。由此，本文以地形起伏度倒数与人均光缆长度的交乘项作为公共数据开放的工具变量。

表 5 报告了两阶段工具变量法的估计结果。第 (1) 列是第一阶段的估计结果，从估计结果看，地形起伏度倒数与人均光缆长度交乘项  $IV$  的估计系数在 1% 的水平下显著为正，表明本文构建的工具变量确实与公共数据开放正相关。第 (2) 列第二阶段的估计结果显示， $Dataopen_{IV}$  的估计系数显著为正，表明在工具变量法下公共数据开放仍然显著促进就业机会的增加。工具变量的 Cragg-Donald Wald F 值大于经验法则 10，通过了弱工具变量检验，并且 Kleibergen-Paap rk LM 值也通过了不可识别检验，说明本文构建的工具变量是有效的。表 6 的结果表明，在进一步考量内生性问题之后，公共数据开放增加就业机会的基准结论没有发生变化。

表 5 工具变量回归结果

	第一阶段	第二阶段
	(1)	(2)
$IV$	59.688 ** (26.637)	
$Dataopen_{IV}$		1.641 * (0.904)
控制变量	控制	控制
城市固定效应	控制	控制
年份固定效应	控制	控制
$R^2$	0.289	0.568
观测值	978	978
Cragg-Donald Wald F	12.233	
Kleibergen-Paap rk LM	10.628 ***	

注：括号内为城市层面聚类稳健标准误；\*、\*\* 和 \*\*\* 分别代表 10%、5% 和 1% 的显著性水平。

资料来源：根据前程无忧、BOSS 直聘、智联招聘、猎聘网、拉勾网五大在线招聘网站、复旦大学数字与移动治理实验室、国家知识产权局专利信息、地级市政府工作报告和《中国城市统计年鉴》数据计算得到。

#### 4. 排除相关政策的影响

在本文研究覆盖的样本期间附近，国家相继推出了若干与公共数据开放平台密切相关的政策，具体包括电子政务、国家大数据综合试验区、智慧城市以及宽带中国等试点项目。为了更准确地识别公共数据开放对就业机会的影响，本文逐一排除了上述政策的影响。

首先，排除国家电子政务综合试点政策的影响。公共数据开放与电子政务在数字化技术运用、提升政府工作效率与服务质量，以及促进政府数字化转型等多个方面展

现出一定的共同点，均强调信息公开和透明的重要性，均有助于增进公众对政府工作的认知与监督。因此，本文在样本中剔除了国家电子政务综合试点城市，以得到更为“干净”的公共数据开放就业机会创造效应。

其次，公共数据开放与国家级大数据综合试验区虽然都致力于推动大数据的发展并贯彻《促进大数据发展行动纲要》，但二者的具体目标并不完全相同。大数据综合试验区更注重推动产业升级和经济转型，而公共数据开放则聚焦于政府数据的开放与应用。为更准确地评估公共数据开放对就业机会的影响，本文剔除了国家大数据综合试验区试点的城市样本。

此外，公共数据开放与智慧城市试点政策都致力于提供新型城市数字公共服务，推动城市数字化转型。两者都强调信息共享和互操作的重要性，以实现更高效的城市管理和公共服务，但智慧城市更侧重于城市整体数字化建设。为了排除智慧城市可能对公共数据开放的就业机会创造效应的影响，本文剔除智慧城市试点的城市样本。

最后，公共数据开放与宽带中国试点都是城市数字基础设施建设中的重要环节，对于提高城市信息化水平和促进数字经济发展具有重要意义。考虑到宽带中国试点可能对公共数据开放产生的影响，本文剔除宽带中国试点的城市样本。

尽管这些政策并没有发生在本文的样本期内，但政策的长期效应可能对本文的研究结论产生影响。为排除这些政策的干扰，本文参考 Cao & Chen (2022) 的做法，剔除受到这些政策影响的样本。表 6 第 (1) 列至第 (4) 列依次是剔除电子政务试点城市、国家大数据综合试验区试点城市、智慧城市试点和宽带中国试点城市的估计结果<sup>①</sup>。从估计结果上看，在排除这些相关政策的影响之后，公共数据开放增加就业机会的效应仍然保持不变，再次说明本文核心结论的稳健性。

表 6 排除相关政策的影响

	剔除电子政务 试点城市样本	剔除国家大数据综合 试验区试点城市样本	剔除智慧城市 试点样本	剔除宽带中国试点 城市样本
	(1)	(2)	(3)	(4)
<i>Dataopen</i>	0.338 *** (0.099)	0.300 *** (0.104)	0.244 ** (0.100)	0.300 ** (0.132)
控制变量	控制	控制	控制	控制

① 如果同时剔除所有四种政策试点的城市，将损失约 3/4 的样本。因此，本文未采用这一做法。

续表

	剔除电子政务 试点城市样本	剔除国家大数据综合 试验区试点城市样本	剔除智慧城市 试点样本	剔除宽带中国试点 城市样本
	(1)	(2)	(3)	(4)
城市固定效应	控制	控制	控制	控制
年份固定效应	控制	控制	控制	控制
R <sup>2</sup>	0.929	0.945	0.954	0.926
观测值	778	728	623	632

注：括号内为城市层面聚类稳健标准误；\*、\*\*和\*\*\*分别代表10%、5%和1%的显著性水平。

资料来源：根据前程无忧、BOSS直聘、智联招聘、猎聘网、拉勾网五大在线招聘网站、复旦大学数字与移动治理实验室、国家知识产权局专利信息、地级市政府工作报告和《中国城市统计年鉴》数据计算得到。

### 5. 其他稳健性检验

为了确保研究结论的稳健性，本文还进行了以下检验：表7第(1)列排除了在线招聘岗位数低于1000的样本，以提高在线招聘需求对全部就业机会的代表性。经此处理后，公共数据开放显著增加就业机会的结论仍然保持不变。第(2)列进一步考虑同省观测值之间潜在的关联性问题，通过聚类到更高层级的省份层面减少可能存在的估计误差。从估计结果看，公共数据开放的估计系数仍然在1%的水平下显著为正，进一步证明公共数据开放对就业机会的积极影响。第(3)列删除了公共数据开放平台上线前一年的样本，以排除预期效应的影响。可以发现，公共数据开放促进就业机会增加的效果仍然存在。综合来看，本文通过多方面的稳健性检验进一步证实公共数据开放对就业机会的促进效果。

表7 其他稳健性检验

	剔除岗位数低于1000样本	聚类到省份层面	删除平台上线前一年样本
	(1)	(2)	(3)
<i>Dataopen</i>	0.244 *** (0.080)	0.324 *** (0.094)	0.291 *** (0.082)
控制变量	控制	控制	控制
城市固定效应	控制	控制	控制
年份固定效应	控制	控制	控制
R <sup>2</sup>	0.938	0.945	0.947
观测值	810	978	884

注：除第(2)列外，其他列括号内为城市层面聚类稳健标准误；\*、\*\*和\*\*\*分别代表10%、5%和1%的显著性水平。

资料来源：根据前程无忧、BOSS直聘、智联招聘、猎聘网、拉勾网五大在线招聘网站、复旦大学数字与移动治理实验室、国家知识产权局专利信息、地级市政府工作报告和《中国城市统计年鉴》数据计算得到。

## 六 进一步研究

### （一）机制检验

在理论分析部分，本文认为公共数据开放可能通过促进创业活跃度、数实融合发展 and 人口集聚三大机制促进就业机会增长。本部分基于全量工商登记注册信息、全量专利信息 and 人口吸引力指数对这三种机制进行实证检验。

#### 1. 创业活跃度机制

根据前文的理论分析，公共数据开放可以缓解信息不对称，为创业者提供更多创业机会，并降低创业门槛。随着创业活动的增多，企业的在线招聘需求也会相应增加。为检验这一机制，本文搜集了全量的工商登记注册信息，并基于此构建城市 - 年份层面的净新注册企业数量指标，用 *Firm* 表示。这一指标的具体定义为：对数化的新注册企业数量减去当期注销企业数量。估计结果见表 8 第（1）列。从估计结果中可以发现，公共数据开放的估计系数为正，且在 5% 的水平下显著，表明公共数据开放显著提升创业活跃度，验证了创业活跃度机制。

#### 2. 数实融合机制

前文理论分析部分认为，数实融合通过催生新兴业态与岗位、促进传统行业数字化转型及产业链延伸显著增加在线招聘需求，而公共数据开放能够打破信息隔阂，为数实融合提供数据支撑与动力。为检验这一机制，本文参考周密等（2024）的思路，通过计算传统产业包含数字产业技术的专利数量构建数实融合指标，用 *DigPhyint* 表示。表 8 第（2）列检验了公共数据开放对数实融合的影响，可以发现，公共数据开放的估计系数在 1% 的水平下显著为正，表明公共数据开放确实显著推动数实融合发展，数实融合机制得以验证。

#### 3. 人口集聚机制

根据理论分析部分，公共数据开放有助于流动人口更高效地选择居住和工作地，促进合理流动和集聚，进而对在线招聘需求产生影响。为衡量这一机制变量，本文采用人口吸引力指数（城市新流入常住人口（驻留超过 2 个月的人口）/全国所有城市新流入常住人口均值）作为代理指标，用 *Laboragg* 表示。这一指数是通过对百度开放平台海量位置数据、百度地图海量交通出行数据进行挖掘计算得到<sup>①</sup>。表 8 第（3）列报

<sup>①</sup> 参见百度地图慧眼、自然资源部陆表系统与入地关系重点实验室发布的《中国城市活力报告》。需要说明的是，表 8 的观测值少于表 3，原因在于部分年份和城市的数据缺失。

告了人口集聚机制的估计结果。从估计结果上看，公共数据开放的估计系数显著为正，表明公共数据开放确实显著促进了人口集聚，从而验证了人口集聚机制。

表 8 机制检验

	<i>Firm</i>	<i>DigPhyint</i>	<i>Laboragg</i>
	(1)	(2)	(3)
<i>Dataopen</i>	0.088 ** (0.037)	0.174 *** (0.064)	0.043 ** (0.017)
控制变量	控制	控制	控制
城市固定效应	控制	控制	控制
年份固定效应	控制	控制	控制
R <sup>2</sup>	0.951	0.925	0.990
观测值	978	978	651

注：括号内为城市层面聚类稳健标准误；\*、\*\* 和 \*\*\* 分别代表 10%、5% 和 1% 的显著性水平。

资料来源：根据前程无忧、BOSS 直聘、智联招聘、猎聘网、拉勾网五大在线招聘网站、复旦大学数字与移动治理实验室、国家知识产权局专利信息、地级市政府工作报告、《中国城市统计年鉴》、各城市市场监督管理局网站和《中国城市活力报告》数据计算得到。

## (二) 不同类型就业机会的差异性影响

接下来，本文进一步考察公共数据开放对不同类型在线招聘需求的差异性影响。值得一提的是，不同于已有的家户微观调查数据，本文的在线招聘数据具有较高的颗粒度，在分城市统计不同类型在线招聘需求时仍然具有较高的代表性。

### 1. 不同岗位就业机会的差异性影响

本文分别统计了不同学历、有无经验要求和不同岗位工资的在线招聘需求数量，以考察公共数据开放对不同岗位就业机会的影响是否存在差异。具体来说，本文将高中及以上记为高学历在线招聘需求 (*Jobs\_hedu*)，将高中以下记为低学历在线招聘需求 (*Jobs\_ledu*)；将有经验要求的记为有经验在线招聘需求 (*Jobs\_exp*)，将无经验要求的记为无经验在线招聘需求 (*Jobs\_nonexp*)；将岗位平均工资及以上的记为高岗位工资在线招聘需求 (*Jobs\_hwage*)，将低于岗位平均工资的记为低岗位工资在线招聘需求 (*Jobs\_lwage*)。

从表 9 的结果看，公共数据开放对不同学历、有无经验要求、不同岗位工资的就业机会均具有显著的正向影响，说明公共数据开放增加就业机会的核心结论具有稳健性。更为重要的是，比较第 (1) 列和第 (2) 列的估计结果，公共数据开放对低学历在线招聘需求的影响高于对高学历在线招聘需求的影响。第 (3) 列和第 (4) 列的结

果显示，相比有经验要求的在线招聘需求，公共数据开放对无经验要求在线招聘需求的影响更大。对比第（5）列和第（6）列的结果发现，公共数据开放主要增加低岗位工资在线招聘需求。整体上看，公共数据开放体现出良好的包容性和均等性，更多地为弱势就业群体创造新的就业机会。

表 9 不同岗位就业机会的差异性影响

	<i>Jobs_hedu</i>	<i>Jobs_ledu</i>	<i>Jobs_exp</i>	<i>Jobs_nonexp</i>	<i>Jobs_hwage</i>	<i>Jobs_lwage</i>
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Dataopen</i>	0.335 *** (0.070)	0.342 *** (0.095)	0.264 ** (0.114)	0.353 *** (0.065)	0.224 ** (0.095)	0.353 *** (0.064)
控制变量	控制	控制	控制	控制	控制	控制
城市固定效应	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
R <sup>2</sup>	0.958	0.937	0.924	0.963	0.926	0.959
观测值	978	978	978	978	978	978

注：括号内为城市层面聚类稳健标准误；\*、\*\*和\*\*\*分别代表10%、5%和1%的显著性水平。

资料来源：根据前程无忧、BOSS直聘、智联招聘、猎聘网、拉勾网五大在线招聘网站、复旦大学数字与移动治理实验室、国家知识产权局专利信息、地级市政府工作报告和《中国城市统计年鉴》数据计算得到。

## 2. 不同行业就业机会的差异性影响

本文分别统计了不同行业的在线招聘需求数量，以考察公共数据开放对不同行业就业机会的影响。具体来说，本文将制造业记为制造业在线招聘需求 (*Jobs\_manu*)，将服务业记为服务业在线招聘需求 (*Jobs\_service*)；将数字行业记为数字行业在线招聘需求 (*Jobs\_dig*)，将余下行业记为传统行业在线招聘需求 (*Jobs\_trad*)；将劳动力密集行业记为劳动力密集行业在线招聘需求 (*Jobs\_labor*)，将资本密集型行业记为资本密集型行业在线招聘需求 (*Jobs\_capital*)，将技术密集型行业记为技术密集型行业在线招聘需求 (*Jobs\_tech*)。

从表10和表11的结果来看，公共数据开放对不同行业的就业机会均具有显著的正向影响，说明公共数据开放增加就业机会的核心结论具有稳健性。比较表10第(1)列和第(2)列的估计结果，公共数据开放对服务业在线招聘需求的影响高于对制造业在线招聘需求的影响。其原因在于，不同行业对数据的需求因其特性而异，服务业更注重个性化、定制化，对数据的多元性需求相对较高，而制造业则更注重生产流程和成本控制，对数据的需求相对较少。比较表10第(3)列和第(4)列的

估计结果发现，相比传统行业在线招聘需求，公共数据开放对数字行业在线招聘需求的影响更大。这是因为数字行业对数据的依赖程度较高，公共数据开放为其提供了丰富的数据资源。这些数据资源有助于数字行业进行更深入的数据挖掘、分析和应用，从而增加对数据科学家、数据分析师等数字行业从业者的需求。对比表 11 第 (1) 列至第 (3) 列的估计结果，公共数据开放的就业机会创造效应在劳动力密集型行业中最大，其次是资本密集型行业，最后是技术密集型行业。这是因为劳动力密集型行业对劳动力的吸纳能力最强，资本密集型行业次之，技术密集型行业最弱。综合来看，公共数据开放主要增加与数据要素密切相关的数字行业、就业吸纳能力较强的劳动力密集型行业和服务业的在线招聘需求，进一步强化了公共数据开放的就业机会创造效应的核心结论。

表 10 不同行业就业机会的差异性影响 (一)

	<i>Jobs_manu</i>	<i>Jobs_service</i>	<i>Jobs_dig</i>	<i>Jobs_trad</i>
	(1)	(2)	(3)	(4)
<i>Dataopen</i>	0.240 *** (0.074)	0.354 *** (0.093)	0.382 *** (0.118)	0.357 *** (0.089)
控制变量	控制	控制	控制	控制
城市固定效应	控制	控制	控制	控制
年份固定效应	控制	控制	控制	控制
R <sup>2</sup>	0.942	0.938	0.926	0.942
观测值	978	978	978	978

注：括号内为城市层面聚类稳健标准误；\*、\*\* 和 \*\*\* 分别代表 10%、5% 和 1% 的显著性水平。

资料来源：根据前程无忧、BOSS 直聘、智联招聘、猎聘网、拉勾网五大在线招聘网站、复旦大学数字与移动治理实验室、国家知识产权局专利信息、地级市政府工作报告和《中国城市统计年鉴》数据计算得到。

表 11 不同行业就业机会的差异性影响 (二)

	<i>Jobs_labor</i>	<i>Jobs_capital</i>	<i>Jobs_tech</i>
	(1)	(2)	(3)
<i>Dataopen</i>	0.375 *** (0.091)	0.357 *** (0.097)	0.275 *** (0.085)
控制变量	控制	控制	控制
城市固定效应	控制	控制	控制
年份固定效应	控制	控制	控制

续表

	<i>Jobs_labor</i>	<i>Jobs_capital</i>	<i>Jobs_tech</i>
	(1)	(2)	(3)
R <sup>2</sup>	0.926	0.937	0.950
观测值	978	978	978

注：括号内为城市层面聚类稳健标准误；\*、\*\*和\*\*\*分别代表10%、5%和1%的显著性水平。

资料来源：根据前程无忧、BOSS直聘、智联招聘、猎聘网、拉勾网五大在线招聘网站、复旦大学数字与移动治理实验室、国家知识产权局专利信息、地级市政府工作报告和《中国城市统计年鉴》数据计算得到。

### (三) 异质性分析

#### 1. 传统要素流通异质性

公共数据开放有助于提升传统要素的流动性，打破信息壁垒，提高资源配置效率，进而释放更大的生产力。因此，本文预期，公共数据开放促进就业机会增加的效果在传统要素流动性较低的地区更加显著。为验证这一观点，本文参考方锦程等（2023）的测度方法，使用历年公路客运量和公路货运量作为人口流动和商品流动的代理变量。本文以这两个变量的中位数为标准，将基准样本分为高低两组，重新估计公共数据开放对就业机会的影响，结果见表12。可以看出，在公路客运量和公路货运量较少的地区，公共数据开放的估计系数在1%的水平下均显著为正。而在公路客运量和公路货运量较多的地区，这一估计系数则均不显著，并且组间估计系数差异显著，即第（1）列和第（2）列之间、第（3）列和第（4）列之间估计系数差异的p值均为0.000。这表明公共数据开放对在线招聘需求的促进作用确实传统要素流动性较低的地区更明显，符合本文的预期。

表12 传统要素流通异质性

	公路客运量较多	公路客运量较少	公路货运量较多	公路货运量较少
	(1)	(2)	(3)	(4)
<i>Dataopen</i>	-0.017 (0.105)	0.682 *** (0.155)	-0.020 (0.108)	0.606 *** (0.151)
控制变量	控制	控制	控制	控制
城市固定效应	控制	控制	控制	控制
年份固定效应	控制	控制	控制	控制
R <sup>2</sup>	0.956	0.927	0.955	0.903
观测值	464	395	462	430

注：括号内为城市层面聚类稳健标准误；\*、\*\*和\*\*\*分别代表10%、5%和1%的显著性水平。

资料来源：根据前程无忧、BOSS直聘、智联招聘、猎聘网、拉勾网五大在线招聘网站、复旦大学数字与移动治理实验室、国家知识产权局专利信息、地级市政府工作报告和《中国城市统计年鉴》数据计算得到。

## 2. 市场化程度异质性

公共数据开放对就业机会的影响与市场化程度相关。本文参考方福前等（2023）的做法，根据樊纲的市场化指数构建方法测算城市层面的市场化程度，并按照市场化程度高低将基准样本划分为两组进行重新评估，估计结果见表 13。从表中的估计结果来看，公共数据开放的估计系数在市场化程度较低的组中显著为正，而在市场化程度较高的组中则不显著，并且两组之间的估计系数差异通过了组间估计系数差异性检验， $p$  值为 0.015。这表明在市场化程度较低的地区，公共数据开放对在线招聘需求的增加作用更明显。这是因为公共数据开放有助于激活市场、促进非国有经济发展、推动产品市场发展、提高要素市场透明度和效率，以及促进市场中介发展，从而对市场化程度较低地区的就业机会产生更加显著的积极影响。

表 13 市场化程度异质性

	市场化程度较高	市场化程度较低
	(1)	(2)
<i>Dataopen</i>	0.108 (0.098)	0.453 *** (0.166)
控制变量	控制	控制
城市固定效应	控制	控制
年份固定效应	控制	控制
$R^2$	0.964	0.932
观测值	475	455

注：括号内为城市层面聚类稳健标准误；\*、\*\* 和 \*\*\* 分别代表 10%、5% 和 1% 的显著性水平。

资料来源：根据前程无忧、BOSS 直聘、智联招聘、猎聘网、拉勾网五大在线招聘网站、复旦大学数字与移动治理实验室、国家知识产权局专利信息、地级市政府工作报告和《中国城市统计年鉴》数据计算得到。

## 3. 信息可达性异质性

信息的可获取性和利用程度在决策过程中起着至关重要的作用，直接影响决策的质量和效果，进而影响公共数据开放对就业机会创造的程度。因此，本文预测在信息可达性较高的地区，公共数据开放的就业机会创造效应更加显著。互联网宽带接入作为信息传递的关键要素，其用户规模的大小直接反映信息可达性的高低。用户规模越大，意味着更多的人能够接入互联网获取信息，从而具有更高的信息可获取性和利用程度。因此，本文利用互联网宽带接入用户规模衡量信息可达性，并按照用户规模高低将基准样本划分为两组进行重新估计，估计结果见表 14。从表中的估计结果来看，

公共数据开放的估计系数在互联网宽带接入用户规模较大的组中显著为正，而在互联网宽带接入用户规模较小的组中则不显著，并且两组之间的估计系数差异通过了组间估计系数差异性检验， $p$  值为 0.090。这表明公共数据开放促进就业机会增加的效果确实在信息可达性更高的组中更加显著，符合本文的预期。

表 14 信息可达性异质性

	互联网宽带接入用户规模较大	互联网宽带接入用户规模较小
	(1)	(2)
<i>Dataopen</i>	0.383 ** (0.146)	0.136 (0.185)
控制变量	控制	控制
城市固定效应	控制	控制
年份固定效应	控制	控制
$R^2$	0.964	0.929
观测值	387	343

注：括号内为城市层面聚类稳健标准误；\*、\*\* 和 \*\*\* 分别代表 10%、5% 和 1% 的显著性水平。

资料来源：根据前程无忧、BOSS 直聘、智联招聘、猎聘网、拉勾网五大在线招聘网站、复旦大学数字与移动治理实验室、国家知识产权局专利信息、地级市政府工作报告和《中国城市统计年鉴》数据计算得到。

## 七 结论和政策建议

数据要素作为数字时代的关键动力，对于推动高质量发展和中国式现代化具有深远影响。在稳定和促进就业方面，数据要素同样发挥着举足轻重的作用。本文利用 6800 万在线招聘岗位数据构建在线招聘需求指标，将地方政府上线公共数据开放平台作为准自然实验，考察公共数据开放对就业机会的影响。研究发现，公共数据开放显著增加就业机会，且这一效果在开放数据量级较大的地区更加明显。机制分析表明，公共数据开放通过激发创业活力、推动数实融合发展以及促进人口集聚发挥作用。异质性结果显示，公共数据开放具有较好的普惠性，给弱势就业群体创造了更多就业机会，体现在主要增加低学历、没有工作经验、低工资岗位的在线招聘需求。同时，公共数据开放主要增加服务业、数字行业以及劳动和资本密集型行业的在线招聘需求。此外，公共数据开放的就业机会创造效应在传统要素流通程度较低、市场化程度较低、信息可达性较高的城市更加显著。本文的研究结论对于深入理解和评估政府公共数据开放的就业机会创造效应，发挥数据要素乘数效应和促进就业都有一定的启发意义。

基于这些研究结论，本文提出以下政策建议。

第一，积极推动公共数据开放，充分发挥数据要素在推动创业、促进数实融合、促进人口聚集和创造就业方面的积极作用。除了基础公共数据，各级政府还应逐步开放更多高价值的数据，如经济、教育、医疗等领域的数据。与此同时，应鼓励不同政府部门之间加强数据整合和共享，消除信息孤岛，形成统一的数据管理和服务平台，最大化数据要素的倍增效应。此外，对于传统要素流通和市场化程度较低的地区，应充分利用数字经济发展的红利和战略机遇，积极推动公共数据开放，从而缩小与其他地区的就业机会差距。

第二，充分发挥公共数据开放在促进社会公平就业中的重要作用。公共数据开放对低学历、无经验和低收入的弱势群体就业具有显著的积极效果。首先，政府应提供针对弱势就业群体的开放数据使用培训，帮助其掌握相关技能，提升就业竞争力。其次，政府应利用开放数据为弱势群体提供职业指导服务，分享市场需求、就业趋势和行业动态信息，帮助其做出更符合个人能力和市场需求的职业选择。

第三，信息基础设施对于公共数据开放所带来的就业创造效应至关重要，因此政府应加大对关键设施的投资力度，如宽带网络、数据中心和云计算，提升数据传输和处理效率。特别是对于偏远和欠发达地区，政府应重点关注其信息基础设施建设，确保所有地区均能公平享有数据开放带来的红利，从而提高社会整体的信息可达性。

## 参考文献：

- 蔡庆丰、王仕捷、刘昊、舒少文（2023），《城市群人口集聚促进域内企业创新吗》，《中国工业经济》第3期，第152-170页。
- 蔡卫星、韦庆芳、林航宇（2023），《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》，《金融研究》第10期，第28-46页。
- 陈晓红、李杨扬、宋丽洁、汪阳洁（2022），《数字经济理论体系与研究展望》，《管理世界》第2期，第208-224页。
- 方福前、田鸽、张勋（2023），《数字基础设施与代际收入向上流动性——基于“宽带中国”战略的准自然实验》，《经济研究》第5期，第79-97页。
- 方锦程、刘颖、高昊宇、董纪昌、吕本富（2023），《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》，《管理世界》第9期，第124 - 142页。
- 李磊、王小霞、包群（2021），《机器人的就业效应：机制与中国经验》，《管理世界》第9期，第104 - 119页。
- 宁光杰、崔慧敏、付伟豪（2023），《信息技术发展如何影响劳动力跨行业流动? ——基于工作任务与技能类型的实证研究》，《管理世界》第8期，第1 - 21页。
- 欧阳伊玲、王愉靖、李平、高昊宇（2024），《数据要素与城投债定价：基于公共数据开放的准自然实验》，《世界经济》第2期，第174 - 203页。
- 彭远怀（2023），《政府数据开放的价值创造作用：企业全要素生产率视角》，《数量经济技术经济研究》第9期，第50 - 70页。
- 王贤彬、陈春秀（2023），《重点产业政策与制造业就业》，《经济研究》第10期，第34 - 54页。
- 夏杰长、苏敏（2024），《以数实融合推动现代化产业体系建设》，《改革》第5期，第12 - 23页。
- 余玲铮、魏下海、孙中伟、吴春秀（2021），《工业机器人、工作任务与非常规能力溢价——来自制造业“企业 - 工人”匹配调查的证据》，《管理世界》第1期，第47 - 59页。
- 张晨、万相昱、姜智超、唐亮（2023），《开放政府数据的经济增长效应研究》，《中国软科学》第2期，第1 - 11页。
- 张叶青、陆瑶、李乐芸（2021），《大数据应用对中国企业市场价值的影响——来自中国上市公司年报文本分析的证据》，《经济研究》第12期，第42 - 59页。
- 赵涛、张智、梁上坤（2020），《数字经济、创业活跃度与高质量发展——来自中国城市的经验证据》，《管理世界》第10期，第65 - 76页。
- 周密、王雷、郭佳宏（2024），《新质生产力背景下数实融合的测算与时空比较——基于专利共分类方法的研究》，《数量经济技术经济研究》第7期，第5 - 27页。
- Bernal, Carolina, Mounu Prem, Juan Vargas & Mónica Ortiz (2024). Peaceful Entry: Entrepreneurship Dynamics During Colombia's Peace Agreement. *Journal of Development Economics*, 166, 103119.
- Brynjolfsson, Erik & Kristina McElheran (2016). The Rapid Adoption of Data-Driven Decision-Making. *American Economic Review*, 106 (5), 133 - 139.
- Callaway, Brantly & Pedro Sant'Anna (2021). Difference-in-Differences with Multiple Time

- Periods*. *Journal of Econometrics*, 225 (2), 200 – 230.
- Cao, Yiming & Shuo Chen (2022). Rebel on the Canal: Disrupted Trade Access and Social Conflict in China, 1650 – 1911. *American Economic Review*, 112 (5), 1555 – 1590.
- Di Addario, Sabrina (2011). Job Search in Thick Markets. *Journal of Urban Economics*, 69 (3), 303 – 318.
- Farboodi, Maryam, Roxana Mihet, Thomas Philippon & Laura Veldkamp (2019). Big Data and Firm Dynamics. *AEA Papers and Proceedings*, 109, 38 – 42.
- Feenstra, Robert, Hong Ma & Yuan Xu (2019). US Exports and Employment. *Journal of International Economics*, 120, 46 – 58.
- Ghasemaghaei, Maryam & Goran Calic (2019). Does Big Data Enhance Firm Innovation Competency? The Mediating Role of Data-Driven Insights. *Journal of Business Research*, 104, 69 – 84.
- Goodman-Bacon, Andrew (2021). Difference-in-Differences with Variation in Treatment Timing. *Journal of Econometrics*, 225 (2), 254 – 277.
- Hershbein, Brad & Lisa Kahn (2018). Do Recessions Accelerate Routine-Biased Technological Change? Evidence from Vacancy Postings. *American Economic Review*, 108 (7), 1737 – 1772.
- Huang, Mian, Chunbing Xing & Xiaoyong Cui (2022). Does College Location Affect the Location Choice of New College Graduates in China? *China & World Economy*, 30 (3), 135 – 160.
- Liu, Mengdi, Ruipeng Tan & Bing Zhang (2021). The Costs of “Blue Sky”: Environmental Regulation, Technology Upgrading, and Labor Demand in China. *Journal of Development Economics*, 150, 102610.
- Ring, Marius (2023). Entrepreneurial Wealth and Employment: Tracing Out the Effects of a Stock Market Crash. *The Journal of Finance*, 78 (6), 3343 – 3386.
- Samuelson, Paul (1954). The Pure Theory of Public Expenditure. *The Review of Economics and Statistics*, 36 (4), 387 – 389.
- Stiglitz, Joseph (2000). The Contributions of the Economics of Information to Twentieth Century Economics. *The Quarterly Journal of Economics*, 115 (4), 1441 – 1478.

## Public Data Openness and Employment Opportunities: Evidence from 68 Million Online Job Postings

Wei Qingfang<sup>1</sup>, Wei Xiahai<sup>2</sup>, Liu Yuan<sup>1</sup> & Yu Lingzheng<sup>2</sup>

(School of Finance, Guangdong University of Finance & Economics<sup>1</sup>;

School of Economics and Finance, Huaqiao University<sup>2</sup>)

**Abstract:** Orderly opening of public data and harnessing digital dividends are key measures to promote the development of the labor market. This article uses 68 million online job postings nationwide to develop an online recruitment demand index. It uses local governments' launch of public data platforms as a quasi-natural experiment to examine their impact on local employment opportunities. The study finds that opening public data increases local online recruitment demand significantly, with a more pronounced effect in regions with larger volumes of open data. The mechanism is that public data openness can invigorate entrepreneurship, promote the integration of digital and traditional economies, and foster population agglomeration, thereby creating employment opportunities. Moreover, the effects of public data openness fully demonstrate inclusivity and equality, as companies are more willing to offer positions to groups with lower education, no work experience, and lower income. Additionally, public data openness primarily boosts job demand in the service sector, digital industry, and labor- and capital-intensive industries. In cities with less traditional factor mobility, lower levels of marketization, and greater information accessibility, the job creation effect of public data openness is more pronounced. This study contributes to a deeper understanding of the relationship between public data openness and labor market development, providing valuable policy insights for leveraging data elements to create employment opportunities.

**Keywords:** data factors, public data openness, employment opportunities, online job posting

**JEL Classification:** J20, O30, H40

(责任编辑：崔慧敏)