

数据要素与企业产能利用率

——基于公共数据开放平台的准自然实验

李 晓 杜宜萱 于晓旭

摘要：数据是新的生产要素，是基础性资源和战略性资源，公共数据开发应用是激活数据要素潜能的关键。本文以我国各省（自治区、直辖市）陆续开放公共数据为准自然实验，利用2008~2022年A股上市公司样本，采用多期双重差分模型检验公共数据开放对企业产能利用率的影响。研究发现，公共数据开放显著提升了企业产能利用率，其中的作用机制是改善区域间要素资源错配（资本要素）、提升政府治理效率、增强企业创新能力，从而促使企业提高产能利用率。进一步研究发现，对于所在省份市场化程度较低的企业、数字化水平较低的企业，以及产业类型为劳动密集型企业，上述作用更加显著。本文丰富了公共数据资源赋能经济高质量发展的研究，为我国企业提高产能利用率、实现高质量发展提供了新的参考。

关键词：数据要素；公共数据开放；产能利用率；要素资源错配

中图分类号：F272；F420 **文献标识码：**A **文章编号：**2095-8838(2025)03-0054-18

一、引言

随着大数据、人工智能等技术的不断革新，企业生产运营、政府行政服务，乃至社会整体运行越来越离不开数据要素的使用。2019年10月，党的十九届四中全会首次将数据列为与劳动、资本、土地、知识、技术和管理并列的生产要素。2025年1月全国数据工作会议披露的数据显示，2024年我国数据要素市场交易规模预计超过1 600亿元，数据资源已成为我国经济高质量发展的源动力。数据资源可分为公共数据和私有数据，私有数据因版权和使用门槛等限制，目前其要素动能受限；而公共数据涵盖范围广、使用门槛较低、流动性较强，能够更好发挥数据要素动能。然而，目前数据要素在我国经济高质量发展过程中发挥的动能与其规模体量尚不匹配，由数据要素驱动的数字经济占国内生产总值的比重仍有提升空间。因此，公共数据开放是充分利用数据资源、有效发挥数据要素价值的重要举措。2012年，上海和北京等地率先上线公共数据开放平台，截至2024年7月，我国已有28个省（自治区、直辖市）上线了公共数据开放平台^①，向社会公众提供教育、交通、医疗、旅游、农业、气象、电力、企业信用信息等

收稿日期：2024-11-25

基金项目：国家自然科学基金面上项目（72472170）

作者简介：李 晓，中央财经大学会计学院副教授，博士生导师；

杜宜萱，中央财经大学会计学院硕士研究生；

于晓旭，中央财经大学会计学院助理教授。

① 数据来源于《中国地方公共数据开放利用报告——省域》。

多领域的的数据资源,便于公众及时获取并利用数据,推动数字经济发展。

公共数据资源与以往的政府信息资源存在显著差异(潘俊等,2023;吴武清等,2024)。一是范围不同。政府信息资源只针对行政机构,而公共数据资源将范围扩大到行政机构、具有公共管理职能的企事业单位等主体,数据类别更丰富、体量更大。二是面向对象不同。以往的政府信息资源多为行政机构内部共享的政务数据,并不直接向社会公众提供获取渠道,而公共数据资源则是将行政服务工作中产生的海量数据经过脱敏处理后,供公众获取使用。三是价值不同。政府信息资源多是经过人为解读、整理、汇总并编制形成的文本形式文件,例如统计报告、年度报告等,这种处理使得公众无法接触原始数据,数据要素的价值释放受限,而公共数据资源提供与社会公众利益相关的、原始的、一手的数据,便于公众充分利用。例如企业可以通过程序接口与公共数据开放平台间实现数据对接,利用工商、司法、交通等领域的公共数据及时调整生产经营决策,推动产业链良性循环。具体而言,在工商领域,企业可以通过公共数据平台获取行业内企业的注册信息、经营范围、股东结构等信息,从而分析市场竞争格局,制定或调整市场进入策略;在司法领域,企业可以获取法律判决、失信被执行人、知识产权侵权等信息,有助于企业规避法律风险,保护自身合法权益;在交通领域,企业可以获取交通流量、道路状况等数据,这些数据对于物流、出行服务等行业企业尤为重要,有助于其优化运输路线和充电方案、提高配送效率、降低运营成本。以山东、福建两省的公共数据开放为例,山东省建成了全国首个海洋数据交易平台,开展海洋地质、地形地貌、水文气象、遥感影像等海洋数据的交易,有助于海洋能源业、渔业、船舶工业等行业企业利用数据合理安排资源的探测开采、船舶的维修改装与海洋通信导航技术的升级;福建省利用公共数据资源搭建了福建省金融服务云平台,调用市场监管、税务、工商等17个定制类数据接口,在一定程度上帮助中小微企业降低综合融资成本。相较于以往的政府信息公开,公共数据资源开放更侧重数据的原始性、未加工性,且面向对象为全体社会公众,有利于激发数据要素动能,创造经济价值。

随着我国经济发展由量变到质变的转变,如何有效提高产能利用率已成为近年来政策部门和学术界重点关注、亟待解决的问题。现有关于产能利用率影响因素的研究主要从宏观经济与微观企业行为两方面展开。宏观方面,学者分别探讨了地方政府税收激励(杨龙见等,2019)、财政政策扩张(郭长林,2016)、对外投资(李雪松等,2017;曹献飞和裴平,2019)、地方政府官员变更(徐业坤和马光源,2019)、区域市场整合(卞元超和白俊红,2021)、环境规制(陈林等,2024)等对产能利用率的影响;微观方面学者则从数字化转型(韩国高等,2022)、股权结构(马新啸等,2021)等角度进行了研究。然而,公共数据这一重要生产要素对于企业提升产能利用率是否有推动作用尚未引起广泛关注。如果公共数据能够被有效利用,可能有助于企业管理层优化资源配置、提高外部利益相关者的监督效率,从而提升企业产能利用率。因此,本文的研究立足于公共数据资源的有效利用,探究企业产能利用率的提升机制。

本文以2008~2022年A股上市公司为样本,使用多期双重差分模型进行实证分析。研究发现,公共数据开放地区企业的产能利用率更高,表明公共数据开放对我国企业产能利用率具有显著提升作用。经过平行趋势检验、安慰剂检验、工具变量法、PSM-DID(倾向得分匹配)、更换变量衡量方式、排除同期其他政策干扰等一系列稳健性测试后,本文的结论依然显著。机制检验表明,公共数据开放通过改善区域间要素资源错配(资本要素)、提升政府治理效率、增强企业创新能力进而提高产能利用率。对于所在省份市场化程度较低的企业、数字化水平低的企业,以及产业类型为劳动密集型企业,公共数据开放对产能利用率的提升作用更显著。

本文的边际贡献主要在于:第一,为公共数据资源赋能我国实体经济高质量发展提供了新的经验证据。现有关于公共数据开放的经济后果研究有降低城投债发行利差(欧阳伊玲等,2024)、降低审计定价(潘俊等,2023),以及促进企业创新(陈艳利和蒋琪,2024)、提高企业全要素生产率(杨秀云和韩奇,2023;吴武清等,2024)等,关于促进产能利用率提升的效应目前只有王海等(2024)进行了研究,但其研究重点在于公共数据开放如何提振企业有效投资,而本文则聚焦于公共数据开放对企业产能利用率本身的影响。第二,为我国企业提高产能利用率提供了新的启示。已有研究揭示了地方官员变更(徐业坤和马光源,2019)、智能制造政策试点(黄卓等,2024)、人力资本

扩张(方森辉和毛其淋, 2021)、数字化转型(韩国高等, 2022)等与我国企业产能利用之间的关系,而本文的研究发现,公共数据开放能够有效提高企业产能利用率,为我国企业提高产能利用率提供了新的参考。第三,此前关于公共数据开放作用渠道的研究多针对其信息效应,即改善信息不对称程度,而鲜有针对数据本身价值进行挖掘,本文则聚焦于数据要素自身的价值,揭示了公共数据开放对企业的积极影响,为充分挖掘数据要素价值提供了更多实证证据。

二、制度背景、理论分析与研究假设

(一)制度背景

近年来,数字经济逐渐成为实现高质量发展的重要驱动力量,而数据要素是数字经济深入发展的核心引擎。为有效发挥数据要素价值,我国已出台多个政策对各层级的公共数据开放作出指引。2015年国务院印发了《促进大数据发展行动纲要》和《关于促进云计算创新发展培育信息产业新业态的意见》,指出我国在大数据发展和应用方面已具备一定基础,拥有市场优势和发展潜力,但也存在政府数据开放共享不足、产业基础薄弱等问题。强调要加快政府数据开放共享,推动资源整合,提升治理能力。2020年《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,推进政府数据开放共享,加快推动各地区各部门间数据共享交换。2021年《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》中明确要求,探索将公共数据服务纳入公共服务体系,构建统一的国家公共数据开放平台和开发利用端口,优先推动企业登记监管、卫生、交通、气象等高价值数据集向社会开放。2012年上海市和北京市率先试点上线了公共数据开放平台,随后浙江省、广东省和贵州省也陆续上线,带动其余省份加快建设公共数据开放平台,2018~2020年期间,省级公共数据开放平台数量快速增长,内嵌的数据接口与数据集也在不断完善、增多。

以北京市公共数据开放平台为例,该平台于2012年正式上线。截至目前,已内嵌115个具有公共管理职能的机关单位,包括市发展改革委、市经济和信息化局、市司法局、市财政局、市统计局、市金融监管局等。平台数据集搜索热词主要包括“行政处罚信息”“事项清单”“法人单位”“规范性文件”“监督检查”“备案信息”等,多为与企业信用信息和监管信息相关的关键词。这表明,平台向社会公众提供了大量关于企业资质、信用、经营状况等方面的数据,有助于公共监督和对数据要素的充分挖掘使用,推动数据资源有效发挥其经济和社会价值。

(二)理论分析与研究假设

现有文献在宏观经济政策与微观企业决策两方面对企业产能利用率的影响进行了大量研究,其中属于数字经济范畴的文献主要集中于宏观层面的数字经济发展和微观企业层面的数字化转型。例如,郑雪梅等(2023)发现,数字经济作为城市经济发展新动能,能够有效提升东部及中部地区规模较小民营企业的产能利用率;罗奇等(2022)认为,“宽带中国”战略下的数字基础设施建设显著提升了企业产能利用率,且这种作用在长期内会逐步增强;韩国高等(2022)及孙帆等(2022)发现,企业数字化转型能够缓解产能过剩、提高企业产能利用率,并可能对同行业企业产生溢出效应;朱国悦和陶锋(2024)研究表明,供应链数字化能够显著提升企业产能利用率,特别是在企业间供应链关系紧密度较高、地区供应链运营环境较好的企业中,这种提升作用更加明显。公共数据开放与以往研究中的数字经济联系密切,但其政策含义和实践意义并不等同,二者所带来的经济后果及内在作用机制也存在差异。现有研究大多认为,数字经济对于产能利用率的影响主要通过缓解信息不对称、加强监督管理以及促进技术创新、提高生产效率等渠道实现,强调数字化的“信息性”,而公共数据开放则强调数据资源的原始性、一手性和低门槛性,而非经过人为解读、加工后的信息资源,因此与以往的数字化研究存在本质区别。

公共数据开放后,各社会主体可以低门槛甚至无门槛获取并使用海量数据资源(方锦程等, 2023),企业管理层能够获取各类数据,根据宏观经济、市场、产业及供应链等数据及时调整生产经营决策,提高决策准确性,改善资源配置效率,进而提升产能利用率。此外,公共数据开放还能帮助企业管理层增强创新能力,从而增加优质产能。因此本文在现有研究的基础上,预期公共数据开放能够提高企业产能利用率,但其中的作用机制更主要是数据要素

本身的价值,具体分为改善区域间要素资源错配、提高政府治理效率、增强企业创新能力三个方面。

作用渠道一：改善区域间要素资源错配

根据有效市场假说,在强式有效市场中,证券价格能够反映所有信息,包括公开信息和内幕信息。在投资者均为理性经济人的前提下,每一位投资者作出的价值判断都相同,从而不存在证券发行者和投资者之间由于理性程度和信息掌握程度不同而导致的供求失衡,即市场摩擦。然而,由于复杂的环境和信息使用者的有限理性,大部分投资者只能获得公开信息,甚至由于信息的获取普遍存在意识成本、获取成本和整合成本,很多公开信息也会被忽视(吴武清等,2024),因此,市场摩擦难以避免,资源错配广泛存在。资本、劳动力、土地、技术与数据被视为当前五大生产要素。经济良性循环的必要条件是各生产要素间有机结合,生产、分配、流通、消费各个环节协同发力,达到动态平衡。

公共数据开放平台一方面可供社会公众无门槛使用或低门槛申请各类原始数据;另一方面也允许社会主体将自己收集整理的数据上传,经审核后开放。这两方面共同保障了公共数据的来源丰富、涵盖范围广、获取成本低。同时,由于公共数据开放平台的使用不受所属区域和部门限制,这有助于畅通数据要素在区域间的流通,打破信息壁垒,进而缓解资本、劳动力、技术的流动阻碍,带动生产、分配等各环节的良性循环。因此,本文预期公共数据开放能够改善区域间要素资源错配。随着要素资源在各区域间的流通配置效率不断提高,区域内企业的融资约束、员工选聘困难、能源短缺等问题可能会得到改善,即生产资料能够充分且有效地配置使用,减少过度投资、投资不足及产能浪费等情况,从而提高企业产能利用率。

作用渠道二：提高政府治理效率

各地区上线公共数据开放平台,政府治理效率可能得到明显改善,原因主要有以下几点:第一,公共数据开放能够帮助政府部门提高决策的科学性和合理性。随着大数据、人工智能、物联网等技术的发展,政府部门在日常管理决策和行政服务中越来越依靠这些技术。公共数据来源真实,体量较大,能够为政府机构的决策提供依据。第二,公共数据开放有助于政府部门提升公共服务水平和社会管理水平。例如,多个省市的公共数据开放网站中设置了医疗卫生接口、交通出行接口和治安管理接口,有助于相关主管部门实时查看数据,根据数据调整资源配置、优化调度安排,进而提升公共服务水平和社会管理水平。第三,公共数据开放便于社会公众评估和监督政府部门。公共数据开放平台上线后,社会公众可以便捷地监督政府的行政决策,评估政府的财政预算计划、政务公开程度等,降低信息不对称程度,提高政府透明度。因此本文预期各地区公共数据开放能够提高政府决策效率、改善政府治理效率,从而化解企业产能过剩,提高产能利用率。

作用渠道三：增强企业创新能力

数据要素对于企业提升创新水平和全要素生产率,进而实现价值释放具有重要意义(谢康等,2020;史丹和孙光林,2024)。公共数据开放可能通过促进数据要素的集成与流通,激发企业技术研发和管理优化,从而增强企业创新能力。具体而言,第一,公共数据开放能够为企业创新提供资源补给。在公共数据开放后,企业的数据资源更加丰富,大大降低了数据驱动型研发项目的门槛,也有助于企业技术研发与实际应用场景相结合,加速技术转化,激发中小企业创新活力。第二,公共数据开放能够促进跨行业协同创新。企业创新受资源获取、外部环境及内部能力等多因素影响。石绍宾和刘菲菲(2025)认为,海量数据资源能够促进企业整合关联资源,中小企业可以利用数据平台,与各协同主体共享创新资源。公共数据开放后,企业数据与信息的流通渠道被进一步打通,有助于企业间创新资源共享,发挥资源互补性,增强协同创新能力,降低创新成本。因此,本文预期公共数据开放能够增强企业创新能力。

创新能力的提升有助于企业在技术层面突破生产瓶颈,释放潜在产能;在管理层面优化资源配置,例如采用精益生产和数字化调度可以减少闲置产能;在流程层面进行工艺再造,对能源物料循环使用,进而使得产能利用率大幅提升。因此,本文预期,公共数据开放将通过增强企业创新能力,从而提高产能利用率。

综上,本文提出如下研究假设:

H：公共数据开放能够提高企业产能利用率。

三、研究设计

(一)数据来源与样本选择

本文公共数据开放平台的相关数据通过手工收集获得,并与复旦大学数字与移动治理实验室发布的《中国地方公共数据开放利用报告——省域》(以下简称《报告》)进行比对和交叉验证。具体步骤如下:首先,在浏览器中输入各省(自治区、直辖市)+“公共数据开放”关键字段,检索是否已搭建官方网站,如已搭建则记录网址和网站名称。对于检索无结果的省(自治区、直辖市),使用《报告》进行查验,如有对应网站则补充记录,如无则记为“尚未实施公共数据开放”。然后,使用各省(自治区、直辖市)的公共数据开放网站名称+“开放时间”或“上线时间”作为关键字段进行检索,查看相关新闻,记录开放时间。对于检索无结果的省(自治区、直辖市),使用《报告》进行补充。最后,对搜集到的网址逐一核验,剔除无效网址,最终得到各省(自治区、直辖市)公共数据开放平台建设情况表,作为本文所使用的公共数据开放数据。在稳健性检验中,本文使用地级市层面数据作为补充,收集方法与省份层面相同。

本文所使用的其他公司财务指标、公司治理指标、宏观经济指标数据均来源于CSMAR(国泰安)数据库和RESSET(锐思)数据库。

本文选取2008~2022年作为样本区间,原因为:为了使用双重差分模型体现政策处理效应,需要保留几期事前年份。此外,2007年起,我国企业开始执行修订后的新会计准则。因此,为了确保样本期间各年数据的可比性,选取2008年作为起始期。在获取原始数据的基础上,按照以下顺序进行数据筛选:(1)剔除金融行业公司样本;(2)剔除ST、PT等公司样本;(3)剔除变量包含缺失值的样本。筛选完成后,得到38 834个公司一年度观测值。为了避免极端值的影响,所有连续变量均在上下1%分位数处进行Winsorize缩尾处理。

(二)回归模型与变量定义

1.模型设计

由于各省(自治区、直辖市)公共数据开放平台是陆续建成并开放的,本文构建多期双重差分模型(1):

$$CU_{i,t} = \beta_0 + \beta_1 Data_Openness_{i,t} + \beta_2 Controls_{i,t} + \sum Year + \sum Firm + \varepsilon_{i,t} \quad (1)$$

其中, $CU_{i,t}$ 表示企业产能利用率,借鉴李雪松等(2017)的研究,使用随机前沿生产函数法进行测算; $Data_Openness_{i,t}$ 表示公共数据开放平台上线的虚拟变量,当企业所在地所属省(自治区、直辖市)建设并开放了公共数据开放平台,则样本年份在平台开放当年及以后取值为1,否则为0; $Controls_{i,t}$ 表示一系列可能对产能利用率存在影响的控制变量; $\varepsilon_{i,t}$ 表示随机扰动项。此外,模型中还控制了年份固定效应和公司固定效应,并使用公司层面的聚类稳健标准误。

2.变量定义

(1)被解释变量:产能利用率(CU)。现有研究对于企业产能利用率的测算方法尚不统一,常见的方法主要包括峰值法、比值法、调查法、数据包络分析法、成本函数法、随机生产前沿函数法。参考李雪松等(2017)的研究,随机前沿生产函数法无需大量有关企业各项投入的数据,并且可以更直接地从企业生产角度确定产出前沿面,将实际产出与前沿产出的比值定义为企业产能利用率。因此,本文采用随机前沿生产函数法,使用营业收入、总资产、员工人数三个指标构建随机前沿生产面,以计算产能利用率。

(2)解释变量:公共数据开放(Data_Openness)。当企业所在地所属省(自治区、直辖市)上线并开放了公共数据开放平台,则样本年份在平台开放当年及以后则取值为1,否则为0。

(3)控制变量。参考以往相关文献,本文控制了一系列公司财务层面和公司治理层面的指标,包括公司规模(Size)、资产负债率(Lev)、总资产收益率(ROA)、成长性(Growth)、第一大股东持股比例(Top1)、固定资产比例(PPE)、上市年龄(Listage)、董事会规模(Board)、两职合一(Dual)、产权性质(Soe),此外,还控制了宏观经济特

征指标, 包括资本密集度(KL)、人力资本水平(Humancapital)和经济发展水平(LnperGDP)。

本文变量的具体计算方法见表1。

表 1 变量定义

变量类型	变量名称	变量符号	变量解释
被解释变量	产能利用率	CU	使用营业收入、总资产和员工人数三个指标, 通过随机前沿生产函数法测算
解释变量	公共数据开放	Data_Openness	当企业所在地所属省(自治区、直辖市)上线并开放了公共数据开放平台, 则样本年份在平台开放当年及以后取值为1, 否则为0
控制变量	公司规模	Size	总资产取自然对数
	资产负债率	Lev	负债总额/资产总额
	总资产收益率	ROA	净利润/总资产
	成长性	Growth	营业收入增长率, (当年营业收入-上年营业收入)/上年营业收入
	第一大股东持股比例	Top1	第一大股东持股数/公司总股数
	固定资产比例	PPE	固定资产净值/总资产
	上市年龄	Listage	上市年龄加1后取自然对数
	董事会规模	Board	董事会总人数取自然对数
	两职合一	Dual	如果公司董事长和总经理由一人兼任则取值为1, 否则为0
	产权性质	Soe	国有企业取值为1, 否则为0
	资本密集度	KL	资本存量除以员工人数后取自然对数
	人力资本水平	Humancapital	企业所在省份高等学校在校生人数/总人口
	经济发展水平	LnperGDP	企业所在省份的人均国内生产总值取自然对数

四、实证结果

(一)描述性统计

表2列示了本文研究样本的描述性统计结果。企业产能利用率(CU)的均值为65%, 中位数为66%, 这表明在样本期内, A股上市公司的平均产能利用率为65%左右, 亟待进一步提高。公共数据开放变量(Data_Openness)的均值为0.501, 本文处理组和控制组的样本量基本相等。控制变量与以往文献进行对比, 均在合理范围内。

表 2 描述性统计

Variable	N	Mean	SD	Min	p25	p50	p75	Max
CU	38 834	0.650	0.099	0.344	0.594	0.660	0.717	0.854
Data_Openness	38 834	0.501	0.500	0	0	1	1	1
Size	38 834	22.170	1.298	19.750	21.240	21.990	22.910	26.190
Lev	38 834	0.431	0.205	0.055	0.267	0.425	0.585	0.889
ROA	38 834	0.035	0.063	-0.249	0.013	0.036	0.066	0.196
Growth	38 834	0.170	0.420	-0.582	-0.032	0.106	0.271	2.667
Top1	38 834	0.341	0.149	0.085	0.225	0.319	0.441	0.743
PPE	38 834	0.212	0.160	0.002	0.086	0.179	0.303	0.695
Listage	38 834	2.214	0.746	0.632	1.652	2.340	2.846	3.368
Board	38 834	2.129	0.200	1.609	1.946	2.197	2.197	2.708
Dual	38 834	0.278	0.448	0	0	0	1	1
Soe	38 834	0.361	0.480	0	0	0	1	1
KL	38 834	13.900	0.963	11.600	13.280	13.850	14.470	16.710
Humancapital	38 834	0.021	0.005	0.011	0.017	0.021	0.024	0.035
LnperGDP	38 834	1.928	0.536	0.543	1.581	1.960	2.317	2.931

(二)基准回归结果

表3列示了模型(1)的回归结果。列(1)中未加入控制变量, 仅控制公司与年份固定效应, Data_Openness的回归系数为0.005, 在1%水平上显著。列(2)中加入了控制变量, Data_Openness的回归系数为0.004, 在5%水平上显

著。这表明，当企业所在地所属省份上线公共数据开放平台后，企业的产能利用率平均提高了0.004，该结果支持了本文的假设。

(三)稳健性检验

1. 平行趋势检验

平行趋势检验是双重差分模型结果成立的必要条件，即处理组和对照组在政策实施之前具有平行的趋势，能够在政策实施后识别出政策冲击产生的净效应。本文构建模型(2)用于多期DID模型的平行趋势检验：

$$CU_{i,t} = \beta_0 + \beta_1 before3_{i,t} + \beta_2 before2_{i,t} + \beta_3 before1_{i,t} + \beta_4 current_{i,t} + \beta_5 after1_{i,t} + \beta_6 after2_{i,t} + \beta_7 after3_{i,t} + \beta_8 Controls_{i,t} + \sum Year + \sum Firm + \varepsilon_{i,t} \tag{2}$$

其中，before_{*i,t*}为公共数据开放前年份的虚拟变量，before3_{*i,t*}为政府公共数据开放平台上线前第三年观测值，若观测属于处理组且所处年份为该地区公共数据开放平台上线前第三年，则取值为1，否则为0；current_{*i,t*}为政府公共数据开放平台上线当年观测值，若观测属于处理组且所处年份为该地区公共数据开放平台上线当年，则取值为1，否则为0；after1_{*i,t*}为政府公共数据开放平台上线后一年观测值，若观测属于处理组，且所处年份为该地区公共数据开放平台上线后第一年，则取值为1，否则为0；以此类推。为了避免模型中存在多重共线性问题，本文将政策实施前一期作为基准组，在回归前将其剔除，重新进行回归检验。回归结果如表4和图1所示，在政策实施之前年份的虚拟变量中，均不显著，政策实施当年也不显著，政策实施后年份的虚拟变量均显著为正，表明处理组与对照组的产能利用率变化趋势满足了平行趋势假设，且在公共数据开放平台上线后，企业产能利用率显著提升。

表3 基准回归结果

	(1) CU	(2) CU
Data_Openness	0.005*** (2.62)	0.004** (2.35)
Size		-0.029*** (-13.95)
Lev		0.095*** (11.14)
ROA		0.300*** (24.75)
Growth		0.033*** (29.01)
Top1		-0.001 (-0.06)
PPE		0.034*** (3.31)
Listage		0.014*** (5.05)
Board		0.007 (1.29)
Dual		-0.002 (-1.40)
Soe		-0.006 (-1.25)
KL		0.017*** (7.42)
Humancapital		0.820** (1.97)
LnperGDP		0.002 (0.24)
cons	0.648*** (753.74)	0.940*** (18.78)
Firm_FE	Yes	Yes
Year_FE	Yes	Yes
Cluster by firm	Yes	Yes
N	38 834	38 834
Adj_R ²	0.660	0.717

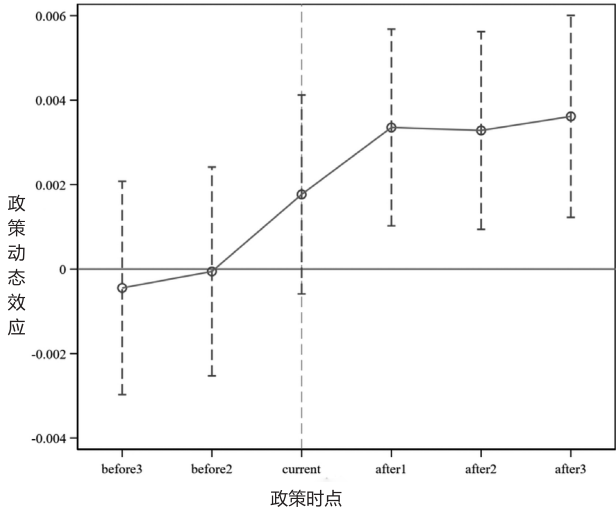


图1 平行趋势检验

2. 安慰剂检验

样本期内企业产能利用率的提高可能是由其他不可观测因素导致的，例如同时期实施的其他政策。本文采用随机化处理组和随机化政策冲击时间的方法，进行500次

表 4 平行趋势检验

	(1) CU
before3	-0.000 (-0.35)
before2	-0.000 (-0.04)
current	0.002 (1.47)
after1	0.003*** (2.82)
after2	0.003*** (2.75)
after3	0.004*** (2.96)
Size	-0.029*** (-39.23)
Lev	0.095*** (27.09)
ROA	0.300*** (47.02)
Growth	0.033*** (44.79)
Top1	-0.001 (-0.18)
PPE	0.034*** (8.59)
Listage	0.015*** (9.54)
Board	0.007** (2.42)
Dual	-0.002** (-2.34)
Soe	-0.006*** (-2.90)
KL	0.017*** (22.35)
Humancapital	0.657*** (3.29)
LnperGDP	0.003 (0.68)
cons	0.943*** (51.53)
Firm_FE	Yes
Year_FE	Yes
Cluster by firm	Yes
N	38 834
Adj_R ²	0.717

回归的安慰剂检验。若所得结果仍能反映出政策效应，则说明本文基准回归的结论不可靠。检验结果如图 2 所示，虚拟政策实施的回归系数集中在 0 附近，显著区别于真实回归系数 0.004，且绝大部分 p 值都大于 0.1，安慰剂检验通过。这在一定程度上表明样本期内企业产能利用率的提升是实施公共数据开放带来的结果，而非其他不可观测因素。因此，本文的结论较为稳健。

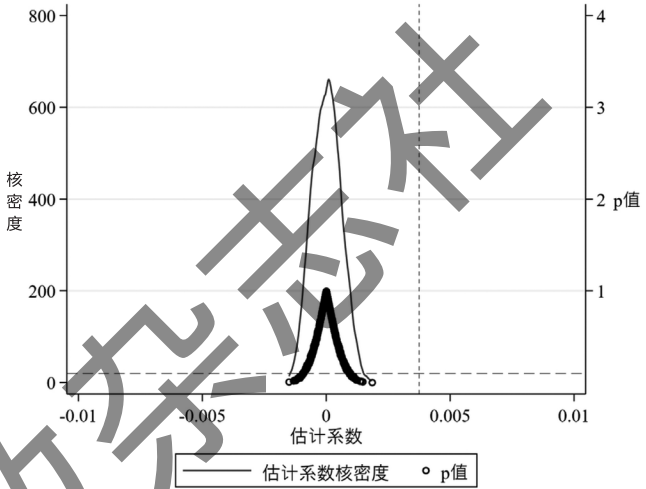


图 2 安慰剂检验

3. 工具变量法

为了进一步缓解反向因果或遗漏变量带来的内生性问题，例如公共数据开放平台更有可能在经济发达的地区率先设立，而这类地区企业可能有较高的产能利用率，从而导致模型中可能存在反向因果或遗漏变量问题，本文使用工具变量法进行检验。

借鉴黄群慧等(2019)和欧阳伊玲等(2024)的研究，采用各省历年政府工作报告中的数字经济词频，经地区生产总值(GDP)标准化后，与该省份 1984 年每百万人邮局数量相乘，作为公共数据开放的工具变量。该工具变量的选取依据是：在互联网普及之前，信息互通主要是通过邮局的信件寄送以及由邮局负责铺设的固定电话实现的，而互联网发展之初也是由电话线拨号接入。因此，人均邮局数量和固定电话数量较多的地区，通常也是互联网兴起较早且发展较快的地区，此类地区更有可能积极搭建并开放公共数据平台。而且，地区数字经济发展程度与是否上线公共数据平台也密切相关，因此满足了工具变量的相关性要求。然而，随着互联网的飞速发展，历史每百万人邮局数量对企业产能利用率的直接影响越来越小，由此满足了工具变

量的外生性要求。

表5列示了工具变量检验的结果。列(1)第一阶段的结果显示,工具变量对公共数据开放的回归系数在1%水平上显著为正,说明信息互通程度越高的地区越有可能上线公共数据开放平台。列(2)第二阶段的结果显示,回归系数在10%水平上显著为正,符合预期。Kleibergen-Paap rk LM统计量的数值为68.35,在1%的显著性水平上拒绝了工具变量不可识别的原假设,满足了工具变量与模型内生解释变量的相关性要求;Cragg-Donald Wald F统计量的值为158.70, Kleibergen-Paap Wald rk F统计量的值为71.28,均大于Stock-Yogo检验10%的临界值16.38,即拒绝“弱工具变量”的原假设。列(3)中将工具变量和解释变量同时纳入回归检验,发现工具变量系数不显著,而解释变量Data_Openness系数在5%水平上显著为正,说明工具变量的确不直接影响被解释变量,满足了排他性条件。综

表5 工具变量法:2SLS

	(1) Data_Openness	(2) CU	(3) CU
IV	0.001*** (8.44)		0.000 (1.19)
Data_Openness		0.022* (1.85)	0.004** (2.18)
Size	0.005 (0.84)	-0.030*** (-13.62)	-0.030*** (-13.92)
Lev	0.053* (1.75)	0.094*** (10.36)	0.096*** (11.27)
ROA	0.152*** (3.53)	0.293*** (21.11)	0.300*** (24.58)
Growth	-0.001 (-0.23)	0.033*** (28.21)	0.033*** (28.56)
Top1	-0.037 (-0.94)	0.001 (0.04)	-0.001 (-0.12)
PPE	-0.092*** (-2.79)	0.039*** (3.65)	0.035*** (3.40)
Listage	0.028*** (2.25)	0.013*** (4.14)	0.015*** (5.09)
Board	-0.004 (-0.16)	0.007 (1.30)	0.007 (1.29)
Dual	0.003 (0.33)	-0.002 (-1.42)	-0.002 (-1.38)
Soe	0.035** (2.14)	-0.008 (-1.53)	-0.006 (-1.28)
KL	-0.007 (-1.08)	0.017*** (7.24)	0.016*** (7.23)
Humancapital	-4.95*** (-2.65)	1.04** (2.10)	0.794* (1.89)
LnperGDP	-0.027 (-0.53)	0.006 (0.61)	0.005 (0.51)
cons	0.537*** (3.18)	0.542*** (3.75)	0.942*** (18.75)
Firm_FE	Yes	Yes	Yes
Year_FE	Yes	Yes	Yes
Cluster by firm	Yes	Yes	Yes
N	37 721	37 721	37 721
Adj_R ²	0.688	0.114	0.719

上,本文所选工具变量是合理且有效的,且在考虑内生性问题后,公共数据开放对企业产能利用率的促进作用依然显著。

4. PSM-DID 检验

为缓解由某些不可观测因素造成的选择偏误问题,本文采用PSM-DID(多时点倾向得分匹配—双重差分模型)方法进行检验。具体而言,将企业所属省份上线了公共数据开放平台并且观测年份在公共数据开放当年及以后的样本划分为处理组,以基准回归中的控制变量作为协变量,使用卡尺距离为0.05的1:1最近邻匹配方法,采用Logit回归为处理组样本匹配特征相似的对照组样本。表6列(1)的结果显示,匹配后样本的回归结果仍在10%的水平上显著为正,ATT(即平均处理效应)在5%水平上显著为正。这表明使用PSM-DID方法后,本文的结论依然稳健。

表6 PSM-DID、更换变量度量检验

	PSM-DID	更换被解释变量衡量方法		更换解释变量衡量方法	
	(1) PSM_CU	(2) CU	(3) CU	(4) CU	(5) CU
Data_Openness	0.003* (1.71)	0.021*** (2.94)	0.014** (2.08)	0.004** (2.26)	0.003* (1.80)
Size	-0.029*** (-12.67)		-0.039*** (-3.80)		-0.029*** (-12.94)
Lev	0.096*** (10.46)		0.051 (1.44)		0.097*** (10.81)
ROA	0.302*** (22.38)		1.062*** (24.15)		0.298*** (23.36)
Growth	0.034*** (26.98)		0.126*** (24.04)		0.033*** (27.54)
Top1	-0.000 (-1.508)		-0.040 (-0.78)		-0.000 (-0.12)
PPE	0.031*** (3.00)		-0.072* (-1.78)		0.030*** (2.71)
Listage	0.014*** (4.01)		0.062*** (3.59)		0.016*** (5.32)
Board	0.006 (1.03)		0.011 (0.47)		0.008 (1.41)
Dual	-0.002 (-1.17)		-0.006 (-0.81)		-0.001 (-0.61)
Soe	-0.007 (-1.25)		-0.006 (-0.36)		-0.006 (-1.25)
KL	0.017*** (7.30)		-0.084*** (-9.75)		0.016*** (6.62)
Humancapital	0.792* (1.75)		1.720 (0.98)		0.814* (1.89)
LnperGDP	0.001 (0.05)		0.017 (0.38)		0.005 (0.48)
cons	0.925*** (17.30)	0.639*** (179.22)	2.386*** (10.60)	0.647*** (771.72)	0.936*** (17.61)
Firm_FE	Yes	Yes	Yes	Yes	Yes
Year_FE	Yes	Yes	Yes	Yes	Yes
Cluster by firm	Yes	Yes	Yes	Yes	Yes
N	31 709	32 193	32 193	34 547	34 547
Adj_R ²	0.714	0.769	0.805	0.662	0.718

5. 更换被解释变量衡量方法

为避免使用单一变量衡量方法导致研究结果可信度不高,本文参考国务院发展研究中心《进一步化解产能过剩的政策研究》课题组(2015)的做法,使用与企业产能利用率高度相关的总资产周转率作为替代指标,重新进行回归估计。总资产周转率为营业收入与资产总额的比值。检验结果如表6列(2)和列(3)所示,更换产能利用率的衡量方式后,本文的结论依然稳健。

6. 更换解释变量衡量方法

本文基准回归聚焦于省份层面的公共数据开放对企业产能利用率的影响。除省级平台外,截至2022年10月,我国已有187个城市(含直辖市、副省级与地级行政区)上线了公共数据开放平台。因此,本文将省级层面替换为城市层面,重新划分处理组进行双重差分检验。结果如表6列(4)和列(5)所示,解释变量的回归系数在1%的水平上显著为正,再次验证了本文结论的稳健性。

7. 排除同期其他政策干扰

除公共数据开放以外,在2008~2022年期间还颁布并实施了多项类似的政策。为排除这些政策的干扰,我们手工收集了四项与公共数据开放较为相似的政策以及一项与企业产能利用率息息相关的政策数据,并纳入基准回归模型进行检验。这五项政策分别是智能制造试点、国家大数据综合试验区、“宽带中国”试点、“互联网+政务服务”以及《国务院关于化解产能严重过剩矛盾的指导意见》。由于智能制造试点、国家大数据综合试验区和化解产能过剩的指导意见是在省级层面实施的政策,所以我们在省级公共数据开放的模型中将其加入控制变量;而“互联网+政务服务”以及“宽带中国”试点则是在地级市层面实施的政策,我们将其纳入城市层面公共数据开放的模型中。检验结果如表7所示,Data_Openness的回归系数依然显著,且各同期政策的回归系数均不显著。这表明本文公共数据开放对企业产能利用率的影响不受此五项政策影响,进一步验证了本文研究结论的稳健性。

五、进一步分析

(一)作用机制检验

前文已述,改善区域间要素资源错配、提高政府治理效率、增强企业创新能力可能是公共数据开放能够提高企业产能利用率的三条作用渠道。本节借鉴江艇(2022)的做法,进行作用机制检验。

1. 改善区域间要素资源错配

借鉴Hsieh和Klenow(2009)、陈永伟和胡伟民(2011)、白俊红和刘宇英(2018)、崔书会等(2019)的方法,使用各地区生产总值、固定资本存量和年平均就业人数,计算各省的资本和劳动产出弹性,进而测算出地区资本错配程度和劳动力错配程度。控制地区人口密度Populationdensity、城镇化水平Urbanization、工业化水平Industrialization、外商直接投资OFDI、政府干预程度Goverintervene、对外开放程度Openness、人均国内生产总值LnperGDP,进行回归。各地区生产总值、固定资本存量及就业人数等数据来源于《中国统计年鉴》。

回归结果如表8列(1)和列(2)所示,资本错配程度对公共数据开放虚拟变量的回归系数在1%水平上显著为负,而劳动力错配程度的回归系数则不显著。这可能意味着公共数据资源的开放更直接地促进了资本要素在地区间的流通,即缓解资本要素的流动障碍,而对劳动力要素的流通尚无明显影响。因此,公共数据开放可能通过改善区域间资本要素的错配程度,从而提升企业的产能利用率。

2. 提高政府治理效率

参考以往文献,本文使用营商环境和企业的政策不确定性感知水平两方面来反映政府治理效率(李光勤等,2024;何雨可等,2024;丁声悒和白俊红,2025;叶永卫等,2025)。营商环境改善和企业的政策不确定性感知水平降低均能在一定程度上表明地方政府治理效率的提升。营商环境指数参考张三保等(2020)的研究,通过构建各省份营商环境评价指标体系得到;企业的政策不确定性感知水平参考方明月等(2023)的研究,通过对企业年报进行文本分析得到。

表 7 排除同期其他政策干扰

	(1) 智能制造试点	(2) 国家大数据综合 试验区	(3) “宽带中国”试点	(4) “互联网+政务 服务”	(5) 化解产能过剩
Data_Openness	0.004** (2.28)	0.004** (2.28)	0.003* (1.85)	0.003* (1.84)	0.004** (2.26)
智能制造试点DID	0.003 (0.68)				
国家大数据综合试验区DID		-0.001 (-0.21)			
“宽带中国”试点DID			-0.000 (-0.16)		
“互联网+政务服务”DID				-0.002 (-0.57)	
化解产能过剩DID					-0.002 (-0.56)
Size	-0.030*** (-14.07)	-0.030*** (-14.07)	-0.029*** (-12.94)	-0.029*** (-12.93)	-0.029*** (-13.96)
Lev	0.096*** (11.32)	0.096*** (11.32)	0.097*** (10.77)	0.097*** (10.78)	0.095*** (11.14)
ROA	0.296*** (24.31)	0.296*** (24.30)	0.297*** (23.22)	0.297*** (23.19)	0.300*** (24.76)
Growth	0.033*** (28.82)	0.033*** (28.77)	0.033*** (27.47)	0.033*** (27.46)	0.033*** (29.01)
Top1	0.000 (0.01)	-0.000 (-0.01)	-0.000 (-0.16)	-0.000 (-0.16)	-0.001 (-0.06)
PPE	0.032*** (3.06)	0.032*** (3.06)	0.030*** (2.66)	0.030*** (2.66)	0.034*** (3.31)
Listage	0.015*** (5.24)	0.015*** (5.24)	0.016*** (5.36)	0.016*** (5.37)	0.014*** (5.06)
Board	0.007 (1.32)	0.007 (1.31)	0.008 (1.38)	0.008 (1.38)	0.007 (1.30)
Dual	-0.002 (-1.20)	-0.002 (-1.21)	-0.001 (-0.61)	-0.001 (-0.61)	-0.002 (-1.41)
Soe	-0.005 (-1.04)	-0.005 (-1.04)	-0.006 (-1.28)	-0.006 (-1.27)	-0.006 (-1.25)
KL	0.016*** (7.24)	0.016*** (7.23)	0.016*** (6.61)	0.016*** (6.60)	0.017*** (7.41)
Humancapital	0.776* (1.88)	0.772* (1.84)	0.795* (1.85)	0.764* (1.74)	0.843** (2.02)
LnperGDP	0.004 (0.46)	0.004 (0.44)	0.005 (0.52)	0.005 (0.52)	0.002 (0.20)
cons	0.951*** (18.66)	0.952*** (18.63)	0.938*** (17.63)	0.939*** (17.61)	0.941*** (18.80)
Firm_FE	Yes	Yes	Yes	Yes	Yes
Year_FE	Yes	Yes	Yes	Yes	Yes
Cluster by firm	Yes	Yes	Yes	Yes	Yes
N	37 448	37 448	34 397	34 397	38 834
Adj_R ²	0.717	0.717	0.718	0.718	0.717

表 8 作用机制检验

	改善区域间要素资源错配		提高政府治理效率		增强企业创新能力	
	(1) 资本错配程度	(2) 劳动力错配程度	(3) 营商环境	(4) 企业政策 不确定性感知	(5) 创新投入	(6) 创新产出
Data_Openness	-0.092*** (-3.15)	-0.008 (-0.65)	1.095*** (10.93)	-0.005* (-1.67)	0.015*** (3.01)	0.038*** (3.54)
Populationdensity	0.003*** (12.81)	-0.001*** (-2.79)	0.077* (1.99)			
Urbanization	-0.046** (-2.62)	0.020*** (3.51)	-0.518 (-0.96)			
Industrialization	-0.340 (-0.65)	0.321 (1.64)	-7.536 (-0.43)			
OFDI	-1.123 (-1.44)	0.564* (1.81)	-28.618** (-2.18)			
Goverintervene	0.287 (0.44)	0.191 (0.62)	97.220*** (4.50)			
Openness	0.314 (1.53)	-0.062 (-1.18)	-0.435 (-0.07)			
LnperGDP	0.041 (1.12)	-0.002 (-0.15)	37.886 (1.59)	0.033*** (2.61)	0.015 (0.63)	0.068 (1.35)
Size				0.005* (1.84)	0.034*** (6.45)	0.083*** (7.68)
Lev				-0.035*** (-3.19)	-0.060*** (-2.76)	-0.080* (-1.84)
ROA				-0.021*** (-1.31)	0.115*** (3.90)	0.105 (1.61)
Growth				0.003* (1.89)	0.009*** (2.73)	-0.007 (-0.95)
Top1				-0.003 (-0.20)	-0.030 (-0.99)	-0.038 (-0.64)
PPE				0.022* (1.82)	-0.039 (-1.48)	-0.004 (-0.09)
Listage				-0.022*** (-4.83)	-0.130*** (-12.29)	-0.265*** (-12.01)
Board				0.015*** (2.01)	0.011 (0.76)	0.025 (0.80)
Dual				-0.004 (-1.28)	0.001 (0.11)	0.012 (1.22)
Soe				0.007 (1.13)	0.003 (0.23)	-0.012 (-0.54)
KL				-0.007** (-2.40)	-0.020*** (-3.74)	-0.038*** (-3.61)
Humancapital				-0.469 (-0.67)	2.003* (1.65)	1.427 (0.58)
cons	1.012 (0.90)	-0.697 (-1.18)	-70.363 (-1.30)	-0.156** (-2.36)	0.210 (1.62)	-0.121 (-0.45)
Firm_FE	No	No	No	Yes	Yes	Yes
Province_FE	Yes	Yes	Yes	No	No	No
Year_FE	Yes	Yes	Yes	Yes	Yes	Yes
Cluster by firm	No	No	No	Yes	Yes	Yes
Cluster by province	Yes	Yes	Yes	No	No	No
N	30 351	30 351	20 661	37 255	29 877	29 877
Adj_R ²	0.963	0.939	0.885	0.287	0.616	0.452

检验结果如表8列(3)和列(4)所示,营商环境的回归系数在1%水平上显著为正,企业政策不确定性感知水平的回归系数在10%水平上显著为负,说明公共数据开放能够提高政府治理效率,具体表现为改善企业所在地区的营商环境和降低企业的政策不确定性感知。因此,公共数据开放通过作用于政府治理效率进而提高产能利用率。

3. 增强企业创新能力

参考以往文献(黎文靖等, 2021; 孙帆等, 2023; 孙薇和叶初升, 2023)的做法, 本文使用创新投入和创新产出两方面来衡量企业创新能力。其中, 创新投入使用研发支出与上期营业收入之比表示, 创新产出使用发明专利被引数的自然对数表示。

检验结果如表8列(5)和列(6)所示, 创新投入和创新产出的回归系数均在1%水平上显著为正, 说明公共数据开放能够促进企业技术研发, 激发企业创新活力, 从而提高产能利用率。

(二) 异质性分析

尽管前文的分析已经证实政府公共数据开放能够提升企业产能利用率, 但二者间的关系可能受到哪些因素影响? 在不同情境下二者的关系是否存在差异? 本文从以下两方面进行探究。

1. 市场化程度

受地理区位、经济政策以及地域文化等因素的影响, 市场化进程不同的地区在政府干预、法治建设及市场资源配置等方面均存在较大差异。公共数据开放对企业产能利用率的提升作用在不同市场环境中可能不同, 市场化程度高的地区本身信息化水平较高、政府治理能力较强、要素资源错配程度可能较低, 从而公共数据开放对企业产能利用率的提升作用可能会被削弱。本文使用中国分省份市场化指数衡量企业所在地市场化程度进行检验。为了体现政府公共数据开放这一外生冲击对不同情境下企业产能利用率的净效应是否存在差异, 分组依据为所使用指标的滞后一期值, 即上一年度的指标值。

分组检验结果如表9列(1)和列(2)所示, 结果显示, 公共数据开放对企业产能利用率的提升作用仅在上年市场化程度较低的地区显著, 而对于市场化程度较高地区的企业无明显影响, 验证了本文的推断。检验结果通过了组间系数差异性检验(检验方法为费舍尔组合检验)。

2. 产业类型

公共数据开放对企业产能利用率的提升作用对于不同产业类型的企业可能存在差异。例如技术密集型企业高度依赖先进技术、知识产权和研发创新能力, 劳动力需求量相对较少; 而劳动密集型企业对技术和设备的依赖程度较低。技术密集型企业可能由于掌握大量先进技术, 自动化和智能化水平较高, 本身的生产运营效率就较高, 公共数据资源开放所带来的资源补给作用在此类企业中可能会被削弱, 相反, 劳动密集型企业受到公共数据开放的资源补给作用可能更为显著。本文针对产业类型进行异质性分析, 分组依据为所使用指标的滞后一期值, 即上一年度的指标值。

分组检验结果如表9列(3)和列(4)所示, 公共数据开放对企业产能利用率的提升作用仅在上年属于劳动密集型的企业样本中显著, 验证了本文的推断。检验结果通过了组间系数差异性检验(检验方法为费舍尔组合检验)。

3. 企业数字化转型程度

随着大数据、人工智能等技术的发展, 其在企业日常生产经营活动中的作用愈发重要, 企业提高自身运行效率和竞争优势离不开数字技术和数字化业务的应用。若企业本身数字化水平不高, 数据挖掘和数据分析能力较为欠缺, 可能不利于投融资及生产等决策效率的提高, 从而导致产能利用率的提高空间受限。公共数据开放后, 这类企业能够获得海量数据资源补给, 并利用这些开放数据训练和优化其机器学习模型, 提升算法的性能和准确性, 从而增强企业的数据分析能力。此外, 开放数据还可以帮助这类企业识别潜在的效率瓶颈, 通过数字化手段进行改进, 从而提升整体运营效率和数字化业务应用能力。而数字化水平较高的企业本身具备较为先进的数字技术、数字化管理模式, 公共数据开放的资源补给作用可能相对较小, 从而对这类企业产能利用率的提升作用可能被削弱。本文针对数字化水平进行异质性分析, 分组依据为所使用指标的滞后一期值, 即上一年度的指标值。

本文参考吴非等(2021)的做法,对上市公司年报中管理层讨论与分析部分进行文本分析,统计“人工智能”“机器学习”“数据挖掘”“自然语言处理”等数字化相关词汇的词频,取自然对数处理后作为企业数字化水平的代理指标。结果如表9列(5)和列(6)所示,公共数据开放对企业产能利用率的提升作用仅在上年数字化水平较低的企业中显著,验证了本文的推断。检验结果通过了组间系数差异性检验(检验方法为费舍尔组合检验)。

表9 异质性分析

	(1) 上年市场化程度高	(2) 上年市场化程度低	(3) 技术密集型	(4) 劳动密集型	(5) 上年数字化水平高	(6) 上年数字化水平低
Data_Openness	0.004 (1.63)	0.007*** (2.67)	0.003 (1.21)	0.008*** (2.85)	0.002 (0.75)	0.005** (2.39)
Size	-0.027*** (-8.56)	-0.036*** (-10.19)	-0.033*** (-11.93)	-0.028*** (-7.35)	-0.031*** (-9.02)	-0.033*** (-11.38)
Lev	0.073*** (5.75)	0.119*** (9.39)	0.106*** (9.43)	0.104*** (7.65)	0.105*** (8.07)	0.080*** (7.16)
ROA	0.328*** (17.93)	0.236*** (13.25)	0.307*** (18.97)	0.294*** (14.04)	0.207*** (13.25)	0.324*** (18.42)
Growth	0.035*** (21.74)	0.033*** (17.93)	0.029*** (17.32)	0.039*** (20.87)	0.034*** (16.34)	0.035*** (21.69)
Top1	-0.012 (-0.73)	0.015 (0.84)	-0.017 (-1.19)	0.005 (0.26)	-0.000*** (-2.96)	-0.000 (-0.80)
PPE	0.011 (0.78)	0.065*** (3.81)	0.009 (0.61)	0.032 (1.84)	0.006 (0.27)	0.052*** (4.24)
Listage	0.022*** (3.58)	0.021*** (3.73)	0.017*** (5.03)	0.013*** (2.50)	0.024*** (3.59)	0.016*** (2.94)
Board	0.010 (1.34)	0.003 (0.43)	0.006 (0.85)	0.008 (0.98)	0.006 (0.73)	0.005 (0.69)
Dual	-0.002 (-0.91)	-0.002 (-0.79)	-0.003 (-1.54)	-0.003 (-1.02)	-0.004 (-1.40)	-0.000 (-0.18)
Soe	-0.002 (-0.40)	-0.006 (-0.71)	-0.000 (-0.04)	-0.011 (-1.15)	-0.001 (-0.20)	-0.010 (-1.62)
KL	0.013*** (3.73)	0.021*** (6.62)	0.011*** (3.45)	0.019*** (5.68)	0.017*** (4.76)	0.018*** (6.28)
Humancapital	-0.122 (-0.18)	1.488 (1.56)	0.661 (1.13)	0.170 (0.24)	0.475 (0.62)	0.800 (1.39)
LnperGDP	0.007 (0.53)	-0.005 (-0.21)	0.005 (0.40)	0.029* (1.65)	-0.024 (-1.23)	0.012 (1.05)
cons	0.946*** (13.12)	1.002*** (10.53)	1.103*** (16.83)	0.832*** (9.06)	1.045*** (11.24)	0.983*** (14.75)
Firm_FE	Yes	Yes	Yes	Yes	Yes	Yes
Year_FE	Yes	Yes	Yes	Yes	Yes	Yes
Cluster by firm	Yes	Yes	Yes	Yes	Yes	Yes
N	16 016	15 725	24 439	13 773	12 463	19 170
Adj_R ²	0.733	0.744	0.729	0.759	0.778	0.742
组间差异检验 经验P值	b0-b1=-0.003 P=0.000		b0-b1=-0.005 P=0.010		b0-b1=-0.003 P=0.010	

六、结论与启示

公共数据开放能否作用于企业,提高产能利用率,是实务界和学术界亟待探究的问题。本文利用2008~2022年A股上市公司数据,分析了公共数据开放对企业产能利用率的影响。研究发现,各地区上线公共数据开放平台显著

提高了企业产能利用率,其中的作用渠道主要是改善区域间要素资源错配(资本要素)、提高政府治理效率及增强企业创新能力。这种效应在地区市场化水平较低的企业、数字化水平较低的企业以及劳动密集型企业中更加显著。本文通过了平行趋势检验、安慰剂检验、工具变量法、倾向得分匹配、更换变量衡量方式、排除同期其他政策影响等稳健性检验,结论依然成立。

根据本文的研究结论,提出以下建议:第一,政府部门应当坚持常态化公共数据开放,并提高数据开放质量和普及度,使公共数据更普惠地为企业所用。同时应引导各市场主体重视并加强公共数据的运用,提高数据资源传播速度和配置效率。特别是在市场化程度较低的地区,政府开放公共数据对于企业获取并利用数据要素、合理制定生产经营决策,进而提高产能利用率有着显著促进作用。因此,这些地区更应重视加快推进公共数据全面开放。第二,鼓励第三方技术机构或行业协会开发低成本、易操作的数据分析工具,降低中小企业数据应用门槛;推动建立统一的数据接口标准和开放平台,解决数据格式碎片化问题,促进跨行业数据融合。第三,企业等各市场主体应高度重视公共数据资源的利用,将数据要素作为资本、劳动力、土地、技术四大生产要素的重要补充。应合理使用公共数据,充分挖掘公共数据的经济和社会价值,降低资源错配程度,合理、精准地提高产能利用率,为我国实体经济高质量发展作出贡献。

主要参考文献:

- [1] 白俊红,刘宇英.对外直接投资能否改善中国的资源错配[J].中国工业经济,2018,(1):60-78.
- [2] 卞元超,白俊红.区域市场整合能否提升企业的产能利用率?[J].财经研究,2021,47(11):64-77.
- [3] 曹献飞,裴平.企业OFDI能促进中国经济高质量发展吗?——基于产能治理视角的实证研究[J].中央财经大学学报,2019,(11):96-104.
- [4] 陈林,陈臻,肖倩冰.产能过剩与环境规制相机选择——基于正式与非正式环境规制视角[J].中国工业经济,2024,(3):62-80.
- [5] 陈艳利,蒋琪.数据生产要素视角下开放公共数据与企业创新——基于建立公共数据开放平台的准自然实验[J].经济管理,2024,46(1):25-46.
- [6] 陈永伟,胡伟民.价格扭曲、要素错配和效率损失:理论和应用[J].经济学(季刊),2011,10(4):1401-1422.
- [7] 崔书会,李光勤,豆建民.产业协同集聚的资源错配效应研究[J].统计研究,2019,36(2):76-87.
- [8] 丁声悒,白俊红.公共数据开放与营商环境优化——来自政府数据平台上线的准自然实验[J].现代财经(天津财经大学学报),2025,(1):23-40.
- [9] 方锦程,刘颖,高昊宇,等.公共数据开放能否促进区域协调发展?——来自政府数据平台上线的准自然实验[J].管理世界,2023,39(9):124-142.
- [10] 方明月,聂辉华,阮睿,等.企业数字化转型与经济政策不确定性感知[J].金融研究,2023,(2):21-39.
- [11] 方森辉,毛其淋.人力资本扩张与企业产能利用率——来自中国“大学扩招”的证据[J].经济学(季刊),2021,21(6):1993-2016.
- [12] 郭长林.财政政策扩张、纵向产业结构与中国产能利用率[J].管理世界,2016,(10):13-33.
- [13] 国务院发展研究中心《进一步化解产能过剩的政策研究》课题组.当前我国产能过剩的特征、风险及对策研究——基于实地调研及微观数据的分析[J].管理世界,2015,(4):1-10.
- [14] 韩国高,陈庭富,刘田广.数字化转型与企业产能利用率——来自中国制造企业的经验发现[J].财经研究,2022,48(9):154-168.
- [15] 何青,琚望静,庄朋涛.如何缓解企业投融资期限错配?基于数字化转型视角[J].数量经济技术经济研究,2024,41(5):113-133.
- [16] 何雨可,牛耕,逯建,等.数字治理与城市创业活力——来自“信息惠民国家试点”政策的证据[J].数量经济技术经济研究,2024,41(1):47-66.

- [17] 黄群慧, 余泳泽, 张松林. 互联网发展与制造业生产率提升: 内在机制与中国经验[J]. 中国工业经济, 2019, (8): 5-23.
- [18] 黄卓, 陶云清, 刘兆达, 等. 智能制造如何提升企业产能利用率——基于产消合一的视角[J]. 管理世界, 2024, 40(5): 40-59.
- [19] 冀云阳, 周鑫, 张谦. 数字化转型与企业创新——基于研发投入和研发效率视角的分析[J]. 金融研究, 2023, (4): 111-129.
- [20] 江艇. 因果推断经验研究中的中介效应与调节效应[J]. 中国工业经济, 2022, (5): 100-120.
- [21] 黎文靖, 彭远怀, 谭有超. 知识产权司法保护与企业创新——兼论中国企业创新结构的变迁[J]. 经济研究, 2021, 56(5): 144-161.
- [22] 李光勤, 李梦娇, 张廷海. 政府数据开放、营商环境与FDI流入[J]. 上海财经大学学报, 2024, 26(6): 49-61.
- [23] 李雪松, 赵宸宇, 聂菁. 对外投资与企业异质性能利用率[J]. 世界经济, 2017, 40(5): 73-97.
- [24] 罗奇, 陈梁, 赵永亮. 数字基础设施建设与企业产能利用率——来自“宽带中国”战略的经验证据[J]. 产业经济研究, 2022, (5): 1-14.
- [25] 马新啸, 汤泰劼, 郑国坚. 混合所有制改革能化解国有企业产能过剩吗? [J]. 经济管理, 2021, 43(2): 38-55.
- [26] 毛捷, 管星华. 地方政府纾困政策的效应研究: 来自上市公司的证据[J]. 经济研究, 2022, 57(9): 82-98.
- [27] 倪婷婷, 王跃堂. 区域行政整合、要素市场化与企业资源配置效率[J]. 数量经济技术经济研究, 2022, 39(11): 136-156.
- [28] 欧阳伊玲, 王愉靖, 李平, 等. 数据要素与城投债定价: 基于公共数据开放的准自然实验[J]. 世界经济, 2024, (2): 174-203.
- [29] 潘俊, 卞子咏, 赵洵, 等. 政府公共数据开放与审计师定价决策——基于构建数据开放平台的准自然实验[J]. 审计研究, 2023, (6): 136-148.
- [30] 申志轩, 祝树金, 汤超, 等. 数字化转型如何影响企业投融资期限错配问题? ——基于资产端和负债端的双重视角[J]. 财经研究, 2024, 50(5): 139-153.
- [31] 石绍宾, 刘菲菲. 国家级大数据综合试验区与企业协同创新[J]. 山西大学学报(哲学社会科学版), 2025, 48(2): 109-117.
- [32] 史丹, 孙光林. 数据要素与新质生产力: 基于企业全要素生产率视角[J]. 经济理论与经济管理, 2024, 44(4): 12-30.
- [33] 孙帆, 杜勇, 胡红燕. 企业数字化转型的去产能效应研究[J]. 软科学, 2023, 37(10): 122-128.
- [34] 孙薇, 叶初升. 政府采购何以牵动企业创新——兼论需求侧政策“拉力”与供给侧政策“推力”的协同[J]. 中国工业经济, 2023, (1): 95-113.
- [35] 陶锋, 王欣然, 徐扬, 等. 数字化转型、产业链供应链韧性与企业生产率[J]. 中国工业经济, 2023, (5): 118-136.
- [36] 王海, 叶帅, 尹俊雅. 公共数据开放如何提振企业有效投资——基于产能利用视角[J]. 中国工业经济, 2024, (8): 137-153.
- [37] 王文甫, 明娟, 岳超云. 企业规模、地方政府干预与产能过剩[J]. 管理世界, 2014, (10): 17-36.
- [38] 吴非, 胡慧芷, 林慧妍, 等. 企业数字化转型与资本市场表现——来自股票流动性的经验证据[J]. 管理世界, 2021, 37(7): 130-144.
- [39] 吴武清, 李祁恒, 章柳漪, 等. 公共数据资源与企业全要素生产率——基于地方政府数据开放的准自然实验[J]. 系统工程理论与实践, 2024, 44(6): 1815-1833.
- [40] 肖土盛, 孙瑞琦, 袁淳, 等. 企业数字化转型、人力资本结构调整与劳动收入份额[J]. 管理世界, 2022, 38(12): 220-237.
- [41] 谢康, 夏正豪, 肖静华. 大数据成为现实生产要素的企业实现机制: 产品创新视角[J]. 中国工业经济, 2020, (5): 42-60.
- [42] 徐业坤, 马光源. 地方官员变更与企业产能过剩[J]. 经济研究, 2019, 54(5): 129-145.
- [43] 杨龙见, 李世刚, 刘盛宇, 等. 增值税留成会影响企业产能利用率吗? [J]. 经济学(季刊), 2019, 18(4): 1397-1418.
- [44] 杨秀云, 韩奇. 公共数据开放能提升企业全要素生产率吗? [J]. 证券市场导报, 2023, (12): 18-30.
- [45] 叶永卫, 余田田, 陶云清, 等. 政府数据要素共享的稳投资效应: 来自公共数据开放平台的证据[J]. 数量经济技术经济研究, 2025, (1): 136-156.
- [46] 张三保, 康璧成, 张志学. 中国省份营商环境评价: 指标体系与量化分析[J]. 经济管理, 2020, 42(4): 5-19.
- [47] 张云, 柏培文. 数智化如何影响双循环参与度与收入差距——基于省级—行业层面数据[J]. 管理世界, 2023, 39(10): 58-83.
- [48] 郑雪梅, 邹粉菊, 徐伟呈. 数字经济、要素资源错配与企业产能利用率[J]. 南京财经大学学报, 2023, (6): 90-98.
- [49] 钟春平, 潘黎. “产能过剩”的误区——产能利用率及产能过剩的进展、争议及现实判断[J]. 经济学动态, 2014, (3): 35-47.
- [50] 周冬华, 万贻健. 数字化转型能有效促进企业跨地域创新吗? ——基于异地子公司创新视角[J/OL]. 南开管理评论,

2024. [http : //kns.cnki.net/kcms/detail/12.1288.F.20240828.0951.002.html](http://kns.cnki.net/kcms/detail/12.1288.F.20240828.0951.002.html).
- [51] 朱国悦, 陶锋. 供应链数字化与企业产能利用率[J/OL]. 当代财经, 2025. [https : //doi.org/10.13676/j.cnki.cn36-1030/f.20240919.002](https://doi.org/10.13676/j.cnki.cn36-1030/f.20240919.002).
- [52] 朱喜安, 马樱格. 数字化转型如何推动企业人力资本结构升级[J]. 经济管理, 2024, 46 (2) : 51-71.
- [53] Gascó-Hernández, M., Martín, E. G., Reggi, L., et al. Promoting the use of open government data: Cases of training and engagement[J]. Government Information Quarterly, 2018, 35 (2) : 233-242.
- [54] Hsieh, C. T., Klenow, P. J. Misallocation and manufacturing TFP in China and India[J]. The Quarterly Journal of Economics, 2009, 124 (4) , 1403-1448.
- [55] Richardson, S. Over-investment of free cash flow[J]. Review of Accounting Studies, 2006, 11 (2/3) : 159-189.

Data Elements and Enterprise Capacity Utilization Rate

——A Quasi-Natural Experiment Based on Public Data Opening

LI Xiao, DU Yixuan, YU Xiaoxu

Abstract: Data, as a new factor of production, represents both a fundamental and strategic resource. The development and opening of public data are key to unlocking the potential of data elements. This paper uses the gradual opening of public data by governments across Chinese provinces (municipalities and autonomous regions) as a quasi-natural experiment. Using a sample of A-share listed companies from 2008 to 2022 and a staggered difference-in-differences (DID) model, we examine the impact of public data opening on enterprise capacity utilization rate. The evidence reveals that public data opening significantly enhances enterprise capacity utilization rate. The underlying mechanisms include alleviating interregional capital misallocation, improving government governance efficiency, and strengthening corporate innovation capabilities, thereby increasing capacity utilization rate. Further analysis shows that these effects are more pronounced for enterprises located in provinces with lower marketization levels, enterprises with lower digitalization levels, and labor-intensive industries. Our study contributes to the research on how public data resources empower high-quality economic development and provides new insights to Chinese enterprises in improving capacity utilization rate and achieving high-quality development.

Key words: data elements; public data opening; capacity utilization rate; factor resource mismatch

(责任编辑 杨亚彬)