

# 中国数字经济核心产业创新创业： 典型事实与指数编制<sup>\*</sup>

戴若尘 王艾昭 陈斌开

**摘要：**数字经济发展事关国家发展大局，推动数字经济核心产业创新创业对我国经济高质量发展至关重要。本文基于北京大学企业大数据研究中心整合的企业大数据，立足于数字经济核心产业的企业家、资本与技术三大核心要素，从新建企业、外来投资、风险投资、专利、商标和软件著作权六个维度，创新性地构建了中国数字经济核心产业创新创业指数，并首次描绘我国 2010—2020 年数字经济核心产业创新创业的动态演变过程和空间分布特征。该指数利用大数据思维构建，基于微观创新主体，关注创新产出而非投入，具有客观性、时效性和多维性的特点。研究发现：第一，过去十年，我国数字产业创新创业水平快速提高；第二，省级数字产业创新创业水平呈现“南热北冷、中部崛起”的空间分化格局；第三，城市层面数字创新指数呈现出地区收敛态势，区域发展不平等水平下降；第四，数字产业创新创业具有地理集聚度，相邻城市发展存在空间关联性；第五，南北区域间差异对于发展不平等水平的贡献度不断提高。

**关键词：**数字经济 创新创业 企业大数据 动态演变 空间格局

## 一、引言

进入 21 世纪，随着互联网的普及以及数字技术的突破，数字经济飞速发展，促进资源广泛重组（江小涓，2017），在全球范围内成为重要的资源要素与发展新动能（Goldfarb & Tucker，2019）。当前，数字经济已经成为中国经济增长的新引擎（戚聿东等，2020），是推动国家经济发展的重要力量（许宪春、张美慧，2020）。根据许宪春和张美慧（2020）估算，2017 年中国数字经济增加值为 5.3 万亿元，占国内生产总值的 6.46%；据中国信息通信研究院（2021）估算，2020 年我国数字经济总规模为 39.2 万亿元，占国内生产总值比重为 38.6%，增速高达 9.7%，其中数字产业化规模达 7.5 万亿元，产业数字化规模超过 31 万亿元。数字经济的持续稳定增长，能有效推动构建双循环新发展格局，提高中国的国际竞争力（郭凯明等，2020）。

作为数字经济的基础，数字经济核心产业<sup>①</sup>（后文简称数字产业）的创新创业行为对数字经济发展起到了重要的推动作用。企业是现代经济发展的微观主体，是创新生态系统的核心（Glaeser et al，2015），企业的创新活动推动了经济的持续增长（Baumol，1996）。熊彼特式的企业家精神通过创业创新，探索新的商业机会，并在创业过程中不断发现新技术、发明新产品，从而推动全要素生产率

<sup>\*</sup> 戴若尘、王艾昭、陈斌开，中央财经大学经济学院，邮政编码：102206，电子邮箱：r.dai@cufe.edu.cn, wangaizz@126.com, chenbinkai@cufe.edu.cn。基金项目：国家社会科学基金重大项目“实施扩大内需战略同深化供给侧结构性改革有机结合研究”（21ZDA034）；国家自然科学基金项目“商事制度改革成效——基于企业微观视角的研究”（71874008）；国家自然科学基金项目“大数据环境下面向宏观经济风险的审计监测预警研究”（71950011）。感谢匿名审稿专家的宝贵意见，感谢北京大学企业大数据研究中心的数据支持，感谢黄嘉敏和马燕来的助研工作，文责自负。

<sup>①</sup> 根据国家统计局发布的《数字经济及其核心产业统计分类（2021）》，数字经济核心产业是指为产业数字化发展提供数字技术、产品、服务、基础设施和解决方案，以及完全依赖于数字技术、数据要素的各类经济活动。

的提高(张维迎、盛斌,2014),创新与创业行为均是企业家精神的具体体现。因此,数字产业的创新创业活动,能够保持中国数字经济的高速增长,推动数字经济和实体经济深度融合,对中国经济未来的高质量发展有着至关重要的意义。

随着数字经济重要性日益凸显,科学准确地描绘数字产业的创新创业情况具有很强的理论和现实价值。本文立足于数字产业的企业家、资本与技术三大核心要素,利用独有的反映中国数字产业中企业创新创业行为的企业大数据,创新性地构建了综合反映中国跨时间跨区域数字产业创新创业发展情况的中国数字经济核心产业创新创业指数<sup>①</sup>(IRIEDEC,后文简称“数字创新指数”),并总结描述2010—2020年中国数字产业创新创业水平动态演变过程与空间分布等典型事实。

具体而言,基于2021年国家统计局公布的《数字经济及其核心产业统计分类(2021)》,本文聚焦数字产品制造业、数字产品服务业、数字技术应用业和数字要素驱动业等四个数字经济核心产业的企业创新创业行为。本文利用由北京大学企业大数据研究中心构建的合并企业工商注册、VCPE投资、专利授权、商标注册与软件著作权登记等数据库的企业大数据,关注注册行业属于数字产业领域的微观企业个体,从数字产业的企业家、资金和技术三大核心生产要素出发,制定新建企业数量、吸引外来投资、吸引风险投资、专利授权数量、商标注册数量和软件著作权登记数量六个维度,分析得到2010—2020年各区域的创新创业活动强度,并构建数字创新指数。该指数为区域一年份层面上的平衡面板数据,本文据此描述中国数字产业创新创业发展的动态演变与空间分布等典型事实。

本文构建的数字创新指数聚焦于企业创新创业行为,具有以下三个突出特点:第一,客观性。本指数以企业为核心,将创业与创新结合,关注企业家精神实现的具体结果,客观反映区域数字经济发展的一个重要角度,即数字产业创新创业发展水平。第二,时效性。本指数结合大数据思维和技术,利用覆盖数字产业全部企业的行政记录大数据,刻画了2010—2020年的中国数字产业创新创业水平的变动。第三,多维性。本指数立足于企业家、资金、技术三大核心生产要素,将原有分散的数据有机整合,构建涵盖创业、投资和创新等多维度的综合评价体系,并提供各维度分项指数,便于学者针对性使用。

基于上述构建的数字创新指数,本文从省级层面和城市层面分析中国数字产业创新创业发展的动态演化和空间分布的典型事实。本文发现:第一,2010—2020年间,中国数字产业创新创业水平快速提高,创新创业绩效显著提升,但2016年之后风险投资、专利授权和软件著作权申请等高质量创新创业指标增长放缓。第二,数字产业创新创业发展呈现“南热北冷、中部崛起”的空间格局,广东省、浙江省等南部省份创新创业绩效持续提高,快速崛起。第三,城市层面的数字产业创新创业发展呈现收敛趋势,低分地区追赶态势明显,但自2018年以来趋势放缓。第四,数字产业创新创业与传统产业相比地理集聚度更高,但由于存在人才和知识的外溢,相邻城市发展存在空间关联性。第五,区域发展的基尼系数逐年下降,主要动力来源于南北区域内部的均等化发展,南北区域之间的差异对于发展不平等水平的贡献度逐年上升。

本文的边际贡献在于:第一,本文基于独有的企业大数据资源,从企业家、资金、技术三大核心生产要素的角度出发构建并公开了数字创新指数,该指数为2010—2020年城市层面的面板数据,包括总指数、分项指数和根据人均指标计算的指数,为学者进一步研究中国数字产业创新创业发展的相关问题提供了数据支持。第二,基于构建的指数,本文首次总结了中国各区域数字产业创新创业发展动态演化和空间分布的典型事实,发现当前数字产业创新创业发展呈现“南热北冷、中部崛起”的空间格局。第三,通过系统比较数字产业与传统产业在空间分布上的特点,本文首次提出中国数字产业创新创业的发展具有地理集聚度高、区域内部空间关联度高、区域发展不平衡问题突出的特点,

<sup>①</sup>中国数字经济核心产业创新创业指数(IRIEDEC)发布平台为北京大学企业大数据研究中心(www.cer.pku.edu.cn),数据下载DOI地址为“<https://doi.org/10.18170/DVN/EYCVSP>”。

并通过基尼系数分解进一步发现我国南北区域数字产业创新创业水平的差异对于发展不平等的贡献度正逐年上升。

## 二、文献综述

作为一种新兴的经济形态，数字经济通过数字产业化和产业数字化两条路径推动了经济持续高质量发展。当前学者围绕数字经济与数据要素在国民经济核算(续继、唐琦,2019)、数据价值测度(许宪春等,2022)、税制结构设计(蒋震等,2021)、数据资产定价与交易(熊巧琴、汤珂,2021)、期权价值(戚聿东等,2019)等方面开展了基础性研究。

数字经济是赋能结构优化和效率提升的重要经济形态,对我国区域数字经济发展水平进行准确测度并探讨其影响已经引起众多专家学者的重视。大多数学者尝试利用统计数据,在省级层面或通过计算相关企业比重(赵云辉等,2019)、相关企业进入退出率(毛丰付、张帆,2021)和自动化岗位数(王林辉等,2020)等数据,或通过构建工业机器人密度(孙早、侯玉琳,2019)、互联网发展水平(韩先锋等,2019)和数字经济发展程度(万晓榆、罗焱卿,2022)等指标,度量了区域数字经济发展,并进行了实证检验。在城市层面,由于数据限制,学者通常利用城市层面的互联网用户数量、宽带用户数量等指标(黄群慧等,2019),或者利用“宽带中国”(赵涛等,2020)、外卖平台兴起(莫怡青、李力行,2022)等冲击间接反映地区的数字经济综合发展情况。柏培文和张云(2021)在上述文章基础上加入了数字产业相关的专利授权数目,考虑了数字经济相关的创新行为。也有文献在指标体系中纳入了数字经济相关企业数量与注册资本指标,从企业数量与规模的角度测算了区域数字经济发展水平(柏培文、喻理,2021);此外,也有少量基于行业信息技术渗透度(杨飞、范从来,2020)和基于企业调查数据中企业信息化程度的研究(何小钢等,2019)。

基于调查和大数据分析等手段,官方和业界智库利用各自数据优势从不同角度构建了一系列反映中国数字经济发展水平的指数。中国信息通信研究院(2017)编制的数字经济指数(DEI)和赛迪顾问(2020)编制的中国数字经济发展指数(DEDI)基于官方统计数据 and 调查数据,从互联网基础、数字产业、营商环境等角度在全国和省级层面度量了数字经济发展水平。业界往往利用自身大数据优势,从不同角度测度了我国区域数字经济发展水平,如新华三集团数字经济研究院(2020)从数字化城市发展与治理的角度构建了城市数字经济指数;腾讯研究院(2020)发布的数字中国指数分不同的终端应用场景,从数字产业、文化、政务、生活四个分指标综合评价城市层面的数字技术应用与发展水平;财新智库和数联铭品(2021)利用大数据联合构建了“中国数字经济指数”,从产业发展、产业融合、产业溢出、基础设施四部分测算了全国和省级层面的数字经济发展水平。部分学者基于业界构建的指数探究了数字经济发展对实体经济(姜松、孙玉鑫,2020)、企业创新(胡山、余泳泽,2022)等方面的影响。

综合来看,尽管已有研究利用数据从多角度测度了我国各区域的数字经济发展水平,但是现有研究存在一定的不足之处:第一,部分指数使用的不同子维度之间可比性较差,同时数据变异频率差异较大,聚合度较低;第二,指数侧重于要素的投入而非产出,要素投入往往存在时间滞后性,短期内无法实现相对应的产出增长,指数结果具有测量误差;第三,部分指数从单一维度对中国数字经济发展程度进行测度,无法较为全面地衡量数字经济发展;第四,鲜有文献从企业创新创业的角度对我国数字产业发展水平进行细致刻画。

为弥补现有研究不足,本文将视角转向企业,从新建企业数量、吸引外来投资、吸引风险投资、专利授权数量、商标注册数量和软件著作权登记数量六个维度创新性地构建出具有区域一年份跨度的中国数字经济核心产业创新创业指数。数字创新指数不仅能够帮助学者们精准判断中国数字产业的发展水平,也能辅助学者验证理论模型的正确性、估计模型参数以及评估不同产业政策的效果(陆毅、孙天阳,2021)。基于该指数,本文从创新创业的角度探究了区域数字产业的发展情况,首次总结了其时空演化的典型事实,并进一步与传统产业对比,分析数字产业创新创业的空间分布特点。

### 三、数据来源、指标计算与指数构建

#### (一)原始数据的收集与清理

本文将视角聚焦于企业,结合大数据思维与技术,运用中国工商企业注册数据库中数字产业的企业基本信息,并合并企业股东数据库、VCPE数据库、专利数据库、商标数据库和软件著作权数据库,利用企业创建、投资行为和创新产出三方面的数据,尝试测算各地区数字产业的创新创业发展水平。为此,本文搜集相关数据库并进行数据清理<sup>①</sup>,以获取数字创新指数的分析指标。下文介绍核心使用的数据和指标体系:

1. 数字经济产业分类。国家统计局于2021年公布了《数字经济及其核心产业统计分类(2021)》,将数字经济产业范围确定为:数字产品制造业、数字产品服务业、数字技术应用业、数字要素驱动业和数字化效率提升业五个大类。其中,前四大类为数字产业化部分,属于数字经济核心产业;第五大类为产业数字化部分,是数字技术与实体经济的融合。

本指数关注数字经济的四大核心产业,即数字产品制造业、数字产品服务业、数字技术应用业和数字要素驱动业。第五大类数字化效率提升业属于产业数字化部分,涵盖了所有统计局认定能够与数字经济结合的传统产业,但由于目前仍缺乏测度数字化改革程度的指标,从产业分类的角度并不能识别出产业数字化的发展程度,因此,指标测算暂时不包含该产业。

本文将四位国民经济分类代码与《数字经济及其核心产业统计分类(2021)》中的六位数字经济代码进行匹配,基于中国工商企业注册数据库的行业信息筛选出属于前四类数字产业的微观企业,这些企业提供的产品和服务是数字经济的直接产出,也是产业数字化发展的基础投入。

此外,数字要素驱动业大类中信息基础设施建设、数据资源与产权交易和其他数字要素驱动业三个中类内部混杂了一些无法归类的传统产业企业,将其纳入指数计算范围会影响结果的准确性。加之上述三个中类在数字产业创业创新的占比极小,因此,本文在计算过程中暂时将属于上述三个分类的企业样本剔除,保留了其他属于数字产业的样本。

2. 中国企业工商注册数据。企业工商注册数据包含了1990年至今所有在中国境内注册的企业经营单位。本文使用的企业数据涵盖有限责任公司、股份有限公司、合伙企业、个人独资企业和内资企业法人五种类型的企业样本。对于数字产业的样本,本文按照企业的注册年份和注册地址行政区划代码(2020年版本)对数据进行加总处理,得到各地区每年在数字产业领域的新建企业数量。该数据能够真实刻画各地区各产业的企业创建行为。在处理过程中,本文对于注册地址、日期等相关变量进行了数据清理<sup>②</sup>。

3. 外来法人投资数据库和 VCPE 数据库。利用中国企业工商注册数据中的企业股东信息,本文筛选出法人股东,并链接法人股东的注册信息,从而构建法人投资数据库。通过比对被投企业的注册地址和投资该企业的法人股东的地址信息,本文能够判断出企业是否获得外来法人投资。对于数字产业获得外来法人投资的企业样本,本文按照投资年份和获得外来投资企业的注册地址对数据进行加总处理,得到各地区每年在数字产业领域获得外来投资的笔数。

北京大学企业大数据研究中心整合中国证券基金协会公示的“私募股权、创业投资基金管理人”名单和清科数据,结合企业股东信息,通过股权穿透识别投资基金及其投资企业,构建了包含基金管理人基本注册信息和其投资企业的中国 VCPE 投资数据库。基于该数据库,本文将样本限制在数字产业领域,按照投资年份和获得风险投资企业的注册地址对数据进行加总,获得各地区数字产业领域每年得到风险投资的笔数。

4. 专利、商标、软件著作权数据。北京大学企业大数据研究中心将国家知识产权局公布的微

<sup>①</sup>数据处理基于北京大学企业大数据研究中心平台。

<sup>②</sup>篇幅所限,此处具体处理过程未予展示,数据清理文档等在线附录详见 <https://www.cer.pku.edu.cn/tab/zhishu-yan-jiu/iriedec>。

观专利申请数据与企业工商注册数据进行匹配,用以度量企业的创新产出情况。由于发明专利从申请到公开到授权的期限过长,专利数据只有公开后才会更新,为了保证指数的时效性,本文只考虑被授权的专利。通过区分专利授权的年份与企业的注册地址,本文对数据进行加总处理,获得了各地区每年数字产业领域企业的发明专利、实用新型专利和外观设计专利三种类型专利的授权件数。同时,被授权的专利质量更高,可以直接帮助企业在市场提供新的产品与服务,与本指标的“度量创新产出而非投入”的设计思路一致。

数据中还整合了同样来自国家知识产权总局的商标注册数据以及软件著作权登记数据,按照相同的处理步骤,将二者与企业工商注册数据相匹配,进行加总,得到各地区每年数字产业领域的商标注册数量与软件著作权登记数量。

## (二)数字创新指数指标体系

通过合并上述数据库,本文从企业家、资金和技术三个数字产业创新创业发展核心要素出发,构建了包含新建企业数量、吸引外来投资、吸引风险投资、专利授权数量、商标注册数量和软件著作权登记数量六个维度共计八个指标的指数体系,用以测算中国数字产业创新创业发展水平。其中,新建企业数量、吸引外来投资和吸引风险投资三个维度的指标从创业角度刻画了各区域每年数字产业的发展水平,专利授权数量、商标注册数量和软件著作权登记数量三个维度侧重于测度每年各区域数字产业的创新产出水平。下文详细阐述上述六大维度的具体含义及使用的具体指标:

1. 新建企业数量。新建企业数量从企业家角度考察了各区域数字产业创新创业活动绩效,参考叶文平等(2018)和赵涛等(2020)的做法,本文选取各区域每年新建企业数量这一指标来衡量区域数字产业的创业活跃水平。企业注册意味着企业家精神的实践,区域新增注册企业的数量能够表征一个地区数字产业企业家要素的增加值。新进入的企业家能够提高区域的市场竞争,促进“优胜劣汰”,实现熊彼特式“创造性毁灭”与资源重新分配,因此,对于创新创业和宏观上的产业转型升级十分重要。新企业的进入也可以反映当地数字经济的活力与区域营商环境的优劣。本维度的具体指标为区域每年数字产业的新增企业注册数量。

2. 吸引外来投资。招商引资能够通过吸引更高质量的企业投资拉动地区就业,并实现知识技术的外溢,从而带动本地产业的发展。外来投资者在投资时也会评估地区的潜能与营商环境,吸引外来投资的数量对本区域数字经济发展十分重要,特别是外来法人投资者,能够为本地数字产业带来更多的资本要素和先进的管理经验。本文重点关注法人投资行为,通过判断进行投资的法人其注册地址与被投资企业是否处在同一地区,识别出各个区域每年获得的外来法人投资数量。吸引外来投资维度指标为区域每年数字产业新增吸引外来投资的笔数。

3. 吸引风险投资。风险投资通过投资新兴企业并参与经营来影响企业的经营状况(张学勇、张叶青,2016),在推动新企业创建(Samila & Sorenson,2011)、技术创新(Kortum & Lerner,2000;吕炜,2002)、企业经营(吴超鹏等,2012)和效率提升(Chemmanur et al,2011)等方面起到重要作用。由于风险投资公司在投资时会严格审查和评估,投资的衡量指标往往是被投资企业能否最终上市,接受风险投资的企业更加可能成长为瞪羚企业或独角兽企业<sup>①</sup>,王兰芳和胡悦(2017)、Sun et al(2020)的研究也表明获得风险投资支持的企业其创新质量往往更高。与其他类型的投资相比,风险投资在

<sup>①</sup>VCPE在某一地区投资的企业往往拥有资本优势,可能出现无序扩张形成垄断的情况。但本文认为获得风险投资的笔数仍然能够反映该地区的创新创业水平,原因如下:第一,企业数目偏少、存在超大型企业的垄断结构垄断并不一定是由于地方保护制度扭曲而产生的实质性垄断,数字产业一直面临着高度的来自新企业进入的竞争压力,如在电商和社交领域不断出现后来者,孙宝文等(2017)认为互联网行业虽存在垄断的结构但并未出现实质性的垄断行为;第二,数字产业内部企业提供的产品和服务具有贸易成本低的特点,服务区域往往不局限于本地,地方保护主义等本地制度扭曲并不能决定数字经济企业的规模,二者相关性不强;第三,从实证角度,吸引风险投资子维度得分与能够反映营商环境的樊纲市场化进程总得分相关系数为0.69,系数在1%水平上显著,表明垄断与区域营商环境相关性较弱。

能够给企业带来更为丰富的资本要素的同时,能够通过助力企业研发(Da Rin & Penas, 2007)、为企业提供行业经验(陈思等, 2017)等途径推动企业创新。本文认为吸引风险投资数量是测算该地区当前数字产业创新创业活跃程度的重要指标。吸引风险投资维度的具体指标为区域在数字产业每年新增吸引风险投资的数量。

4. 专利授权数量。专利代表了经济中技术要素的引入与发明。张杰等(2016)认为与专利申请数量相比,专利授权数量可以避免虚假、不合格的申请专利的存在以及“创新假象”问题,更能反映地区的创新能力,因此,本文选取专利授权数量而非专利申请数量作为衡量地区创新绩效的指标。专利分为发明专利、实用新型专利和外观设计专利三类,其中,发明专利需要经过实质审查,而实用新型专利和外观设计专利只要经过形式审查(龙小宁、王俊, 2015),因此,发明专利质量更高(邵朝对等, 2021),与另外两种专利相比更能体现创新能力(张杰、郑文平, 2018)。而实用新型专利创新程度介于外观设计专利和发明专利之间(毛昊等, 2018)。考虑到专利质量具有异质性,本文以权重高低来表征不同类型专利的质量,将三种专利在专利维度中的相对权重依次设定为 5 : 2 : 1。将中国知识产权总局的专利数据与企业工商注册数据匹配,依据企业注册地信息、授权年份和专利类型对数据进行加总处理,本文能够获得各区域数字产业中所有存续企业在每一年的专利授权数量。专利授权数量维度共包含地区数字产业新增发明专利授权数量、新增实用新型专利公开数量和新增外观设计专利公开数量三个指标。

5. 商标注册数量。注册商标能够间接度量企业的产品创新与产品质量。参考 Hu et al(2021)使用出口占 GDP 的比重、注册商标数、专利申请数和通过一级及以上质量检查的商品抽样比例四个指标来衡量我国制造品质量的做法,本文认为企业注册商标的行为能够反映其品牌意识与产品创新的质量。本文将来自中国知识产权总局的商标数据与企业工商注册数据进行匹配,并按照商标注册时间和区域对数据进行加总,获得区域一年份层面数字产业中所有存续企业的商标注册数<sup>①</sup>。商标注册数量维度指标为各地区数字产业新增商标注册数量。

6. 软件著作权登记数量。软件是数字产业赋能产业数字化的关键载体,一个地区软件著作权的登记数量能够反映该地区企业在软件开发领域的创新成果,特别是数字技术应用业企业的创新产出,因此,软件著作权是衡量数字产业内部企业创新行为的有效指标。本文将来自中国知识产权总局的软件著作权登记数据与企业工商注册数据进行匹配,按年统计各地区数字产业中所有存续企业的软件著作权登记数。软件著作权登记数量维度由各地数字产业的新增软件著作权登记数量这一指标来度量。

### (三)各指标原始值的描述统计

本节首先描绘了 2010—2020 年中国各维度指标原始值的动态演变过程以及占全国总数的比重<sup>②</sup>。

图 1 展示了数字创新指数各分项指标原始值的时间趋势与占全国总数的比重。如图 1 所示,总体来看,2010—2020 年间,新建企业数量、吸引外来投资、吸引风险投资、专利授权数量、商标注册数量和软件著作权登记数量六个维度的指标大体呈现出逐年递增的趋势。特别是在 2014 年商事制度改革后,多项指标在绝对数目上实现迅速增长。具体而言,图 1(a)反映了新建企业数量这一指标的动态演化过程,2010 年数字产业新建企业为 70249 家,2020 年这一数量已经增长为 604888 家,增幅

<sup>①</sup>数据显示,数字产业内部企业平均每年申请商标数量为 2.46 件,申请商标数量的 99%分位数为 21 件,平均值超过 4 件的行业有广播、电信和非金融机构支付服务三行业。由于数据中不存在反映商标质量的变量,为了验证数据并未存在商标“注水”问题,本文删除申请商标数目在 99%分位数以上的企业,重新构造区域一年份层面的商标申请数量数据,结果表明新旧两种数据的相关系数达到 0.99,表明一些企业进行虚假商标注册的行为并没有令本文结果产生重大偏差。

<sup>②</sup>在探讨指标原始值的时间趋势时,将专利授权数量维度的发明专利授权数量、实用新型专利公开数量和外观设计专利公开数量三个子指标数量进行加总处理。

高达 8.61 倍。与此同时,数字产业的创业热度持续上升,新建企业数量占当年全国总数的比例也在逐年增加,从 2010 年的 4.56% 提升至 2020 年的 7.97%。2017 年数字产业新建企业数目占全国总数的比例为 7.11%,这与许宪春和张美慧(2020)估算的 2017 年数字经济增加值占 GDP 的比例 6.46% 十分接近,说明若从企业家角度进行观测,数字产业的创业水平能够部分反映数字经济的发展规模。从资本要素的角度,2020 年数字产业共获得外来投资 23504 笔,占当年全国总数比例为 11.57%;而 2010 年数字产业内部获得外来投资 1305 笔,占比仅为 5.19%。2020 年数字产业获得风险投资共 1414 笔,占全国总数比重约为 12.29%。从创新产出的角度,对于专利授权维度,2020 年三种专利授权共计 254072 件,占比约 10%;数字产业的商标注册共计 442681 件,占比约 10.86%;数字产业软件著作权登记数量提升明显,2020 年总数已突破 44 万件,占全国总数比重约为 29.81%。

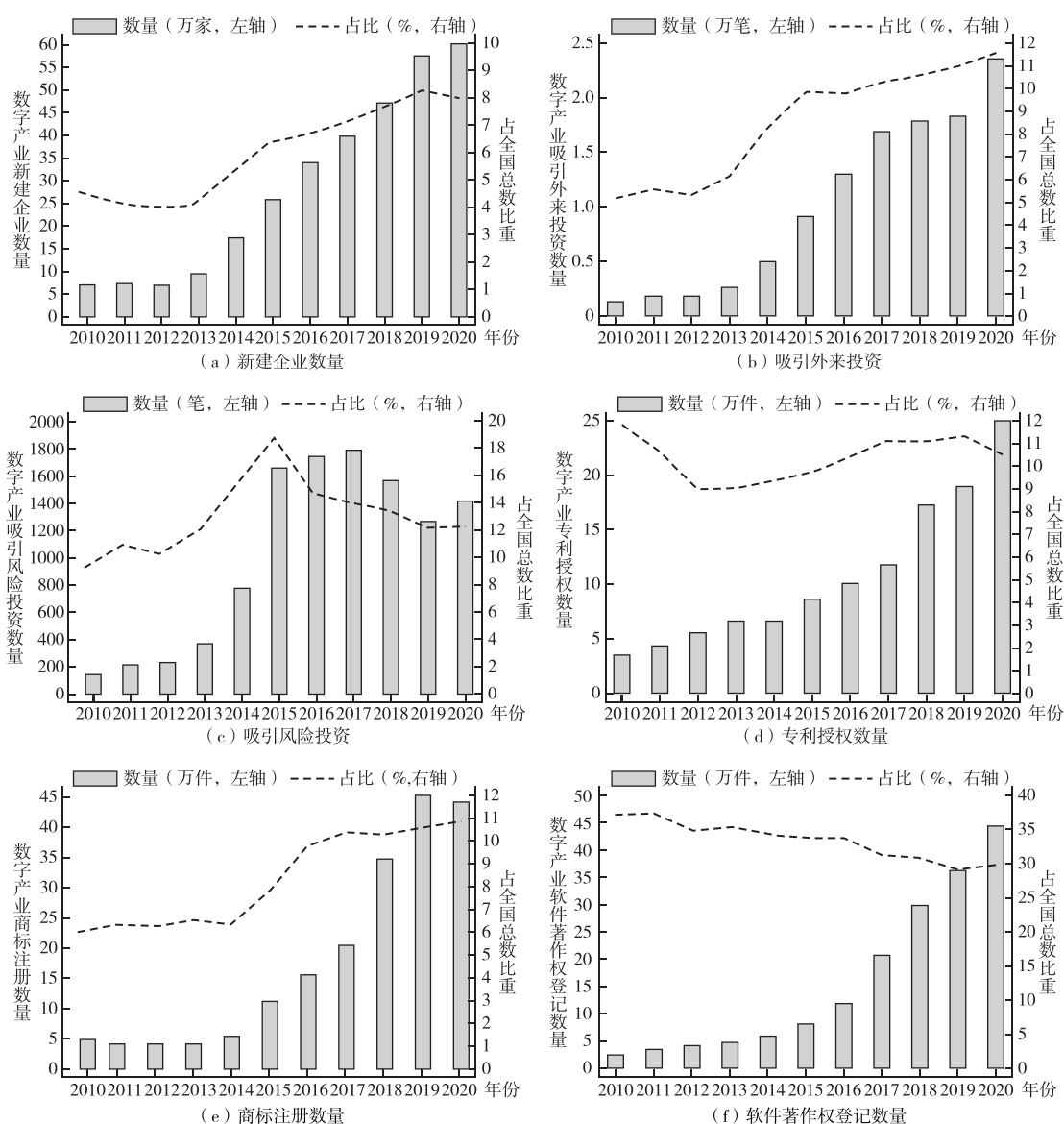


图 1 数字创新指数各分项指标原始值的时间趋势

注:作者通过全国工商企业注册数据库、VCPE 投资数据库、专利数据库、商标注册库以及软件著作权库数据计算得出。

值得注意的是,对于新建企业数量和吸引外来投资两维度的指标,数字产业占比在过去十年间保持着稳定的增速,新建企业数目占比从 2010 年的 4% 增长到当前的 8%,吸引外来投资占比从 5%

上升到11%。而 VCPE 投资和商标注册二维度的指标,数字产业占比在 2010—2015 年之间虽保持快速增长,但之后却逐渐失去动能:VCPE 投资占比从 2015 年最高的 18%下降当前的 12%,商标注册占比在 2016 年之后维持在 11%左右。专利授权和软件著作权登记两项指标的占比在这十年间没有增长,专利授权占比维持在 10%~11%之间,软件著作权登记占比从 2010 年的 35%缓慢下降到当前的 30%。当前我国数字产业创业数量激增,但对于 VCPE 投资、专利授权等更侧重于反映相对高质量创新创业活动的指标,近年增长趋缓。

(四)数字创新指数构建

在明确数字创新指数的指标体系和原始值的变化趋势后,本文首先进行了 KMO 检验以比较不同维度指标之间的相关性。本文六个维度的八个指标 KMO 值为 0.91,满足进行主成分分析的条件。根据计算可知,第一主成分的特征值为 7.09,对总方差贡献率达到 88.62%。可见,数字创新指数的指标体系具有一定的合理性。

在确定指标体系的合理性后,本文对各指标进行标准化(无量纲化)处理,给予各指标相应权重后进行加权平均计算。关于权重的选择,本文基于“中国区域创新创业指数”的指标权重(戴若尘等,2021),并结合数字产业创新创业的特点,沿用其关于新建企业数量、吸引外来投资和吸引风险投资三个指标的权重,适度调低实用新型专利、外观设计专利和商标注册三个指标的权重,并对新增的软件著作权指标设定权重。因此,本文计算的数字创新指数可以与区域创新指数进行对比,本文在后续章节会比较数字经济与传统经济的差异,从而分析数字产业的特殊性。具体的权重设定如表 1 所示,新建企业数量的权重为 20%,吸引外来投资所占权重为 15%,吸引风险投资所占权重为 25%,专利授权数量的权重为 20%(区域创新指数为 25%),商标注册数量的权重为 10%(区域创新指数为 15%),软件著作权登记的权重为 10%。专利授权数量维度中的三个分项指标发明专利授权数量、实用新型专利公开数量、外观设计专利公开数量权重设置为 5:2:1,占总指数的权重分别为 12.50%、5%和 2.50%(区域创新指数为 12.50%、7.50%和 5%)。

表 1 中国数字创新指数指标及权重

维度名称	分项指标	所占权重
新建企业数量	新增企业注册数量	20.00%
吸引外来投资	新增外来法人投资的笔数	15.00%
吸引风险投资	新增风险投资的企业数量	25.00%
专利授权数量	新增发明专利授权数量	12.50%
	新增实用新型专利公开数量	5.00%
	新增外观设计专利公开数量	2.50%
商标注册数量	新增商标注册数量	10.00%
软件著作权登记数量	新增软件著作权登记数量	10.00%

注:维度、指标及其权重为作者设计。

在确定权重后,本文将各指标数据加 1 取自然对数后进行标准化处理得到标准值,根据标准值确定各区域一年份层面指标的分位数,将分位数乘以 100 得到各分项指数得分。在计算总量得分时,本文将各指标的标准值与表 1 的权重相乘,得到区域一年份层面数字创新指数的标准值,通过计算分位数并乘以 100,最终得到数字创新指数总量得分。

本文使用固定权重进行分项指标的加总,而非使用主成分分析法来确定权重,原因如下。首先,作为一套具有区域一年份跨度的面板数据,数字创新指数能够逐年更新,精准刻画我国各区域数字产业创新创业绩效。但采取主成分分析法确定的权重可能会随时间发生变动,造成指标权重分配不稳定(樊纲等,2011),而固定权重在实际应用中能够保证指数跨年可比。为了进一步判断本文所设定的权重是否合理,参照樊纲等(2011)的比较方法,本文使用主成分分析法提取的系数与固定权重同时对数字创新指数进行测算,结果显示两种方法测算的指数得分与排名相关系数皆达到 0.99,这



一结果表明采取固定权重计算指数具备一定的合理性。

其次，与主成分分析法相比，本文所设定的权重优势在于其能将创新创业质量纳入考虑范围。由于创新创业行为并不是同质的，在对各项指标赋权时，本文同时将企业规模、创新质量等因素纳入考虑，将表征高质量创新创业的指标赋予更大的权重。例如，表 2 展示了不同类型的企业注册资本的平均值与标准差，数据显示获得外来法人投资的企业注册资本均值达到 8172.3 万元，获得 VCPE 投资的企业注册资本平均超过 1 亿元，远超自然人投资企业与本地法人投资企业的注册资本均值。相比其他两种类型的企业，获得 VCPE 投资的企业和外来法人投资的企业具有更丰富的资本要素，因此，从资金角度对企业的创新创业行为进行考察时，本文对吸引外来投资和吸引风险投资赋予了较高权重（分别为 15% 和 25%），从而反映规模层面的差异。如果只考虑新企业进入，或者按照主成分分析法得到的权重构建指数，则无法反映某些地区由于进入资本规模更大带来的更高水平的创新创业。又如，在专利授权维度，一般而言发明专利的创新能力更高，实用新型专利次之，外观设计专利技术门槛相对较低（李兵等，2016；毛昊等，2018）。但由于数据限制，本文缺乏能够直接测度专利质量的指标，若按照主成分分析法所得的接近 1 : 1 : 1 的权重进行加权则不能体现三种专利的差异性。因此，在设定权重时，本文充分考虑了专利质量的异质性，将三种专利的权重之比定为 5 : 2 : 1，以高权重表征高创新质量。

表 2 不同类型的企业的注册资本(单位:万元)

企业类型	平均值	标准差
自然人投资企业	345.3	3970.4
本地法人投资企业	4636.9	103841.4
外来法人投资企业	8172.3	181477.6
VCPE 投资企业	11822.4	162422.0

注：作者基于中国企业工商注册数据计算得出。

为检验本文构建的指数是否有效，本文使用当前文献常用的衡量数字经济发展的指标，包括互联网宽带用户数、移动电话用户数、电信业务收入 and 信息技术从业人员数（黄群慧等，2019），探索本文构建的数字创新指数与上述指标的相关关系。作为一项新型基础设施，互联网是数字经济发展的重要平台，其广泛应用能够起到提升创业概率（周广肃、樊纲，2018）和促进创新要素流动的作用（黄群慧等，2019），因此，本文认为互联网发展程度应与数字创新指数存在相关关系。如表 3 所示，以互联网宽带用户数、移动电话用户数、电信业务收入、信息技术从业人员数四个指标的对数值为自变量，以城市层面的数字创新指数对数值为因变量，本文控制了城市与年份的双向固定效应进行指标间的相关性分析。参考赵涛等（2020），本文在回归中控制了以下变量：使用当年实际使用外资额与地区生产总值之比代理外资占比；使用财政预算内收入与财政预算内支出之比代理政府财政自主权；使用人口密度的对数来代理当地城市化水平；使用金融机构存贷款余额与地区生产总值之比代理当地金融发展程度；使用当年地区生产总值的对数代理当地经济发展程度。如第(2)(6)列所示，表征互联网发展与产出的互联网宽带用户数量、电信业务收入两个变量与数字创新指数均具有正相关关系，且分别在 1% 和 5% 的水平上显著。移动电话用户数目的系数虽然为正却不显著，其原因可能在于：移动电话的普及率水平较高，跨期变化差异不大。而信息技术从业人员数量的增长并未促进数字产业创新创业水平的提高，反而存在负向相关关系，本文认为这反映了其他相关的数字产业（如硬件制造、自媒体和电商等非计算机和软件行业）发展占比较大，忽略这些产业可能会导致总体发展水平的衡量偏差。综上，数字创新指数与文献现有的衡量数字经济发展程度的指数中反映数字基建的指标存在较高的相关关系，与移动电话用户和计算机软件行业规模的相关性较弱。数字创新指数能够反映所有数字产业中企业的创业创新行为，与现有指标存在一定的互补性。

表3 数字创新指数与其他衡量数字经济发展指标的相关关系

变量	数字创新指数对数							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
互联网宽带用户数对数	0.1160*** (0.0323)	0.1110*** (0.0325)						
移动电话用户数对数			0.0277 (0.0860)	0.0135 (0.0867)				
电信业务收入对数					0.0633** (0.0263)	0.0588** (0.0257)		
计算机服务和软件业从业人员数对数							-0.0731*** (0.0275)	-0.0701** (0.0277)
外资占比		-0.1552** (0.0707)		-0.1514** (0.0722)		-0.1431** (0.0714)		-0.1461** (0.0691)
财政自主权		0.0899 (0.1352)		0.1130 (0.1366)		0.0859 (0.1361)		0.1167 (0.1374)
城市化水平		-0.3898 (0.2946)		-0.3905 (0.2957)		-0.4057 (0.2964)		-0.3452 (0.2915)
金融发展程度		-0.0321 (0.0249)		-0.0395 (0.0250)		-0.0388 (0.0253)		-0.0369 (0.0240)
经济发展程度		-0.0225 (0.1732)		0.0230 (0.1747)		-0.0009 (0.1729)		0.0471 (0.1720)
年份固定效应	控制	控制	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	控制	控制
常数项	2.0357*** (0.4221)	4.9638 (4.6779)	3.1351** (1.2850)	5.0515 (4.7204)	2.1804*** (0.5700)	4.6927 (4.7225)	4.1598*** (0.2291)	4.9509 (4.6462)
观测值	1599	1599	1599	1599	1599	1599	1599	1599
R <sup>2</sup>	0.8665	0.8672	0.8654	0.8662	0.8660	0.8667	0.8660	0.8668

注:自变量与控制变量数据来自《中国城市统计年鉴》,覆盖年份为2010—2016年。括号中为稳健标准误,\*\*\*、\*\*和\*分别表示在1%、5%和10%的水平上显著。

#### 四、省级层面数字产业创新创业发展的动态演化与空间分布

##### (一)中国数字产业创新创业发展的动态演化和空间分布

根据测算的数字创新指数,本文发现我国数字创新指数得分总体水平呈现逐年增加的趋势。表4展示了2010—2020年中国各省份数字创新指数得分的描述性统计(因数据限制,未测算我国港澳台地区,下同)。如表4所示,2010年省级指数平均得分仅为26.65分,中位数得分为18.48分,尽管平均得分在2019年有些许下降,但十年来依旧能够观察到明显的增长趋势。2020年,省级得分平均值达69.70分,是2010年的2.62倍,中位数得分也增长至79.77分。

表4 数字创新指数省级得分描述性统计

年份	平均值	标准差	最小值	中位数	最大值
2010	26.65	22.46	0.29	18.48	82.70
2011	29.79	24.17	0.88	22.58	86.51
2012	31.03	23.88	1.47	24.63	85.63
2013	36.11	24.98	2.05	31.67	90.91
2014	44.50	26.27	4.40	43.70	95.89
2015	53.87	26.18	5.87	56.30	97.65
2016	60.09	24.72	10.85	60.70	98.53
2017	64.49	23.25	11.44	66.57	99.71
2018	67.75	22.66	13.20	73.90	99.41
2019	67.64	24.48	12.61	76.54	99.12
2020	69.70	23.99	20.53	79.77	100.00

数据来源:数字创新指数。下同。

表 5 展示了 2020 年中国各省份数字创新指数得分及排名。2020 年,广东省、浙江省、山东省、四川省和上海市得分均超过 95 分,是我国数字产业创新创业的领军省份,创新创业实力强劲;而尾部省份则表现不佳,内蒙古自治区、青海省与西藏自治区得分均未超过 30 分,实力有待增强。当前头部省份与尾部省份发展差异明显,例如,广东省排名第一,得分是数字创新指数得分最低的西藏自治区的 4.87 倍。

表 5 2020 年中国省级数字创新指数得分及排名

省份	得分	排名	省份	得分	排名	省份	得分	排名
广东	100.00	1	湖南	83.58	12	新疆	52.79	23
浙江	98.83	2	江西	82.99	13	黑龙江	50.44	24
山东	97.07	3	海南	82.40	14	辽宁	48.97	25
四川	96.48	4	河南	81.52	15	吉林	48.68	26
上海	95.60	5	河北	79.77	16	甘肃	38.71	27
福建	92.96	6	陕西	77.42	17	宁夏	34.02	28
北京	90.03	7	贵州	68.04	18	内蒙古	27.86	29
安徽	87.68	8	云南	66.28	19	青海	21.99	30
湖北	86.22	9	山西	65.98	20	西藏	20.53	31
江苏	85.04	10	天津	58.06	21			
重庆	84.16	11	广西	56.60	22			

在观测到数字产业创新创业水平迅速提升的同时,本文发现当前我国数字产业创新创业发展空间格局呈现出“南热北冷、中部崛起”的分布特征。具体而言,本文发现我国南北地区<sup>①</sup>的发展程度出现分化,平均而言南部省份发展程度优于北部省份。例如,以广东省、浙江省、四川省和上海市为代表的南部省份数字产业创新创业实力强劲,在 2020 年全国得分前五强中共占据四个席位,而北部地区仅山东省和北京市创新创业绩效较优,得分超过 90 分;从低分省份来看,广大北部地区发展潜力仍需激活,全国得分未超过 50 分的省份共计七个,除西藏自治区外其余省份都位于北部地区。除可观测到明显的南北差异外,数字产业创新创业发展空间格局同样呈现出明显的“中部崛起”的特征。自 2018 年起,以安徽省和湖北省为代表的中部省份得分位居全国前列,至 2020 年,安徽省和湖北省排名进入全国前十,除山西省外,中部五省份得分均超过 80 分,可见中部省份在数字产业领域表现突出,发展潜力巨大。

## (二) 中国数字产业创新创业发展活力:2020 年

进一步,将测算的各分项指数进行加权,本文能够从企业家、资金和技术三大核心生产要素的角度细致描摹 2020 年中国数字产业创新创业发展活力空间格局。基于省级数字创新指数,图 2(a)展示了 2020 年省级数字创新指数得分的具体情况,图 2(b)、图 2(c)和图 2(d)分别从企业进入(新建企业数量)、吸引投资(吸引外来投资和吸引风险投资)和创新产出(专利授权数量、商标注册数量和软件著作权登记数量)三个角度描绘了中国各省份数字产业的创新创业绩效<sup>②</sup>。

如图 2(a)所示,从总量维度来看,广东省、浙江省、四川省、上海市和福建省五个南部省份指数得

<sup>①</sup>本文以秦岭、淮河为界,将中国省份区分为南部地区与北部地区。具体而言,北部地区包括北京市、天津市、河北省、山西省、内蒙古自治区、辽宁省、吉林省、黑龙江省、安徽省、山东省、河南省、陕西省、甘肃省、青海省、宁夏回族自治区、新疆维吾尔自治区;南部地区包括上海市、江苏省、浙江省、福建省、江西省、湖北省、湖南省、广东省、广西壮族自治区、海南省、重庆市、四川省、贵州省、云南省、西藏自治区。由于数据限制,目前南北地区分类暂未包括我国港澳台地区。

<sup>②</sup>本文在描绘 2020 年中国各省份数字产业创新创业绩效时,将数字创新指数包含的六维度指标分为三个角度:企业家(新建企业数量)、资金(吸引外来投资和吸引风险投资)和技术(专利授权数量、商标注册数量和软件著作权登记数量),并按照表 1 的权重对三个角度包含的子维度得分进行重新加权计算(例如在投资角度:吸引投资得分 =  $\frac{15\%}{15\%+25\%} \times \text{吸引外来投资得分} + \frac{25\%}{15\%+25\%} \times \text{吸引风险投资得分}$ )。

分均超过 92 分,数字产业创新创业水平领先全国;北部地区数字产业创新创业发展程度较低,潜力有待开发,除山东省和北京市外,其余北部省份得分均低于 90 分。从企业进入、吸引投资和创新产出三个角度对各省的表现加以考察,南北地区在数字产业创新创业领域的表现同样存在异质性:在企业进入角度,广东省和浙江省是企业注册最为活跃的两个省份,得分均超过 99 分,四川省、福建省和海南省排名四至六名,而北部地区中仅山东省和河北省两省得分超过 90 分,分别为 98.24 分与 90.32 分,排名位于第三和第七,其余北部地区发展活力略显不足;在吸引投资角度,南北地区差距同样较为明显,北部地区发展空间巨大,仅有山东省得分超过 90 分,全国排名第四,而南部省份中浙江省得分为 99.08 分,排名第一,广东省、四川省和上海市得分均超过 93 分,对于投资的吸引能力更强;在创新产出角度,广东省、浙江省和上海市排名前三,创新产出领先全国,福建省与四川省得分分别为 94.01 分和 93.49 分,分别排第六和第七,而北部地区只有北京市和山东省得分超过 90 分,排名第四和第五,与之相对,大部分北部地区发展潜力仍需激活。

除展现出明显的南北差异外,我国数字产业创新创业同样可以观测到中部崛起的空间分布特征。得分显示,安徽省、湖北省、湖南省、江西省和河南省中部五省得分均超过 81 分,排名进入全国前十五名,中部省份正成为数字产业创新创业活动的核心区域。在企业进入角度,江西省、湖南省和河南省三个中部省份对于企业家有较强的吸引力,排名进入全国前十,安徽省和湖北省分别排名第十三、第十四,得分超过 82 分;在吸引投资角度,江西省、安徽省、湖北省和湖南省四省排名第十至第十三名,得分均超过 80 分,获得投资能力较强,数字产业活跃度较高;在创新产出角度,湖北省和安徽省分别排第八和第九,河南省位列第十一,表现较为突出。

综上,数字创新指数显示相比于北部地区,南部省份更具创新创业活力,数字产业创新创业发展程度较高,而除北京市和山东省外,广大北部地区发展相对乏力,创新创业活力亟须激发;安徽省和湖北省等中部省份崛起势头明显,正成为全国数字产业创新创业的核心区域。中国数字产业创新创业已形成“南热北冷、中部崛起”的空间格局。

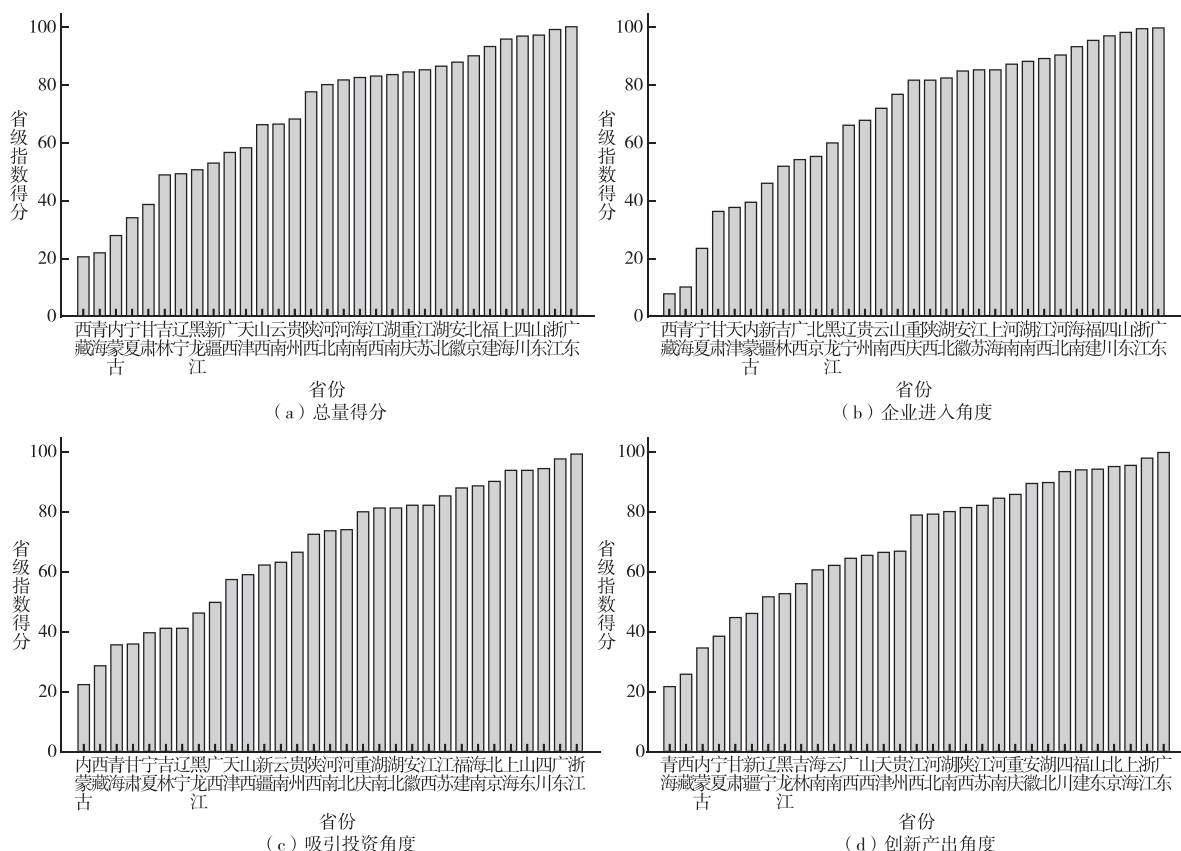


图 2 中国数字产业创新创业发展活力:2020 年

## 五、城市数字产业创新创业的空间格局

中国同一省份内不同城市的发展差异较大,城市层面的指数能够更好地反映数字产业创新创业发展在区域层面的分布特点,同时也能够帮助识别出目前发展水平较高的城市,对政策制定起到辅助作用。本文考虑了城市人口规模的差异,新计算了人均数字创新指数,从城市层面更细致地描绘数字产业创新创业发展的空间格局。本文进一步构建了反映数字产业发展的收敛性、集聚度、外溢性等特征的指标并进行分析,强调数字产业创新创业发展呈现城市群内部协同发展的格局,同时也存在南北区域层面的发展不平衡问题。

### (一) 中国城市层面数字产业创新创业发展格局:2020年

相较于省级指数,市级指数可以对不同区域的创新创业绩效进行更精准的描摹。图3展示了2020年总量与人均层面城市数字创新指数排名前20名的城市<sup>①</sup>。图3(a)为市级总量指数得分情况。在本文测算的指数中,深圳市与杭州市总量得分位列第一和第二,上海市、成都市和北京市排名紧随其后。从空间分布的角度来看,市级得分同样显现出“南热北冷、中部崛起”的空间格局。市级得分前20名中仅有6个城市位于北部地区,其余城市均位于南部地区。若按照指数得分将城市划分为5个得分梯队,南部地区共有41个城市得分位列全国前20%,而北部地区则有26个城市得分进入第一梯队。具体而言,以深圳市、杭州市为代表的珠江三角洲、长江三角洲城市群和成渝地区中部分城市排名靠前,引领全国数字产业创新创业发展;北部地区中北京市、青岛市和济南市等城市同样排名位于前20%,但西北、东北地区的部分城市得分较低,仍有较大的发展空间。中部6省共有14个城市得分位于全国前20%,以武汉市、长沙市和合肥市为代表的中部城市表现优异,推动省份发展,助力形成数字产业创新创业新的增长极。

为了排除人口规模差异对各地创新创业绩效的影响,衡量区域数字产业的创新创业密度,本文进一步在城市层面计算了人均数字创新指数。图3(b)展示了2020年人均数字创新指数排名前20名的城市。在校准城市人口规模的差异后,深圳市和杭州市优势仍然明显,领跑全国数字经济发展,厦门市、东莞市和海口市位于第三至第五名,在人均层面,排名前5的城市皆位于南部地区。若同样按照指数得分,将城市划分为5个得分梯队,与总量指数较为相似,京津冀城市群、山东半岛城市群、中原城市群、长三角城市群、长江中游城市群、珠三角城市群、成渝城市群等多个城市群数字产业创新创业密度较高,部分城市排名位居前列,创新创业绩效突出。

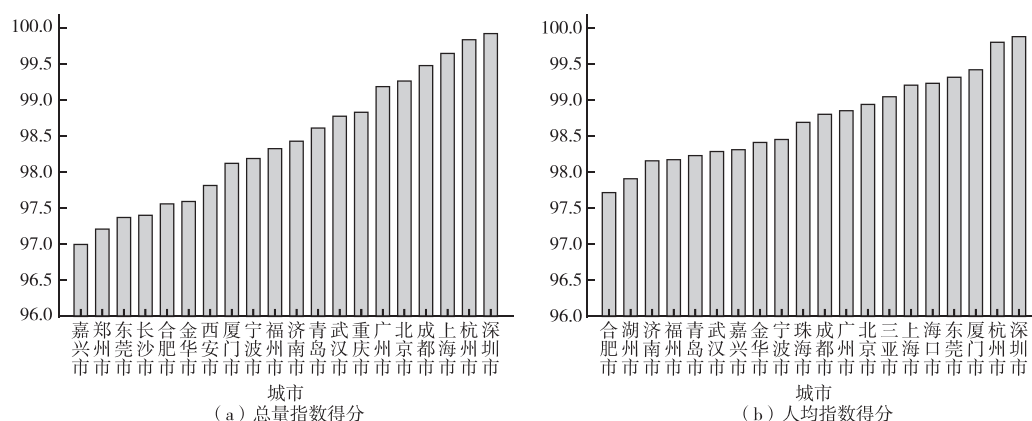


图3 2020年市级数字创新指数得分情况

### (二) 数字产业创新创业发展的地区收敛性

市级数字创新指数得分显示,中国数字产业创新创业水平存在显著的地区差异性。与传统经济

<sup>①</sup>在计算市级得分时,本文将4个直辖市纳入测算框架,与除海南省三沙市外的332个地级市共同排名。篇幅有限,本文在此仅展示了总量指数得分与人均指数得分排名位于前20名的城市。

形式不同,数字经济的发展使得资源、信息的流动能够突破地域的限制,推动企业面向全国的市场。孙杰(2020)认为在数字经济时代跨行业、跨区域的竞争形式将逐渐常态化,数字经济对于区域协调发展与资源优化配置有重要的推动作用。为进一步探明我国区域间数字产业创新创业水平差距的变化趋势,本文应用 $\sigma$ 收敛模型来刻画区域经济发展的动态演化过程。参照郭峰等(2020),本文利用公式(1)进行检验:

$$\sigma_t = \sqrt{\frac{1}{n} \sum_{i=1}^n \left( \ln score_{it} - \frac{1}{n} \sum_{i=1}^n \ln score_{it} \right)^2} \quad (1)$$

其中, $\ln score_{it}$ 代表城市*i*在第*t*年数字创新指数的对数值, $\sigma_t$ 代表第*t*年数字创新指数的 $\sigma$ 收敛系数。若观测到系数 $\sigma_t$ 逐年下降,即可认为数字产业创新创业发展水平趋于收敛。图4展示了2010—2020年中国市级数字创新指数的 $\sigma$ 收敛系数。如图4所示, $\sigma$ 收敛系数由2010年的1.20快速下降至2018年的0.49,出现较快的区域收敛。但在2018—2020年间, $\sigma$ 收敛系数的维持在0.48~0.49,没有出现持续收敛的情况。

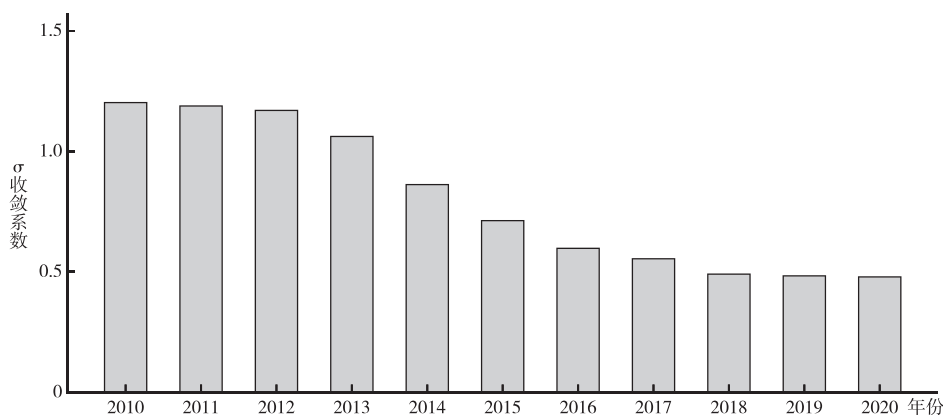


图4  $\sigma$ 收敛系数变动情况:2010—2020年

注:作者根据数字创新指数绘制。

本文同时计算了2010—2020年市级数字创新指数90%分位得分与10%分位得分的比值,反映数字产业创新创业发展的高分组与低分组随时间的相对变化趋势,并与戴若尘等(2021)基于总产业计算的区域创新指数进行比较。如图5所示,高低分组指数比值近年呈现出下降的趋势,由2010年的17.42倍下降到2020年的2.50倍,表明地区间创新创业发展水平逐渐趋同,低分城市追赶势头明显。但区域创新指数显示,自2012年起,高低分组比值已稳定在2以下,2020年这一比值已下降至1.43,低于数字创新指数。

综上,在2010—2020年我国数字产业内部低分城市虽然在绝对水平上和相对水平上都实现了突破式的发展,区域之间数字产业创新创业水平处于收敛状态,发展不平衡程度随时间下降。但近几年区域收敛速度达到瓶颈,呈现出稳定的区域发展差异。同时,与传统产业相比,数字创新指数得分较低的城市发展速度更慢、追赶高分城市难度更高,发展不平衡问题更加突出。

### (三)数字产业创新创业的地理集聚度分析

与传统经济模式相比,数字经济因其规模效应巨大、运输成本低、市场可达度高,导致地理集聚度会更高(江小涓、罗立彬,2019)。本文首先计算了数字产业在城市层面各指标的赫芬达尔指数,得到中国数字产业的地理集聚度。图6分别展示了数字产业与传统产业在各个指标上的地理集聚度,在创新创业的六个维度,数字产业的地理集聚度都更高,表明中国数字产业的集聚效应显著。与传统产业相比,高度的空间集聚意味着数字产业的创新创业活动更为集中,这可能是上文发现数字产业发展不平衡问题突出的一个重要原因。

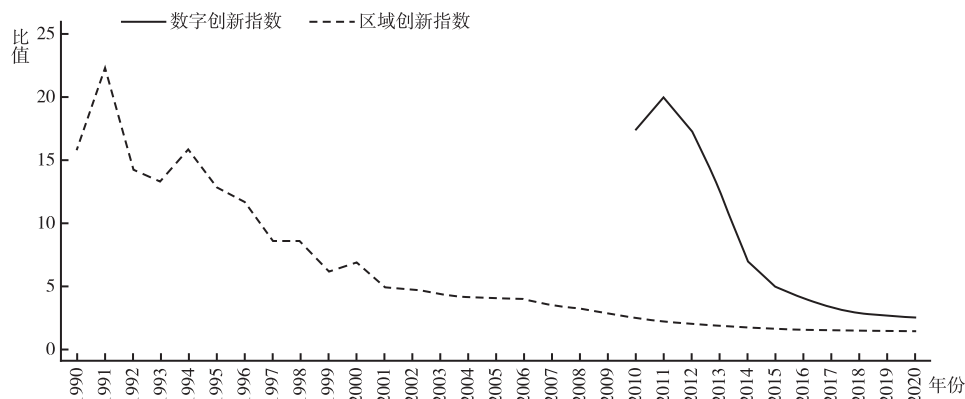


图5 1990—2020年市级数字创新指数90%分位与10%分位得分比值变化

注:作者根据数字创新指数和区域创新指数(戴若尘等,2021)绘制。

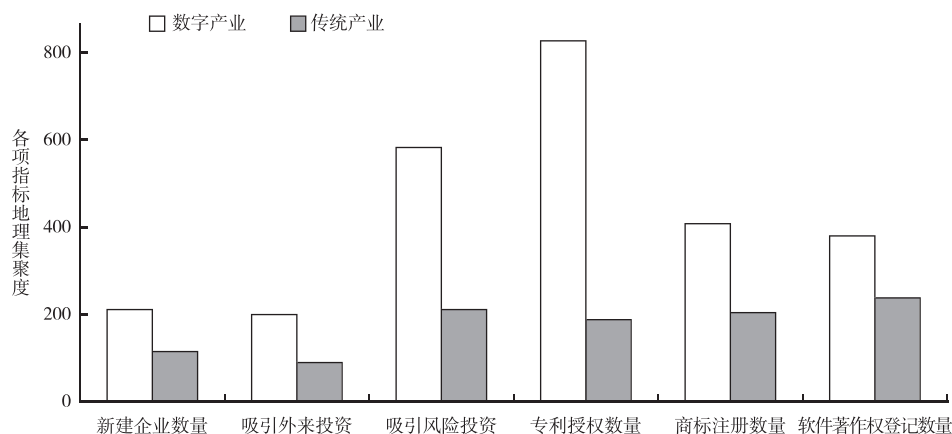


图6 2020年数字产业与传统产业地理集聚度

注:作者根据数字创新指数各项子指标计算,各指标地理集聚度已乘以10000。

本文进一步利用 Moran 指数检验了数字产业创新创业的空间相关性,分析邻近城市是否出现集聚与协同发展。作为一种新兴的经济形态,数字经济虽然能够克服传统贸易壁垒,在数字化平台上与远距离市场实现直接互动,但是其以数据为核心要素,高度依赖高技术人才,相邻城市存在着人才流动与知识外溢,邻近城市的发展对本城市数字经济发展仍有显著的影响(刘文明等,2020)。参考郭峰等(2020),本文使用空间距离权重矩阵计算了2010—2020年数字创新指数的 Moran 指数,并绘制了2010年、2014年、2018年和2020年四年数字创新指数局部 Moran 指数散点图,结果如图7所示。总体而言,我国数字产业创新创业水平自2010年起便存在显著的空间正相关性,由图7可以看出大多数城市位于一、三象限,即数字产业创新创业绩效较高的地区周边指数得分也较高、发展程度较低的城市与得分较低的城市相邻,这表明数字经济与传统经济相比虽然能够降低对地理条件的依赖,但仍具有明显的空间相关性。四年的 Moran 指数表明我国数字产业创新创业水平的空间自相关性不断提高,由2010年的0.25增长至2020年的0.34。人才、知识、技术和信息等因素的集聚和扩散推动相邻城市的协同发展,当前我国数字产业创新创业呈现出多核发展的态势,长三角、珠三角城市群、成渝城市群、京津冀城市群和山东半岛城市群等多个城市群内部的城市数字创新指数得分较高,空间联动效应突出。

#### (四)区域发展不平等的分解与南北差异

本文分析了中国数字产业创新创业水平在城市层面的收敛性、集聚度与空间关联度,发现数字产业虽然因较低的贸易成本导致各个维度指标都存在更高的空间集聚度,但总体发展在区域层面呈

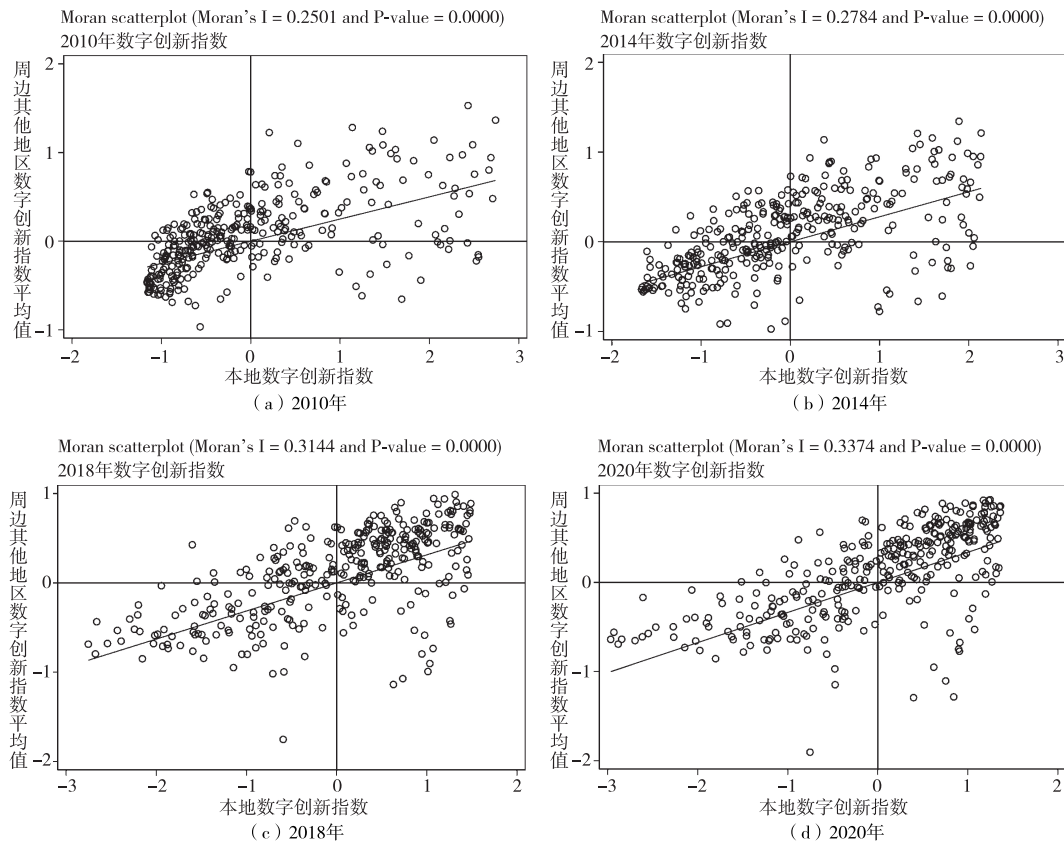


图7 数字创新指数局部 Moran 指数散点图

注：2010—2020年的 Moran 指数均为正且 P 值均为 0.0000。篇幅所限，本文在此仅展示 4 年局部 Moran 指数散点图。

现出逐渐收敛性，区域不平等得到缓解。这与相邻城市之间逐渐加强的空间关联度有关，先发展的城市可以通过知识和人才外溢等渠道带动周边城市的发展。以空间关联为主的外溢方式虽然会推动区域(城市群)内部实现均等化发展，但是区域之间的发展不平衡问题却会逐渐突出。本节从南北区域的差异入手，将能够衡量区域不平等程度的基尼系数进行分解，进一步刻画区域内部与区域间的不平等程度对于中国总体发展不平等程度的贡献度。

本文先使用市级数字创新指数对我国南北城市数字产业创新创业水平差异进行深入剖析，并进一步明晰区域间差异的动态演变过程。当前，我国南北城市数字产业创新创业水平存在差距，2020年市级得分前 20 名中仅有 6 个北部城市，前 50 名中北部城市占比为 32%。如图 8 所示，若以上文的标准划分南北，我国南北城市数字产业创新创业发展水平一直存在明显的得分差异，南部地区相较而言得分较高，2010 年区域间得分均值差异为 6.25 分，近年来差距已超过 10 分。

本文利用 Dagum 基尼系数刻画过去十年城市数字产业创新创业发展的不平等程度，并将其分解为南北区域内部差异(群内差异)、南北区域间差异(群间差异)以及交互差异(超变密度)，结果如表 6 所示。具体而言，总体基尼系数由 2010 年的 0.46 下降至 0.18，表明近年来南北地区数字产业创新创业发展不平衡程度呈现逐年下降的趋势，与之前的地区收敛性结果相一致。分解结果表明，与总体系数呈下降态势相同，2010—2020 年南北地区群内差异和超变密度呈现明显的下降趋势，而群间差异近年来则在 0.04~0.06 之间波动。从贡献率来看，南北地区群内差异近年占比虽有所下降，但仍超过 48%，这表明地区内的差异对总体差异的影响作用虽然有所减弱，但仍是数字产业创新创业地区发展不平衡的主要原因；超变密度刻画了地区之间发展存在的交叉重叠问题，其近年来贡献率出现不断降低的趋势，表明样本交叉重叠问题对区域数字产业创新创业水平



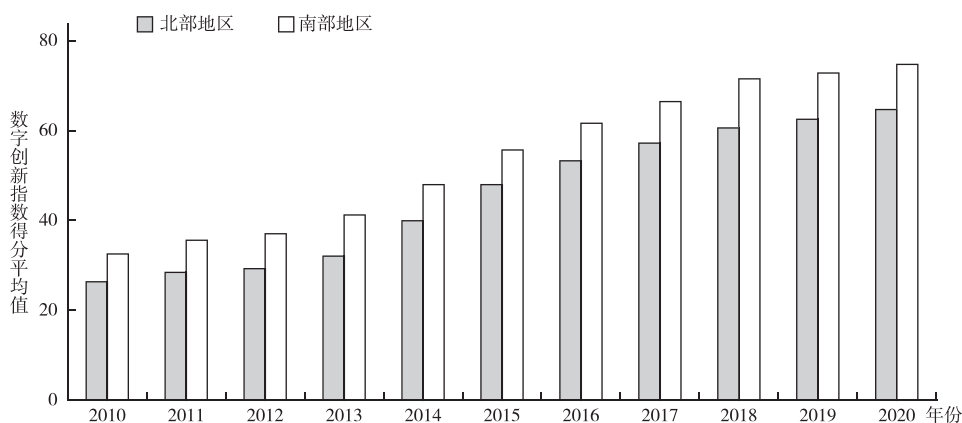


图8 南北地区市级数字创新指数平均值对比：2010—2020年

差距的影响逐年下降；群间差异贡献率占比最小，但总体而言呈现出上升趋势，由 11.53% 增长至 20.68%。

表6 2010—2020年中国南北地区数字创新指数基尼系数分解

年份	总体	群内贡献	群内贡献率	群间贡献	群间贡献率	超变密度	超变密度贡献率
2010	0.4613	0.2284	49.51%	0.0532	11.53%	0.1797	38.96%
2011	0.4484	0.2216	49.42%	0.0541	12.07%	0.1727	38.51%
2012	0.4432	0.2187	49.35%	0.0573	12.93%	0.1672	37.71%
2013	0.4158	0.2044	49.15%	0.0611	14.70%	0.1503	36.15%
2014	0.3349	0.1648	49.22%	0.0461	13.77%	0.1239	37.01%
2015	0.2838	0.1398	49.26%	0.0388	13.68%	0.1052	37.06%
2016	0.2457	0.1208	49.17%	0.0357	14.54%	0.0892	36.29%
2017	0.2186	0.1070	48.97%	0.0376	17.20%	0.0739	33.83%
2018	0.1920	0.0931	48.48%	0.0411	21.41%	0.0578	30.11%
2019	0.1832	0.0891	48.63%	0.0374	20.40%	0.0567	30.97%
2020	0.1760	0.0855	48.56%	0.0364	20.68%	0.0541	30.76%
均值	0.3093	0.1521	49.07%	0.0453	15.72%	0.1119	35.22%

综上，由于区域内部的城市能够协同发展，南北区域之间的差异对区域发展不平等的贡献度正逐渐上升。南北地区虽都出现了高分城市，但相比于南方地区，北方地区平均而言发展程度较低，数字创新指数 2020 年得分前 10 名中仅有北京市、青岛市与济南市位于北方地区，发展模式为单核发展，存在大量地区没有享受到外溢效应。南方地区由于多核发展，且容易形成城市群，城市的空间关联度和外溢性较大，因此总体发展更快速。

## 六、结论与政策启示

本文聚焦于企业这一核心市场主体，结合大数据思维，从企业家、资金与技术三大推动数字产业发展的核心要素出发，利用中国数字产业的企业大数据，创新性地构建了中国数字经济核心产业创新创业指数。该指数能够帮助分析描述中国数字产业创新创业水平的动态演化过程和空间分布特征，并为进一步开展针对中国数字产业创新创业问题的经济学实证研究提供数据支持。

本文构建的指数将数字产业企业的创业行为和创新行为有机结合，能够从创新创业角度对地区数字产业发展水平进行细致刻画。基于构建的数字创新指数，本文总结描绘了中国数字产业创新创业在动态演化与空间分布方面的典型事实。本文主要结论如下：第一，在动态演化方面，2010—2020 年间中国各地区数字创新指数不断提高，创新创业绩效显著提升，但 2016 年之后高质量创新创业占比增速不断放缓；第二，在空间格局上，数字产业创新创业发展已逐步形成“南热北冷、中部崛起”的

空间格局;第三,中国数字产业创新创业发展的区域不平等程度逐渐下降,其原因为相邻城市存在空间关联,协同发展;第四,数字产业与传统产业相比有着更高的地理集聚度,区域之间不平等的贡献度逐渐加强。

2021年发布的《“十四五”数字经济发展规划》提出了数字经济核心产业增加值占GDP比重从2020年7.8%提升到2025年的10%的规划目标。企业是市场经济的主体,是创新体系的核心,数字产业化水平的提高离不开活跃的企业家精神,因此,产业政策的制定需要关注如何更好地激发企业家精神。本文基于数字创新指数,从多个维度分析了中国数字产业创新创业在过去十年内的动态演变与空间分化特点,并提出三点政策建议:第一,数字产业化发展需要进一步激发草根创业者的创业创新热情,特别是具备高技能、具有创新精神的企业家,鼓励具有高科技含量的优质创新行为,从而进一步实现数字产业的分工细化。第二,加强互联网、5G和软件开发平台等新兴基础设施的建设,基础设施的建设与完善能够进一步降低创业门槛,提高区域内数字产业的创业创新水平。第三,与传统行业相比,数字经济行业存在贸易成本低、产业集聚度高、依赖高知识人才等特点,在空间分化中呈现出城市群之间协同发展的特点。因此,区域发展模式需要从单核发展模式转移到多核发展模式,引导并鼓励各个城市之间进一步分工协作,打造基于数字经济的完备产业链。具体而言,一方面推动数字产业化发展较快的区域(例如浙江省与广东省)进一步提高自身的规模效应,形成专业的数字经济产业集群;另一方面鼓励落后区域优先发展先发城市,并充分发挥先发城市的外溢作用,从新的细分行业加入中国数字经济产业链中,促进区域均衡发展。充分发挥数字经济的规模优势特点,从国内一体化与顶层设计的角度在全国范围内构建完备的数字产业链,应该是当前中国推动构建双循环新发展格局和提高国际竞争力的工作重点与改革方向。

本文目前只针对数字经济的核心产业,出于数据的缺乏,暂时未衡量产业数字化部分的创新创业活动。希望未来可以通过文本识别等大数据手段、结合专利商标等信息构建反映中国在产业数字化程度的区域性指数,例如逐步兴起的机器人流程自动化(RPA)行业与工业机器人行业(王永钦、董雯,2020)。此外,目前的指数并不具备与国际上其他指数的可比性,通过国际比较判断中国数字产业创新创业水平在世界范围内的所处阶段也是未来研究的重点。

#### 参考文献:

- 柏培文 张云,2021:《数字经济、人口红利下降与中低技能劳动者权益》,《经济研究》第5期。
- 柏培文 喻理,2021:《数字经济发展与企业价格加成:理论机制与经验事实》,《中国工业经济》第11期。
- 财新智库 数联铭品,2021:《2021年5月中国数字经济指数报告》,https://yun.ccxe.com.cn/indices/dei。
- 陈思 何文龙 张然,2017:《风险投资与企业创新:影响和潜在机制》,《管理世界》第1期。
- 戴若尘 祝仲坤 张晓波,2021:《中国区域创新创业指数构建与空间格局:1990—2020》,北京大学企业大数据研究中心工作论文。
- 樊纲 王小鲁 朱恒鹏,2011:《中国市场化指数——各地区市场化相对进程2011年报告》,经济科学出版社。
- 郭峰等,2020:《测度中国数字普惠金融发展:指数编制与空间特征》,《经济学(季刊)》第4期。
- 国家统计局,2021:《数字经济及其核心产业统计分类(2021)》,http://www.stats.gov.cn/tjgz/tzgb/202106/t20210603\_1818129.html。
- 郭凯明 潘珊 颜色,2020:《新型基础设施投资与产业结构转型升级》,《中国工业经济》第3期。
- 韩先锋 宋文飞 李勃昕,2019:《互联网能成为中国区域创新效率提升的新动能吗》,《中国工业经济》第7期。
- 何小钢 梁权熙 王善骞,2019:《信息技术、劳动力结构与企业生产率——破解“信息技术生产率悖论”之谜》,《管理世界》第9期。
- 胡山 余泳泽,2022:《数字经济与企业创新:突破性创新还是渐进性创新?》,《财经问题研究》第1期。
- 黄群慧 余泳泽 张松林,2019:《互联网发展与制造业生产率提升:内在机制与中国经验》,《中国工业经济》第8期。
- 姜松 孙玉鑫,2020:《数字经济对实体经济影响效应的实证研究》,《科研管理》第5期。
- 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。
- 江小涓 罗立彬,2019:《网络时代的服务全球化——新引擎、加速度和大国竞争力》,《中国社会科学》第2期。
- 蒋震 苏京春 杨金亮,2021:《数字经济转型与税制结构变动》,《经济动态》第5期。

- 李兵 岳云嵩 陈婷,2016:《出口与企业自主技术创新:来自企业专利数据的经验研究》,《世界经济》第12期。
- 刘传明 尹秀 王林杉,2020:《中国数字经济发展的区域差异及分布动态演进》,《中国科技论坛》第3期。
- 龙小宁 王俊,2015:《中国专利激增的动因及其质量效应》,《世界经济》第6期。
- 陆毅 孙天阳,2021:《以科学规范的经济学研究方法讲好中国故事》,《经济学(季刊)》第5期。
- 吕炜,2002:《论风险投资机制的技术创新原理》,《经济研究》第2期。
- 毛丰付 张帆,2021:《中国地区数字经济的演变:1994—2018》,《数量经济技术经济研究》第7期。
- 毛昊 尹志锋 张锦,2018:《中国创新能够摆脱“实用新型专利制度使用陷阱”吗》,《中国工业经济》第3期。
- 莫怡青 李力行,2022:《零工经济对创业的影响——以外卖平台的兴起为例》,《管理世界》第2期。
- 戚聿东 刘翠花 丁述磊,2020:《数字经济发展、就业结构优化与就业质量提升》,《经济学动态》第11期。
- 戚聿东 孙洁 李峰,2019:《数字期权理论研究进展》,《经济学动态》第5期。
- 赛迪顾问,2020:《2020中国数字经济发展指数白皮书》,https://www.ccidgroup.com/info/1096/21309.htm。
- 邵朝对 苏丹妮 王晨,2021:《服务业开放、外资管制与企业创新:理论和中国经济》,《经济学(季刊)》第4期。
- 孙宝文 荆文君 何毅,2017:《互联网行业反垄断管制必要性的再判断》,《经济学动态》第7期。
- 孙杰,2020:《从数字经济到数字贸易:内涵、特征、规则与影响》,《国际经贸探索》第5期。
- 孙早 侯玉琳,2019:《工业智能化如何重塑劳动力就业结构》,《中国工业经济》第5期。
- 腾讯研究院,2020:《数字中国指数报告(2020)》,https://www.tisi.org/16456。
- 万晓瑜 罗焱卿,2022:《数字经济发展水平测度及其对全要素生产率的影响效应》,《改革》第1期。
- 王兰芳 胡悦,2017:《创业投资促进了创新绩效吗?——基于中国企业面板数据的实证检验》,《金融研究》第1期。
- 王林辉 胡昆明 董直庆,2020:《人工智能技术会诱致劳动收入不平等吗——模型推演与分类评估》,《中国工业经济》第4期。
- 王永钦 董雯,2020:《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》,《经济研究》第10期。
- 吴超鹏 吴世农 程静雅 王璐,2012:《风险投资对上市公司投融资行为影响的实证研究》,《经济研究》第1期。
- 新华三集团数字经济研究院,2020:《中国城市数字经济指数白皮书(2020)》,http://deindex.h3c.com。
- 熊巧琴 汤珂,2021:《数据要素的界权、交易和定价研究进展》,《经济学动态》第2期。
- 续继 唐琦,2019:《数字经济与国民经济核算文献评述》,《经济学动态》第10期。
- 许宪春 张美慧,2020:《中国数字经济规模测算研究——基于国际比较的视角》,《中国工业经济》第5期。
- 许宪春 张钟文 胡亚茹,2022:《数据资产统计与核算问题研究》,《管理世界》第2期。
- 杨飞 范从来,2020:《产业智能化是否有利于中国益贫式发展?》,《经济研究》第5期。
- 叶文平 李新春 陈强远,2018:《流动人口对城市创业活跃度的影响:机制与证据》,《经济研究》第6期。
- 张杰 高德步 夏胤磊,2016:《专利能否促进中国经济增长——基于中国专利资助政策视角的一个解释》,《中国工业经济》第1期。
- 张杰 郑文平,2018:《创新追赶战略抑制了中国专利质量么?》,《经济研究》第5期。
- 张维迎 盛斌,2014:《企业家:经济增长的国王》,上海人民出版社。
- 张学勇 张叶青,2016:《风险投资、创新能力与公司IPO的市场表现》,《经济研究》第10期。
- 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。
- 赵云辉 张哲 冯泰文 陶克涛,2019:《大数据发展、制度环境与政府治理效率》,《管理世界》第11期。
- 中国信息通信研究院,2017:《中国数字经济发展白皮书(2017年)》,http://www.caict.ac.cn/kxyj/qwfb/bps/201804/t20180426\_158452.htm。
- 中国信息通信研究院,2021:《中国数字经济发展白皮书》,http://www.caict.ac.cn/kxyj/qwfb/bps/202104/t20210423\_374626.htm。
- 周广肃 樊纲,2018:《互联网使用与家庭创业选择——来自CFPS数据的验证》,《经济评论》第5期。
- Baumol, W. J. (1996), “Entrepreneurship: Productive, unproductive, and destructive”, *Journal of Business Venturing* 11(1):3–22.
- Chemmanur, T. J. et al(2011), “How does venture capital financing improve efficiency in private firms? A look beneath the surface”, *Review of Financial Studies* 24(12):4037–4090.
- Da Rin, M. & M. F. Penas(2007), “The effect of venture capital on innovation strategies”, NBER Working Paper, No. 13636.
- Glaeser, E. L. et al(2015), “Entrepreneurship and urban growth: An empirical assessment with historical mines”,

- Review of Economics and Statistics* 97(2):498—520.
- Goldfarb, A. & C. Tucker(2019), “Digital economics”, *Journal of Economic Literature* 57(1):3—43.
- Hu, X. et al(2021), “Crisis-induced innovation: Quality upgrading in Chinese industrial clusters”, *Journal of Law, Economics, and Organization* 37(3):571—606.
- Kortum, S. & J. Lerner(2000), “Assessing the contribution of venture capital to innovation”, *RAND Journal of Economics* 31(4):674—692.
- Samila, S. & O. Sorenson(2011), “Venture capital, entrepreneurship, and economic growth”, *Review of Economics and Statistics* 93(1):338—349.
- Sun, W. et al(2020), “Big data analytics for venture capital application: Towards innovation performance improvement”, *International Journal of Information Management* 50:557—565.

### Innovation and Entrepreneurship in Core Industries of Digital Economy in China: Stylized Facts and Index Compilation

DAI Ruochen WANG Aizhao CHEN Binkai

(Central University of Finance and Economics, Beijing, China)

**Abstract:** Digital economy’s development affects the level of national development, and promoting innovation and entrepreneurship in core industries of digital economy is crucial for the high-quality development of China’s economy. Utilizing a novel firm-level big data collected by the Center for Enterprise Research of Peking University, and focusing on three key factors in core industries of digital economy as entrepreneurs, capital and technology, we innovatively construct the Index of Regional Innovation & Entrepreneurship in Digital Economy in China (IRIEDEC) from six dimensions, i. e. business entry, outside investment, venture capital, patent, trademark, and copyright. The index for the first time captures the dynamic evolution process and spatial distribution characteristics of innovation and entrepreneurship of core industries in China’s digital economy from 2010 to 2020. IRIEDEC is constructed through big data thinking, which treats enterprises as micro subjects in digital economy and focuses on innovative output instead of input, with characteristics of objectivity, timeliness and multi-dimensionality. Our study finds that; firstly, the level of innovation and entrepreneurship of core industries in China’s digital economy has improved rapidly in the past decade; secondly, IRIEDEC at provincial level presents a spatial pattern of “hot in the south, cold in the north, and rising in the middle”; thirdly, IRIEDEC at municipal level demonstrates a trend of regional convergence, and the inequality level of regional development decreases; fourthly, the innovation and entrepreneurship of core industries in China’s digital economy is geographically agglomerated, and the development of adjacent cities is spatially correlated; fifthly, the contribution of the differences between the two regions is constantly increasing.

**Keywords:** Digital Economy; Innovation and Entrepreneurship; Firm-level Big Data; Intertemporal Change; Spatial Pattern

(责任编辑:何伟)

(校对:刘洪愧)