

数字知识流动如何促进区域协调发展

——兼论经济增长和平衡发展双重目标

马述忠，张道涵，胡增玺

[摘要] 区域协调发展需要在区域协同的基础上实现人均意义上的区域经济持续增长和平衡发展。数字经济时代，数字知识已成为技术进步和经济增长的源泉，是促进区域协调发展的重要力量。本文基于大语言模型方法和专利引用信息测度了数字知识的跨区域流动，进一步结合夜间灯光和人口栅格数据，深入考察了数字知识流动对区域协调发展的影响，赋予传统增长理论数字时代的新内涵。研究发现，由发达地区流入欠发达地区的数字知识能够促进区域协调发展。机制分析表明，在欠发达地区产业与数字知识适配的前提下，改善供给侧对需求的适配性以缓解要素禀赋差距，是数字知识流动促进区域协调发展的关键路径。异质性分析表明，东部发达城市向中西部欠发达城市的数字知识流动、大型发达城市向小型欠发达城市的数字知识流动，以及发展水平相近城市间的数字知识流动促进区域协调发展的效果更好。拓展分析发现，城市内部中心区县流向外围区县的数字知识能够显著缓解“中心—外围”格局，并且数字知识流动促进区域协调发展的作用是长期存续的。因此，政府应通过优化区域数字产业结构、建立灵活的需求反馈机制等方式，有针对性、有优先级地促进数字知识的跨区域流动，发挥长期积极效应。

[关键词] 数字知识流动；区域协调发展；知识适配性；需求适配性；要素禀赋差距

[中图分类号] F424 **[文献标识码]** A **[文章编号]** 1006-480X(2025)02-0080-19

一、引言

党的十八大以来，中国在区域协同方面取得积极进展，但区域发展差距依然较大。2023年12月，中华人民共和国国家发展和改革委员会、国家数据局印发的《数字经济促进共同富裕实施方案》强调，要推动区域数字协同发展。2024年《政府工作报告》强调，要深入实施区域协调发展战略，提高区域协调发展水平。可见，如何在数字经济时代推动区域协调发展，是中国实现共同富裕和经济社会高质量发展的重要议题。

[收稿日期] 2024-09-18

[基金项目] 国家社会科学基金重大项目“数字贸易壁垒对中国对外贸易高质量发展的影响及对策研究”(批准号23&ZD084)。

[作者简介] 马述忠，浙江大学中国数字贸易研究院教授，博士生导师，管理学博士；张道涵，浙江大学中国数字贸易研究院博士研究生；胡增玺，中央财经大学国际经济与贸易学院助理教授，经济学博士。通讯作者：胡增玺，电子邮箱：zengxi_hu@163.com。感谢匿名评审专家和编辑部的宝贵意见，文责自负。

实现区域协调发展需要兼顾经济增长和平衡发展双重目标。具体来说,在新形势下促进区域协调发展,需要尊重生产要素向优势区域集中的客观经济规律。区域协调发展的目标并不是追求区域之间要素总量或经济规模的均等,而是在区域协同的基础上实现人均意义上的区域经济持续增长和区域间平衡发展(陆铭和李鹏飞,2022)。然而,现实困境是,中国改革开放以来,发达地区对欠发达地区的“虹吸效应”使生产要素的空间布局发生巨大变化(洪银兴,2022),要素向发达地区集聚虽然强化了发达地区的规模经济优势,但也加剧了区域分化和“马太效应”。因此,考虑到供给侧结构性改革有利于疏解要素环境变化的压力,也是发展数字经济的工作主线,优化欠发达地区的数字经济供给体系以提升其需求适配性,是缓解要素禀赋分布不均、促进区域协调发展的突破口之一。

在中国,知识流动是区域间经济互动性的主要来源(张勋和乔坤元,2016),并且在局部地区展现出促进经济空间收敛的作用(朱丰毅和桂文林,2022)。鉴于数字知识流动是数字经济时代区域协同的重要形式,在欠发达地区产业与数字知识适配的前提下,由发达地区流入欠发达地区的数字知识有利于改善供给侧对需求的适配性,进而缓解要素禀赋差距,最终助力中国实现区域协调发展的双重目标,即兼顾区域经济增长和平衡发展。^①具体来说,数字经济环境下消费者的个性化需求大量涌现,但是传统供给体系依赖标准化、批量化的生产实现盈利,难以快速调整以获取“长尾效应”带来的溢价(Brynjolfsson et al., 2011; 郭继文和马述忠, 2022)。数字知识的流入能够通过引入先进的数字技术和管理经验,赋予供给侧更强的适应性和灵活性,助力供给侧精准识别偏好、快速响应需求,进而增加有效供给。因此,对于欠发达地区而言,数字知识流入有利于在不改变初始要素禀赋的前提下,提高全要素生产率和资源配置效率,实现生产可能性曲线的外推(荆文君和孙宝文, 2019),进而缓解当地与发达地区由要素禀赋差距引致的发展不平衡问题,助力区域间的动态协调发展。

尽管数字知识流动对中国推动区域协调发展具有重要意义,但是目前学术界还未曾探讨二者之间的关系。由发达地区流入欠发达地区的数字知识是否能够促进区域协调发展?如果可以,那么这种积极作用是否以欠发达地区产业与数字知识适配为前提?若得到肯定的答案,那么,改善供给侧对需求的适配性进而缓解要素禀赋差距,是否为数字知识流动促进区域协调发展的关键路径?回答以上问题,不仅有利于填补数字知识流动与区域协调发展关系的理论空白,也能为中国进一步实现数字经济时代公平与效率的统一提供政策启示。鉴于此,考虑到专利引用记录是代理知识流动的理想指标,本文采用大语言模型识别专利所涵盖的数字技术,再利用数字技术专利的引用情况刻画数字知识流动,进一步结合长时间序列夜间灯光和人口栅格数据,系统探讨了数字知识流动促进区域协调发展的作用及其内在机制。

与既有研究相比,本文的边际贡献主要在于:^①随着新增长理论通过引入知识要素修正了新古典增长理论的外生技术进步假说(Romer, 1986; Lucas, 1988),大量研究探讨了传统经济模式下知识对区域经济增长的驱动作用,然而关注区域协调发展这一重要现实议题的研究相对较少,更未曾有研究考察数字经济时代的全新知识形态——数字知识如何作用于区域经济。本文首次系统探讨了数字知识流动对区域协调发展的影响,拓展了传统增长理论的分析维度和解释边界,赋予了新古典增长理论、新增长理论等经典理论在数字经济时代的新内涵。^②既有研究大多从技术创新和要素

^① 为避免歧义,本文中的“数字知识”是指内嵌于数字技术的数字对象(数字知识的详细定义及其和数字技术的关系参见后文第三部分),而不是指所有以数字形式流动的知识。

集聚视角探讨知识流动对区域经济的影响(Owen-Smith and Powell, 2004; 赵勇和白永秀, 2009; 白俊红等, 2017), 且没有对知识的流向给予过多关注, 因此, 更多聚焦于知识流动和经济增长的因果关系, 而非解构知识流动影响区域差距的内在逻辑。考虑到数字知识流动本质上体现了城市间的经济互动, 本文以“在发展中促进相对平衡”为研究思想, 基于供给侧对知识和需求的双重适配性和要素禀赋的区域差距, 深入探讨了两两城市间数字知识流动影响其相对发展差距的理论机制, 揭示了数字知识助力供给侧结构性改革的内在逻辑, 为中国推动区域协调发展和共同富裕提供了可参考的实践路径。③统计专利引用情况是量化知识流动的经典做法(Jaffe et al., 1993), 但可能由于准确识别专利中的数字技术存在一定难度, 在数字经济蓬勃发展的背景下, 聚焦于数字知识的量化研究仍有待探讨。本文利用大语言模型领域的BERT(Bidirectional Encoder Representations from Transformers, BERT)模型, 通过识别专利摘要文本中涵盖的数字技术及其引用情况, 较为精准地测算了区域数字知识流, 丰富了大语言模型在经济学中的应用场景。此外, 本文基于夜间灯光栅格、人口栅格等高颗粒度地理信息系统(Geographic Information System, GIS)数据, 克服了长时间序列条件下城市层面统计数据的稀缺和口径不统一问题, 还在手动配准近代舆图的基础上构造了基于有线电报局网点的历史GIS工具变量, 相关指标的量化测度方式能够为后续研究提供有益参考。

余文结构安排如下: 第二部分为文献综述和理论分析; 第三部分阐述实证策略, 包括介绍研究采用的大语言模型识别方法、模型设定和变量数据; 第四部分为实证分析, 包括基准回归、稳健性检验、内生性问题分析和机制检验; 第五部分为异质性分析; 第六部分为结论与政策启示。

二、文献综述和理论分析

1. 文献综述

学术界针对知识流动与区域经济发展关系的研究源于20世纪中叶新古典增长理论对外生技术进步的探讨。随后, Arrow(1962)关于“干中学”(Learning by Doing)的开创性研究系统论述了知识的经济含义, 并刻画了知识积累的外部性。进一步地, 新增长理论在索洛增长模型的基础上区分研发部门与生产部门, 从而引入知识要素以内生化技术进步, 表明知识积累是经济持续增长的根本动力(Romer, 1986; Lucas, 1988)。在上述观点基础上, 大量研究表明, 知识的跨区域流动可以显著促进流入地的经济增长(赵勇和白永秀, 2009)。一方面, 知识流入能够推动流入地的技术创新, 进而降低长期生产成本并提高生产率(Owen-Smith and Powell, 2004; 曹湛等, 2022)。另一方面, 知识的流入往往伴随着生产要素和产业在当地的集聚(Singh, 2005; 朱丰毅和桂文林, 2022)。这带来的规模经济效应会进一步提高创新部门效率(蒋伏心和高丽娜, 2012), 放大知识促进技术创新和经济增长的积极作用(白俊红等, 2017)。

遗憾的是, 上述研究大多聚焦于知识流入对区域经济增长的影响, 较少关注区域间平衡发展这一深层次的协调发展目标, 也没有对知识的流向给予太多关注, 更没有在数字经济蓬勃发展的现实背景下将数字知识作为核心探讨对象, 拓展现有基于传统经济形态的增长理论, 检验数字知识流动带来的经济增长是否兼顾效率和公平。

尽管数字知识流动和区域协调发展的关系仍属学术空白, 考虑到数字知识催生出各类数字经济新业态, 探讨数字经济与区域差距的研究值得关注。此类研究大多认为, 数字经济能通过促进要素流动、推动合作创新等机制降低区域间尤其是城市间的发展差距(邓慧慧等, 2022; 姚常成和沈凯琦, 2023)。不过, 也有研究指出, 特定类型的数字经济可能扩大市际差距(叶堂林和王雪莹, 2023)。

随着可得数据微观化,市域尺度的研究日益增多。例如,彭刚等(2023)发现,城市发展数字经济会扩大内部区县间的发展差距。然而,郭峰等(2023)指出,数字经济能缩小城市边界区和中心区之间的差距;马述忠等(2024)以跨境电子商务为研究对象,验证了数字经济弥合城市内部县际差距的作用。总之,即使在同一研究尺度上,学术界也未能就数字经济和区域协调发展的关系达成共识。需要说明的是,数字经济和数字知识流动对区域协调发展的影响存在底层逻辑的不同,并不能一概而论。数字经济影响区域协调发展的关键在于单个城市自身的经济增长和当地不同部门相对生产效率的变化(洪俊杰等,2024),而数字知识流动的影响机理则能更多追溯到城市间的经济互动(张勋和乔坤元,2016)。鉴于数字知识和数字经济密切相关,上述研究对数字经济普惠逻辑的刻画,能够为本文对数字知识的探讨提供有益启示。

2. 理论分析

区域协调发展是人均意义上区域内经济持续增长和区域间经济平衡发展的有机结合(陆铭和李鹏飞,2022),其关键在于欠发达地区的持续快速增长(叶堂林和王雪莹,2023)。在数字经济时代,由发达地区流入欠发达地区的数字知识是区域协调发展的重要源动力。这种积极作用可以追溯至数字知识流的两类属性:①作为跨区域知识流的属性。数字知识流作为跨区域知识流的重要形式,不仅具有结构化、正式化和规范化的特征,还能够传递市场机会和潜在需求的新信息。因此,外部知识的流入能够打破区域技术创新边界,促进知识与技术在不同区域间的整合与重组,从而帮助欠发达地区弥补本地知识的匮乏,避免内部知识边际收益递减带来的“低端锁定困局”(Owen-Smith and Powell, 2004)。②数字知识自身的内在属性,包括自生长性、无边界性等。一方面,数字知识可重复编程的特征允许其被动态迭代和无限重塑(Yoo et al., 2010),这种自生长性极大加快了知识在区域间的扩散速率。另一方面,数字经济时代诞生了大量具有开放共享特点的新业态,这些新业态的普及和深化有利于缓解知识的部分排他性,进而打破组织边界和产业边界。这种无边界性提升了数字知识使用主体的多元性,增大了“创造性破坏”的发生概率(Guellec and Paunov, 2017),从而进一步放大了数字知识流入欠发达地区的边际效益。基于上述分析,本文提出:

假说1:由发达地区流入欠发达地区的数字知识能够促进区域协调发展。

欠发达地区产业与数字知识适配,是数字知识流动促进区域协调发展的前提。知识对经济发展的驱动作用必须经历被流入地吸收和转化的过程(叶静怡等,2019)。实际上,供给侧的知识适配性在很大程度上影响着长期经济增长率(蒋伏心和高丽娜,2012),而不同地区由于经济基础、产业结构、人才储备等因素的差异,对知识的吸收和转化能力不尽相同。换言之,如果欠发达地区对知识的适配性不足,就无法有效承接外部知识流以弥补发展短板、追赶发达地区。在数字经济时代,这种知识适配性具体表现为数字产业的发展水平。一方面,数字产业能够将流入的数字知识与本地资源禀赋和产业需求相结合,使知识转化为更具针对性的生产力。例如,针对资源型地区的绿色发展需求,数字产业能够结合数字知识引入先进的环境监测技术并优化清洁生产方案;在以消费品制造为主的地区,数字产业则可利用数字知识开发智能供应链管理技术以加快市场响应速度。另一方面,数字产业还能够通过构建开放式的生态系统,促进数字知识的扩散与共享,尤其是上下游产业部门间的知识溢出(赵勇和白永秀,2009)。

改善欠发达地区供给侧对需求的适配性以缓解要素禀赋差距,是数字知识流动促进区域协调发展的关键路径。在传统经济模式下,增加要素投入进而使生产点不断接近生产可能性曲线,是最普遍的增长路径(荆文君和孙宝文,2019)。然而,发达地区往往对要素禀赋具有“虹吸效应”,这会增强“马太效应”,损害欠发达地区的长期可持续增长能力,阻碍区域差距的弥合(洪银兴,2022;叶

堂林和王雪莹,2023)。实际上,要素禀赋的空间分布不均正是中国区域发展不平衡、不协调的主要原因之一(Zhang and Zhang,2003)。在上述困境难以迅速改善的情况下,本文认为,数字知识流入欠发达地区能够通过改善供给侧对需求的适配性,提高当地的全要素生产率和资源配置效率,进而提升生产部门的运作效率并强化研发部门的创新能力,从多维度缓解要素禀赋差距问题。值得注意的是,除了资本、劳动力等传统经济模式下的主要生产要素,作为数字时代的新型生产资料,数据要素也可能成为数字知识流动促进区域协调发展的重要抓手。

具体来说,数字经济时代,以平台为代表的数字技术实现了消费需求的线上加总,即使是较为小众的需求,也能在汇总后形成可观的规模(Brynjolfsson et al.,2011)。相比传统经济模式下企业通过批量化、标准化生产形成规模经济的盈利逻辑,识别并满足消费者的个性化需求以获取差异化溢价逐渐成为供给侧的新增长点(郭继文和马述忠,2022)。基于数字知识,欠发达地区企业可以通过引进和开发在线平台、大数据、云计算等数字技术提高全要素生产率。与此同时,从发达地区流入的数字知识往往还包含先进管理经验与前沿市场信息,能够帮助欠发达地区企业更高效精准地在不同部门间分配生产资料,更好地保证要素投入的充裕性与持续性(荆文君和孙宝文,2019)。换言之,即使在资本、劳动力等传统要素禀赋不足的情况下,欠发达地区企业仍然能通过把握市场消费趋势,充分利用既有的数据要素开发个性化、定制化的产品服务,获取“长尾效应”带来的差异化溢价(Brynjolfsson et al.,2011)。因此,数字知识流入欠发达地区能通过提升需求适配性提高全要素生产率和资源配置效率,使有限的要素禀赋在各部门间实现更为充分的协同优化,从而产生更高的边际收益,进而缓解要素禀赋的空间分布不均问题,助力区域协调发展。基于上述分析,本文提出:

假说2:在欠发达地区产业与数字知识适配的前提下,数字知识流入当地有利于改善供给侧对需求的适配性,进而缓解要素禀赋差距,最终促进区域协调发展。

三、实证策略

1. 识别内嵌数字知识的数字技术:大语言模型方法

(1)概念界定:数字知识和数字技术。鉴于数字知识是本文的核心研究对象,且其内嵌于数字技术(Hund et al.,2021),需要先对二者的概念和关系进行界定。单纯从技术角度讲,在数字经济背景下,一种特别有趣的非物质“对象”(Object)是遵循一组语法规则的比特串(Bitstrings),而数字知识对应的“数字对象”则被定义为组成部分包括一个或多个比特串的对象(Faulkner and Runde,2019;Hund et al.,2021)。具象来说,数字知识本身可以是纯非物质对象(如代码),但也可以是混合体(如嵌入代码的设备)。个人和社会群体能以不同的方式使用数字知识,赋予其技术身份,塑造其边界(Faulkner and Runde,2009)。在此基础上,数字技术的概念便诞生了,即已被社会认可的数字知识。换句话说,数字知识与数字技术是一体两面,数字知识在被赋予含义时成为数字技术,且该过程由社会行为者(如用户)基于应用目的完成(Yoo et al.,2010;Hund et al.,2021)。从一般性角度考虑,本文将采纳这一社会技术视角的界定。

(2)分析对象:数字技术专利文本。考虑到专利引用情况是量化知识流动的经典指标(Jaffe et al.,1993),数字技术专利的引用情况是刻画数字知识流动的理想数据。现有的一些专利分类系统包含某些特定的数字技术,例如,国家知识产权局发布的《数字经济核心产业分类与国际专利分类参照关系表(2023)》提供了部分数字经济核心产业相关专利的国际专利分类(International Patent Classification,IPC)。然而,要想精确识别与数字技术有关的专利却较为困难。这是因为,数字技术

通常被认为是一种通用目的技术(General Purpose Technologies, GPT),这个标签是极为宽泛的。现实中,该术语本质上涵盖所有以比特串为基础的技术,这跨越许多技术部门和行业,涉及大量不同类型的专利。

与此同时,文本文档是社会科学家丰富而重要的数据来源(Gentzkow et al., 2019),而专利摘要恰好有涵盖专利名称、所属技术领域、需要解决的技术问题、主要技术特征、用途等方面的文字说明。鉴于此,本文以《数字经济及其核心产业统计分类(2021)》提供的108个数字经济核心产业类别所对应的具体数字技术为基准,以1985年《中华人民共和国专利法》实施以来至2019年在国家知识产权局申请并公开的所有专利为分析对象,进行是否含有数字技术或者属于哪种数字技术的文本识别。《数字经济核心产业分类与国际专利分类参照关系表(2023)》在《数字经济及其核心产业统计分类(2021)》的基础上匹配了每个数字经济核心产业分类对应的专利IPC分类代码,本文细致检查该分类后发现两个方面的问题:一是类别缺失问题。该对照仅提供了86个小类的对照,而《数字经济及其核心产业统计分类(2021)》提出的数字经济核心产业小类共108个(删除无实际意义含“其他行业”表述的小类),且上文也提到了数字技术作为GPT横跨许多技术部门和行业,涵盖许多不同的专利类别,因此,不能仅仅通过几个类别就完全识别专利中内含的某一数字技术。二是一个专利分类号往往对应多个类别,且有些类别中的专利分类号所包含的专利很多都与该类别几乎不相关。这是因为,IPC专利号往往基于技术主体和物理特征,而数字技术的本质则是功能和应用。这意味着该分类标准存在“二类错误”,即缺乏完备性和错误归纳。因此,本文将基于完备的108个细分小类进行数字技术识别。^①诚然,本文可以逐一根据专利摘要所描述的信息进行人工识别,然而,由于专利数据数以千万计,这一方法可行性较低,因此,需要采用自然语言处理领域前沿的大语言模型来识别。^②

2. 模型设定

中国城市间的显著差距仍是区域发展不协调的重要体现。例如,中国东部、中部、西部城市人均收入比虽然从2000年的2.4:1.2:1.0缩小至2019年的1.4:1.0:1.0,但是中西部与东部城市之间的差距仍然偏大,同时高收入行业也大多分布在发达城市(洪银兴, 2022)。尽管一些嵌入城市群落的中小城市有异军突起之势,但其与中心发达城市之间依然存在无法被忽视的差距(姚常成和沈凯琦, 2023)。因此,城市间协调发展是本文的主要关注点。鉴于数字知识流动具有明确的方向性,为了从经济增长和平衡发展双重目标出发,更细致地刻画两两区域之间相对发展水平的动态变化,本文遵循“在发展中促进相对平衡”的区域发展新理论,通过如下模型考察数字知识流动对区域协调发展的影响:

$$Bal_{it} = \alpha_0 + \alpha_1 \ln DKF_LtoH_{it-1} + \alpha_2 \ln DKF_HtoL_{it-1} + \beta X'_{it-1} + \lambda_i + \lambda_t + \varepsilon_{it} \quad (1)$$

① 实际上,根据参照关系表提供的IPC分类号,本文在1985—2019年的所有专利中共识别出8016911条数字专利。然而,根据本文构建的有监督的机器学习算法,共识别出13779356条数字专利。通过匹配专利申请号,本文发现根据关系表的IPC分类号识别出的8016911条数字专利有1146112匹配不上根据算法识别出的数字专利。为了更进一步探究这些专利是否内嵌或者本身就是一种数字技术,本文从中随机抽取了500条来进行每个人“背对背”识别,发现其中共462条与数字技术无关,主要是因为其标题和摘要具有迷惑性,因此在分类过程中可能存在错误归纳,可以认为这100余万条专利中有92%是与数字技术无关的。而对根据本文算法识别的、且没有匹配上根据关系表IPC分类号识别的数字技术专利共6908557条,根据最终训练和微调后算法的95.19%准确率,本文认为这接近700万条的专利中约有95%是数字技术专利。关于数字技术的具体分类、类别描述及其对应的数字知识类型参见《中国工业经济》网站(ciejournal.ajcass.com)附件。

② 基于BERT模型的具体识别思想和步骤参见《中国工业经济》网站(ciejournal.ajcass.com)附件。

其中, i 表示城市对, t 表示年份。 Bal 表示每个城市对的协调发展程度, 是城市对的两个城市中, 相对欠发达城市的发展水平与相对发达城市的发展水平之比。 $\ln DKF_LtoH$ 为相对欠发达城市流向相对发达城市的数字知识流量, $\ln DKF_HtoL$ 则是方向相反流量。两个解释变量都进行了加 1 取对数处理, 这一方面能够缓解数据的长尾形态和异方差问题, 另一方面可以规避忽视零知识流量带来的样本选择偏误问题。 X' 是一系列控制变量, 包括城市之间是否开通高铁 HSR 、城市之间的技术相似度 TS , 以及城市对整体的知识产权保护水平 IPP 。 λ_i 是城市对层面的固定效应, 用以控制距离等非时变变量。 λ_t 表示年份层面的固定效应。 ε_{it} 为随机扰动项。

为缓解反向因果问题, 本文解释变量做滞后一期处理, 相应地, 模型中所有涉及对两个城市相对发展水平高低的判断也均基于各城市滞后一期的发展水平。因此, Bal 的取值有可能大于 1。出现这种情况说明 $t-1$ 年相对欠发达的城市在 t 年实现了经济反超, 符合区域差距动态变化的现实逻辑。反之, 若不基于滞后的发展水平判断, 在城市相对发展水平发生反转的情况下就会错误地识别数字知识的流向, 进而造成偏误。

式(1)存在一个模型设计的潜在问题, 即 $\ln DKF_LtoH$ 和 $\ln DKF_HtoL$ 之间可能存在较为严重的共线问题, 因此, 在通过式(1)初步验证假说 1 后, 本文主要使用下式做进一步分析:

$$Bal_{it} = \alpha_0 + \alpha_1 \ln DKF_net_{it-1} + \beta X'_{it-1} + \lambda_i + \lambda_t + \varepsilon_{it} \quad (2)$$

其中, 变量符号及含义均与式(1)相同。区别在于, 本文不再同时控制 $\ln DKF_LtoH$ 和 $\ln DKF_HtoL$, 取而代之的是将发达城市净流入欠发达城市的数字知识量 $\ln DKF_net$ 作为核心解释变量。需要说明的是, 由于净流量为负时无法进行取对数处理, 此处 $\ln DKF_net$ 的计算方式为 $\ln DKF_HtoL$ 与 $\ln DKF_LtoH$ 之差。

3. 变量和数据说明

(1) 被解释变量。作为城市对中相对欠发达城市与相对发达城市的发展水平之比, 被解释变量 Bal 本身由各城市的经济增长情况决定, 同时能很好地刻画区域平衡发展的深层次目标, 是代理协调发展水平的合理指标。其中, 城市的发展水平使用人均夜间灯光亮度代理。其合理性在于, 在长时间序列条件下, 中国城市层面经济社会数据存在时空连续性不足的问题, 而夜间灯光亮度能有效代理区域经济活动强度和发展水平 (Henderson et al., 2012), 且相关栅格数据具有覆盖面大、分辨率高和更新频繁等优势 (Gibson et al., 2021), 结合行政区划矢量图, 还能克服因区划变动导致的统计数据不可比问题, 因此, 许多国内学者将其应用于中国区域经济的研究中 (刘修岩等, 2017; 邓慧慧等, 2022; 彭刚等, 2023; 郭峰等, 2023)。为了在尽可能延长样本期的同时兼顾研究的时效性, 本文最终使用 Zhong et al. (2022) 研发的 2000—2020 年中国长时间序列类 EANTLI 夜间灯光数据集。与其他同类数据相比, 该数据具有更高的稳定性、连续性和经济相关性, 能很好地量化区域长期经济发展水平。借鉴 Henderson et al. (2012) 分解夜间灯光亮度的思想, 本文结合类 EANTLI 数据、LandScan 人口栅格数据和中国城市行政区划矢量图,^① 在利用 ArcGIS 实现更适配中国地形的亚洲北半球兰伯特等角圆锥 (Asia North Lambert Conformal Conic) 投影坐标系转换后, 测算了各城市的人均夜间灯光亮度, 进一步得到城市对层面的被解释变量。

(2) 核心解释变量。 $\ln DKF_HtoL$ 、 $\ln DKF_LtoH$ 和 $\ln DKF_net$ 均来源于国家知识产权局知识产权

^① LandScan 全球人口分布数据来自 East View Cartographic, 由美国能源部橡树岭国家实验室 (ORNL) 开发。LandScan 运用 GIS 和遥感等创新方法, 是全球人口数据发布的社会标准, 也是全球最为准确、可靠、基于地理位置的、具有分布模型和最佳分辨率的全球人口动态统计分析数据。关于该数据更详细的介绍可参见其网站 (<https://landscan.ornl.gov/>)。

出版社。该数据包包含名称、申请日、专利类型、分类号、参考文献、邮编、地址和摘要等专利基础信息。为了识别工作的顺利开展,本文首先对专利数据进行了一些预处理工作:①授权的发明专利会公开第二次,为确保唯一性,剔除其在授权前公开的信息。②由于专利数据缺乏能直接识别所在省市区县的信息,本文依据专利信息中的邮编和地址信息进行识别:第一步,对于邮编能够直接识别的省市区县,予以匹配;第二步,对于邮编缺失和无法识别的专利,本文采用MGeo多任务多模态地址预训练底座并结合人工校准进行识别。③由于外观专利并不具备数字技术的内涵,将外观专利剔除,仅保留发明专利和实用新型专利。随后,本文基于前文所述的BERT大语言模型方法识别数字技术专利,并利用专利引用信息测算了各城市之间的数字知识流动情况。具体来说,在城市对的两座城市之中,一座城市的专利(包括数字技术专利和非数字技术专利)每有一次对另一座城市的数字技术专利引用,记为一次数字知识流入。

(3)控制变量。城市之间互通高铁为要素流动提供了快速通道,这不仅会影响数字知识的流动情况,也可能会进一步影响城市之间的相对发展差距。本文根据国家铁路局网站和中国铁路12306网站的高铁线路名称、开通时间以及经停站数据,手动整理了样本期内两两城市的高铁互通情况,记为 HSR 。该指标取1时,代表城市对已互通高铁,反之取0。城市之间的技术相似度也可能影响区域差距。一方面,城市间的技术相似度越高,欠发达城市吸收数字知识的效果可能越好。另一方面,这种技术相似度也可能加剧发达城市对欠发达城市的“虹吸效应”,进而扩大差距。技术相似度借鉴叶静怡等(2019)的方法计算,数据来源于国家知识产权局知识产权出版社。此外,城市的知识产权保护制度也会在很大程度上影响知识流动对区域经济的作用(白俊红等,2017),因此,本文控制了城市对整体的知识产权保护水平 IPP ,以城市对中的知识产权示范城市数量代理,取值0、1或2,数据来源于国家知识产权局。

(4)描述性统计和特征性事实分析。①为规避突发公共卫生事件的干扰,最终样本时间跨度为2001—2019年,覆盖由296个地级及以上城市单位两两组成的43660个无序城市对。在完成变量匹配并剔除少量极端值后,本文尝试对数字知识流动的特征进行初步观测。具体来说,本文设定在一个城市对中,若发达城市在样本期内的平均发展水平属于最高的25%,且欠发达城市属于最低的25%,该城市对的发展差距较大,否则差距较小。数据显示,数字知识在发达城市和欠发达城市之间的流动在21世纪初较低,不过随着时间的推移逐渐增多,且在差距较大的情况下,从发达城市流入欠发达城市的数字知识的增长倍率显著高于其他知识流。可能的原因是,数字知识流入这些发展水平属于最低分位的城市后,能发挥更加积极的作用,因此,自发的知识流增长更快,这也暗示数字知识流动可作为区域协调发展路径的可靠性。同时,当城市间差距较小时,虽然数字知识流的增速较慢,但其流动量显著更高。一个可能的解释是,发展水平相近的城市相互间的知识适配性更强,因此,数字知识流入后能转化的经济价值更高,自发的流动也更多。对此观点的进一步检验参见异质性分析部分。

四、实证分析

1. 基准回归和稳健性检验

表1第(1)—(3)列展示了基于式(1)的回归结果。第(1)列的回归仅包括核心变量和年份固定

① 描述性统计结果和特征性事实分析数据参见《中国工业经济》网站(ciejournal.ajcass.com)附件。

效应,第(2)、(3)列逐步加入城市对固定效应和控制变量。结果显示,lnDKF_HtoL的估计系数显著为正,说明从发达城市流入欠发达城市的数字知识的确能够弥合区域发展差距,促进区域协调发展;相应地,lnDKF_LtoH的估计系数显著为负,同样符合预期,因为数字知识流入发达城市会促进其经济增长,在其他条件不变的情况下,会在客观上放大发达城市和欠发达城市之间的差距。进一步地,第(4)列展示了式(2)的估计结果,在使用lnDKF_net作为核心解释变量以排除可能的共线性问题后,其估计系数显著为正,假说1得到验证。为了增强基准回归结论的可靠性,本文也做了一系列稳健性检验,如替换变量测度方式、剔除部分样本等,均得到了稳健的结果。^①

表1 基准回归结果

变量	(1)	(2)	(3)	(4)
lnDKF_HtoL	0.0143*** (0.0010)	0.0179*** (0.0007)	0.0211*** (0.0007)	
lnDKF_LtoH	-0.0111*** (0.0010)	-0.0271*** (0.0007)	-0.0235*** (0.0007)	
lnDKF_net				0.0222*** (0.0005)
控制变量	否	否	是	是
城市对固定效应	否	是	是	是
年份固定效应	是	是	是	是
观测值	789394	789394	773145	773145
R ²	0.0669	0.7128	0.7149	0.7149

注：*、**和***分别表示在10%、5%和1%的水平上显著；括号内为聚类在城市对层面的稳健标准误；常数项估计结果省略，控制变量包括HSR、TS和IPP。以下各表同。

2. 内生性问题分析

虽然本文在模型设定和变量计算过程中做了大量滞后处理,并采用较为严格的固定效应,但是考虑到城市经济发展水平越高,其吸收的数字知识量可能越大,且两个城市当期的相对差距和滞后一期的差距往往高度正相关,因而反向因果问题依然可能存在,同时遗漏变量问题无法被完全消除。鉴于此,本文通过工具变量法进一步处理潜在的内生性问题。

本文以数字技术专利的引用情况代理数字知识流动,而由于近现代通信基础设施的普及,大幅提高了编码知识的远距离传输效率,以专利引用为代表的编码知识流动通常高度依赖通信网络。当代的城市间通信网络虽然与数字知识流动具有很强的相关性,但是囿于内生性问题,并不是合适的工具变量。与此同时,历史基础设施的分布通常对当代经济活动具有长期且外生的影响(Duranton and Turner, 2012)。就历史通信网络而言,一方面,其在很大程度上影响着当代城市间通信网络的形成,从而塑造着数字知识流动的路径,因此,其与核心解释变量之间具有正相关性;另一方面,决定历史通信网络分布的通常是过去的政治目的和军事需求,而非当代的经济活动,因此,相关指标也具有很强的外生性。鉴于此,本文尝试构建能代理历史通信网络强度的工具变量。

本文首先使用1984年固定电话机数据构造指标,这是能够直接获取的最早的城市层面通信基础设施数据。本文计算了每个城市对之中,两座城市的固定电话机数量的乘积,进一步计算其与全

^① 稳健性检验内容参见《中国工业经济》网站(ciejournal.ajcass.com)附件。

部城市对该指标加总之比,以此量化两两城市之间的历史通信网络强度。由于该指标为截面数据,借鉴 Nunn and Qian(2014),本文将该指标与全国层面发达城市净流入欠发达城市的总数字知识流量交乘,得到最终的工具变量,记为 IV_hist_tel 。相应两阶段最小二乘回归结果稳健,一阶段工具变量和二阶段拟合的核心解释变量估计系数均显著为正。但遗憾的是,观察样本量可以发现,由于行政区划变动等原因,以文本形式记载的历史通信基础设施数据存在较为严重的缺失问题。尤其是在本文以城市对为样本的情况下,只有两座城市都不存在数据缺失问题才是有效样本,这进一步放大了该问题。考虑到历史 GIS 数据能很好解决行政区划变动导致数据缺失的问题,对中国近现代通信网络地图电子化以获取相关数据,成为一个可行的选择。

为了获取更为合适的历史通信网络工具变量,本文尝试采用 ArcGIS 软件对光绪三十三年(1907年)《中国电线图》进行数字化地理配准,以此构建城市对层面的历史通信网络强度数据。该地图为上海通商海关注册处印制的《大清邮政舆图》中的一部分,详细记录了清朝末期全国范围内的有线电报局及电报线路分布情况。自“洋务运动”以来,清朝积极引入以电报为代表的通信技术。截至 1907 年,全国有线电报网络已基本覆盖沿海主要港口、内陆经济中心和部分战略节点,且初衷大多服务于军事需求。与此同时,该地图绘制于清朝灭亡前夕,意味着 1907 年的通信网络布局在很大程度上定型,并在此后长期影响中国的区域经济和知识流动格局。此外,采用有线电报数据的另一个优势在于,相比固定电话,电报线的连接更加能直接反映由城市对构成的通信网络。相关具体工作包括:①地图数字化与配准。通过地理配准工具,以现代行政区划图为准,选取清晰的地理标识作为控制点,完成《中国电线图》到现代地理坐标系的精准对齐。②电报局点要素创建。根据《中国电线图》上标注的有线电报局位置,逐一手动绘制点要素,以最大限度减少行政区划变动带来的误差。③利用空间分析工具汇总每个城市的历史电报局数量。^①

基于该城市层面数据,与前文做法一致,本文计算了每个城市对之中,两座城市的电报局数量的乘积,进一步计算其与全部城市对的该指标加总值之比,以此量化两两城市之间的历史通信网络强度。最终将该指标与全国层面发达城市净流入欠发达城市的数字知识流量交乘,同样记为 IV_hist_tel 。相应两阶段回归结果均符合预期,再次验证了数字知识流动促进区域协调发展的积极作用。

3. 机制检验

下面将聚焦于数字知识流动促进区域协调发展的内在逻辑,依次检验知识适配性、需求适配性以及要素禀赋差距在上述过程中扮演何种角色,以论证假说 2。本文的检验方式基于调节效应思想,即在基准回归模型的基础上进一步构建核心解释变量与调节变量的交乘项模型,并估计和检验交乘项的系数。

(1)知识适配性。如果数字知识流动促进区域协调发展的前提是欠发达城市供给侧适配于数字知识,考虑到数字产业是承接数字知识的关键,那么,欠发达城市的数字产业发展水平越高,供给侧的知识适配性越强,数字知识流入当地后其促进区域协调发展的边际作用应当越强,因此,代理欠发达城市数字产业水平的指标应当具有正向调节效应。借鉴洪俊杰等(2024)的做法,本文主要根据城市的数字经济企业数量测算调节变量。具体步骤为:①从 RESSET 企业大数据平台获取企业工商登记数据,包括企业名称、所属行业、所属城市、注册时间和经营范围等,并根据

^① 《中国电线图》扫描件和配准结果数据,以及内生性问题检验结果参见《中国工业经济》网站(ciejournal.ajcass.com)附件。

所属行业对企业进行初步筛选^①。②根据《数字经济及其核心产业统计分类(2021)》的行业解释和企业信息做进一步的手动筛选,如果某行业本身就是该分类中罗列的产业,那么直接认定该行业的企业为数字经济企业;如果行业中存在部分产业属于数字经济范畴,则根据该分类的行业解释设置关键词,结合企业的经营范围筛选出属于数字经济核心产业的企业。③考虑到数字产业的发展与迭代速度较快,加总各城市对中相对欠发达城市每年新增的数字经济企业数并取对数,以此代理供给侧对数字知识的适配性,记为 $Compat_kno$ 。为尽可能保证调节变量的外生性,在式(2)的基础上,本文在解释变量中加入了滞后一期的 $Compat_kno$ 及其与核心解释变量的交乘项,相应结果见表2第(1)列。结果显示,交乘项系数显著为正,欠发达城市数字产业发展水平具有正向调节效应,符合预期。②为了增强结论的稳健性,本文还替换了 $Compat_kno$ 的计算方式,将其定义为欠发达城市每年新增数字经济企业的注册资本额,交乘项模型结果见表2第(2)列,同样符合预期。

表2 关于知识适配性和需求适配性的机制检验结果

变量	(1)	(2)	(3)	(4)
	知识适配性:数字产业发展水平		需求适配性:偏好差异化程度	
$\ln DKF_net$	-0.0111*** (0.0040)	-0.0366*** (0.0040)	0.0206*** (0.0007)	0.0230*** (0.0002)
$\ln DKF_net \times Compat_kno$	0.0049*** (0.0006)	0.0048*** (0.0003)		
$\ln DKF_net \times Compat_dem$			0.0012*** (0.0003)	-0.0002 (0.0014)
控制变量	是	是	是	是
城市对固定效应	是	是	是	是
年份固定效应	是	是	是	是
观测值	772873	772319	707268	702638
R ²	0.7156	0.7172	0.7143	0.7147

注:各调节变量本身均已得到控制,以下各表同。

(2)需求适配性。提升需求适配性是数字知识流动促进区域协调发展的关键之一,而其着力点在于满足消费者差异化偏好(马述忠等,2024)。因此,在欠发达城市面临的市场范围内,消费者偏

① 与洪俊杰等(2024)的做法一致,本文只保留属于“信息传输、软件和信息技术服务业”“批发和零售业”“租赁和商务服务业”“科学研究和技术服务业”4个一级行业和“计算机、通信和其他电子设备制造业”“通用设备制造业”“专用设备制造业”“电气机械和器材制造业”“印刷和记录媒介复制业”“文教、工美、体育和娱乐用品制造业”“化学原料和化学制品制造业”“仪器仪表制造业”“土木工程建筑业”“建筑安装业”“房屋建筑业”“机动车、电子产品和日用产品修理业”“货币金融服务”“其他金融业”“广播、电视、电影和录音制作业”“新闻和出版业”16个二级行业的企业。

② 在上述调节效应检验中,一个潜在的担忧是知识流入可能对同期的数字产业建立存在影响,但好的调节变量应当具有较强的外生性,不受处理变量或结果变量的影响(江艇,2022)。鉴于此,本文考察了知识流入对城市数字产业的影响。整体而言,直到流入后第4期,数字知识才会对新数字产业的建立产生显著影响,这很好地缓解了本文的担忧。

好的差异化程度越高,需求适配性提升的空间就越大,潜在的差异化溢价也就越高,数字知识流入促进区域协调发展的效果应当越强。遗憾的是,由于目前缺乏覆盖中国全国范围的消费数据库,直接测度偏好差异化程度并不现实。借鉴郭继文和马述忠(2022)的做法,本文通过城市进口的分散化程度来近似反映消费偏好差异化程度。其合理性在于,进口偏好在一定程度上反映了市场整体的消费模式,差异化的进口倾向应当意味着消费者对各类产品的需求都更为多元,这也会提高本地产品的市场匹配难度。为兼顾调节变量的外生性和数据的可得性,本文基于中国海关数据库2000年数据构造指标。本文将初始的微观企业进口数据汇总至城市—进口国—产品层面后,依据泰尔指数公式测算了城市从不同国家进口同一类产品金额的离散程度,进一步基于此城市—产品层面指标得到城市层面的均值。^①该指标值越小,表示进口离散程度越高,那么该城市消费者的偏好差异化程度也越高。因此,本文取该指标的倒数来代理各城市消费者的偏好差异化程度,并将城市对中欠发达城市的该指标记为 *Compat_dem*,相应调节效应检验结果见表2第(3)列,调节效应为正,符合预期。此外,本文还尝试将 *Compat_dem* 替换为和欠发达城市同省其他城市消费者的偏好差异化程度均值,结果如第(4)列所示。然而,此时交乘项的估计系数不再显著。可能的解释是,即使在数字知识的助力下,城市匹配消费者差异化偏好和个性化需求的能力依然存在瓶颈,可能大多止步于市域尺度。因此,为充分发挥数字知识流动促进区域协调发展的积极作用,城市亟待进一步增强匹配外部市场消费者差异化偏好的能力。

(3)要素禀赋差距。本文认为,缓解要素禀赋差距是数字知识提升欠发达地区需求适配性、进而促进区域协调发展的关键路径。因此,如果欠发达城市的各类要素禀赋和发达城市的差距越大,数字知识流入当地后,弥合区域差距的边际效应越强。反过来说,代理欠发达城市要素禀赋相对可及性的调节变量应具有负向调节效应。为了满足调节变量的外生性要求,本文的测算目标是样本期之前欠发达城市相对发达城市的初始禀赋可及性。

资本要素方面,本文计算了欠发达城市和发达城市2000年的人均金融机构贷款余额之比,以此代理欠发达地区初始对资本要素的相对可及性,记为 *Gap_k*。相应调节效应检验结果见表3第(1)列,交乘项系数估计结果显著为负,符合预期。为了增强结论的可靠性,本文还采用人均固定资产净值构造指标,相应检验结果见表3第(2)列,同样稳健。

劳动力要素方面,吸纳中高端劳动力是数字知识助力欠发达城市提升需求适配性的着力点。这是因为上述积极作用的发挥具有一定门槛,需要参与要素市场的劳动者具有熟练运用在线搜索、远程办公等数字技术的能力(白俊红等,2017)。由于受教育程度是劳动力层次的重要划分依据,本文根据第五次全国人口普查数据,计算了欠发达城市和发达城市2000年文盲率之比并进行逆序处理,以此代理初始时期欠发达地区对中高端劳动力的相对可及性,记为 *Gap_l*。同时,本文采用了平均受教育年限之比做稳健性检验。相应结果见表3第(3)、(4)列,结果均显示指标具有负向调节效应,符合预期。

除了以资本和劳动力为代表的传统生产要素,数据要素已成为数字时代的核心生产力之一,可能成为数字知识流动促进区域协调发展的重要抓手。与此同时,数字知识流动本身可被视为一种区域数字鸿沟的弥合路径。鉴于此,本文还尝试考察数字知识流动能否缓解数据要素分布不均的问题。为兼顾指标的科学性和可得性,本文从数据生成的视角切入,使用欠发达城市和发达城市2000年国际互联网用户数之比,代理欠发达城市对数据要素的相对可及性,记为 *Gap_d*。相应调节

^① 为缓解零贸易流带来的偏差,本文使用HS2位码层面的产品数据。

表3 关于要素禀赋差距的机制检验结果

变量	(1)	(2)	(3)	(4)	(5)	(6)
	资本要素差距		劳动力要素差距		数据要素差距	
$\ln DKF_{net}$	0.0245*** (0.0006)	0.0229*** (0.0006)	0.0198*** (0.0007)	0.0397*** (0.0031)	0.0225*** (0.0006)	0.0228*** (0.0006)
$\ln DKF_{net} \times Gap_k$	-0.0036*** (0.0006)	-0.0017*** (0.0004)				
$\ln DKF_{net} \times Gap_l$			-0.0016*** (0.0003)	-0.0182*** (0.0003)		
$\ln DKF_{net} \times Gap_d$					-0.0004*** (0.0001)	-0.0012*** (0.0003)
控制变量	是	是	是	是	是	是
城市对固定效应	是	是	是	是	是	是
年份固定效应	是	是	是	是	是	是
观测值	716252	487379	721113	721113	623635	623635
R ²	0.7216	0.7241	0.7166	0.7191	0.7232	0.7239

效应的检验结果见表3第(5)列,交乘项系数的估计结果显著为负,说明数字知识流动缓解要素分布不均作用对数据要素同样适用。鉴于数据要素具有梅特卡夫定律(Metcalfé's Law)提出的无形性和外部性,本文也采用国际互联网用户数占总人口比例之比重新计算该指标,交乘项模型估计结果同样稳健,见第(6)列。

五、异质性分析

《数字经济促进共同富裕实施方案》指出,要加强数字经济东西部协作。这为推动区域数字协同发展、助力数字经济促进共同富裕指明方向。然而,当前区域间经济互动性弱依然是掣肘中国区域协调发展的重要瓶颈(张勋和乔坤元,2016)。与此同时,东密西疏的知识网络特征可能会进一步加剧数字知识流动促进区域经济增长的不对称性(蒋伏心和高丽娜,2012;曹湛等,2022)。因此,为更好地促进方案思想落实,有必要考察东部城市间、中西部城市间,以及东部和中西部城市间数字知识流动促进区域协调发展的作用存在何种差异。

本文首先对东部地区城市间,以及中西部城市间数字知识流动的积极作用进行对比。具体来说,当城市对中的两座城市均属于东部城市时,变量 $East_East$ 取1,否则取0;均属于中西部城市时,变量 $West_West$ 取1,否则取0。表4第(1)、(2)列相应的调节效应检验结果显示,在中西部地区内,数字知识流动的促进区域协调发展的作用更强。可能的解释是,相比东部地区,中西部城市整体的数字技术水平较低,提升空间更大,此时数字知识流动能为当地欠发达城市带来更高的边际收益。

当然,本文更加关心的还是东部和中西部城市之间的跨区域数字知识流动。本文定义当城市对中的两座城市分别属于东部和中西部城市时,变量 $East_West$ 取1,否则取0。然而,表4第(3)列的调节效应检验结果显示,这种横跨东西的数字知识流动并没有带来额外的收益。一个可能的解释是,这种数字知识流动其实包含两种情况:一是城市对中相对发达的城市为东部城市,欠发达城

表4 区域特征异质性分析结果

变量	(1)	(2)	(3)	(4)
	东部—东部	中西部—中西部	东部—中西部	
$\ln DKF_{net}$	0.0232*** (0.0006)	0.0210*** (0.0006)	0.0223*** (0.0008)	0.0175*** (0.0019)
$\ln DKF_{net} \times East_East$	-0.0055*** (0.0013)			
$\ln DKF_{net} \times West_West$		0.0040*** (0.0012)		
$\ln DKF_{net} \times East_West$			-0.0001 (0.0010)	
$\ln DKF_{net} \times East_dev$				0.0060*** (0.0021)
控制变量	是	是	是	是
城市对固定效应	是	是	是	是
年份固定效应	是	是	是	是
观测值	773145	773145	773145	352095
R ²	0.7149	0.7149	0.7149	0.7252

市为中西部城市；二是位于中西部地区的城市相对发达。考虑到东部地区整体的数字创新更活跃（胡增玺和马述忠，2023），从中西部地区流入东部地区的数字知识可能无法促进当地经济发展，因此，本文无法在全样本条件下观测到显著的结果。于是，本文定义虚拟变量 $East_dev$ ，仅当城市对中的东部城市更发达时，该变量取1，否则取0。第(4)列结果表明，在两座城市分别属于东部和中西部城市的前提下，该变量具有正向调节效应。可能的解释是，得益于对口支援和东西协作等制度在实践中不断完善（张可云等，2023），东部发达城市流入中西部欠发达城市的数字知识能起到较为显著的积极作用。此外，考虑到东部和中西部城市间经济社会特征的差异，应当大于东部内部或中西部内部城市间的差异，第(3)列结果的一个潜在解释是，当城市之间的差异过大时，欠发达城市可能无法很好吸收从发达城市流入的数字知识。为深入探讨该问题，本文还尝试基于城市对的经济社会差异展开分析。

在经济社会异质性上，一方面，本文关注以人口数量为代表城市规模的差距。参考国务院印发的《关于调整城市规模划分标准的通知》，本文根据城市在2000年的人口定义指标 $Scale$ ，当城市属于II型小城市、I型小城市、中等城市、II型大城市、I型大城市、特大城市和超大城市时， $Scale$ 分别取值1—7。进一步构造变量 Gap_scale ，其计算方式为城市对之中，相对发达城市的 $Scale$ 减去欠发达城市的 $Scale$ 。相应交乘项模型的检验结果见表5第(1)列，交乘项系数显著为正，这说明发达的大城市能够通过数字知识流动有效带动欠发达小城市的发展。一个可能的解释是，大城市的数字知识更多在规模经济的市场环境下诞生，而在数字技术整合在线需求的背景下，这些发达大城市的知识流入欠发达小城市后，能够激发小城市充分利用在线市场获益。但是反过来，发达小城市的知识流入欠发达的大城市，无法起到类似的效果，因为大城市已经处于一个较大的市场中，数字知识打破市场边界的边际效用较小。为了增强结论的稳健性，本文采用1989年的四级城市旧标准重

表 5 经济社会特征异质性分析结果

变量	(1)	(2)	(3)	(4)
	城市对规模差距		城市对发展水平差距	
$\ln DKF_{net}$	0.0222*** (0.0005)	0.0224*** (0.0005)	0.0236*** (0.0006)	0.0218*** (0.0006)
$\ln DKF_{net} \times Gap_{scale}$	0.0017*** (0.0004)	0.0108*** (0.0021)		
$\ln DKF_{net} \times Gap_{light}$			-0.0102*** (0.0010)	-0.0115*** (0.0009)
控制变量	是	是	是	是
城市对固定效应	是	是	是	是
年份固定效应	是	是	是	是
观测值	773145	773145	773145	773145
R ²	0.7149	0.7149	0.7149	0.7238

新计算相关指标,结果没有发生实质性变化,见第(2)列。另一方面,延续前文分析思路,本文进一步探讨了城市间发展水平差距的不同如何影响数字知识流动促进区域协调发展的作用。具体来说,当一组城市对中,欠发达城市样本期内人均夜间灯光亮度属于最低的25%,同时发达城市属于最高的25%时,变量 Gap_{light} 取值1,否则为0。第(3)列的结果显示,交乘项的估计系数显著为负,说明当城市发展水平相近时,数字知识流动促进区域协调发展的作用更强。这很可能是因为在发展水平相近的城市间,流入地对数字知识的吸收能力更强,知识适配性更高,同时也在一定程度上解释了为何前文会发现差距较小的城市间,自发的数字知识流动更多。为了提高结论的稳健性,本文也仅根据各城市2000年的发展水平重新进行分位计算,相应结果见第(4)列,交乘项系数依然显著为负,符合预期。

此外,本文还针对核心议题进行了一些拓展分析。一方面,市域尺度的分析表明,从城市外部流入的数字知识并不会显著影响其内部区县间的协调发展水平,不过从中心区县向外围区县的净数字知识流入则能够缓解“中心—外围”结构差异。另一方面,长期效应分析表明,随着时间的推移,累积的数字知识对发达城市的积极作用会逐渐减弱直至消失,但对欠发达城市的经济增长具有更稳定的驱动作用。换言之,数字知识流动促进区域协调发展的作用在长期会得到进一步强化。与此同时,虽然数字知识流动具有培育新数字产业的长期效应,但这一过程并非一蹴而就,而是存在明显的滞后性。^①

六、结论与政策启示

在数字经济时代,作为构成数字技术的必要条件,数字知识流动是促进区域协调发展的重要力量。本文采用前沿的大语言模型BERT技术和专利摘要文本,首次量化了数字知识的跨区域流动,并结合类EANTLI夜间灯光数据集和LandScan人口栅格数据,基于经济增长和平衡发展双重目标,

^① 拓展分析的详细内容参见《中国工业经济》网站(ciejournal.ajcass.com)附件。

考察了数字知识流动对区域协调发展的影响。研究发现,由发达地区流入欠发达地区的数字知识能够促进区域协调发展,并且这一积极作用以欠发达地区供给侧对数字知识的适配性为前提,以数字知识改善需求适配性进而缓解要素禀赋差距为关键路径。异质性分析表明,当数字知识从东部发达城市流向中西部欠发达城市时、从大规模发达城市流向小规模欠发达城市,以及在发展水平相近城市间流动时,其促进区域协调发展的效果更好。在市域范围上,城市内部中心区县流向外围县区的数字知识能够显著缓解“中心—外围”结构差异。长期看,相比发达城市,数字知识对欠发达城市的经济增长具有更稳定的驱动作用。

随着中国进入扎实推动共同富裕的新阶段,本文的研究结论对于如何在不断做强做优做大数字经济的同时助力区域协调发展、促进公平与效率更加统一具有丰富的政策启示意义。

(1) 优化区域数字产业结构,改善供给侧对数字知识的适配性。鉴于数字知识的有效流动依赖于流入地供给侧对外部知识的契合度,政府应根据地方产业结构和技术需求,因地制宜制定知识输出策略,引导特定数字知识的流入。例如,优先向农业主导地区引入数字农业技术,如精准农业、智能灌溉等,而在制造业发达地区则侧重于工业互联网和智能制造技术。同时,地方高校和科研机构是推动知识流入的关键角色,建议设立数字知识引导基金,支持地方高校与发达地区科研力量合作,通过联合实验室、研发中心等形式引入前沿技术并加速本地化转化。高校应承担起技术人才的培养责任,保障数字知识的持续适配。此外,针对欠发达地区劳动力技术培训的薄弱环节,政府应加强数字技能基础教育和职业培训的结合,通过在线学习平台推广高质量的技能课程,提升劳动力对数字技术的应用能力和地方企业的技术承接能力,确保流入地与发达地区的技术能力差距持续缩小。

(2) 建立灵活的需求反馈机制,增强流入地供给侧利用数字知识改善需求适配性的能力。数字知识流动促进区域协调发展的关键在于,助力流入地供给侧满足消费者的差异化需求。为此,地方政府应通过定期调研建立灵活的需求反馈机制,推动数字经济应用场景的拓展,使流入的数字知识能够更好地贴合国内大市场的消费需求。建议地方政府设立数字经济需求对接平台,整合企业需求与外部技术资源,动态匹配供需信息,提升数字知识的应用效率,强化其促进新数字产业建立的长期效应。在政策引导下,应激励地方企业积极参与数字化转型,推动农村电商、智能制造和数字农业等领域发展,提升对数字技术的利用效率。例如,在数字农业领域,通过物联网和大数据技术优化生产流程,推动精准农业发展;而在工业领域,依托工业互联网提升全要素生产率。此外,地方企业需通过大数据分析和市场调研,动态调整数字知识的吸收方向,确保其与消费者日益多样化的需求相匹配,从而全面打通国内市场。考虑到数据要素在供需匹配中的关键作用,政府还应推动区域间数据共享机制建设,通过设立跨区域数据协同平台,消除“数据孤岛”现象,为地方企业提供高质量的市场洞察力。同时,为帮助后发地区在外部市场的竞争中突破差异化偏好的瓶颈,建议建立全国性区域消费偏好数据库,并配套数据要素市场化流通规则。地方企业可以依托该数据库,结合人工智能算法开发定制化产品和服务,从而提升后发地区对外部市场需求的响应速度和竞争力。

(3) 有针对性、有优先级地促进数字知识的跨区域流动。根据异质性分析结论,政府应优先推动东部大规模发达城市与中西部小规模欠发达城市之间的合作,建立长期技术交流机制,促进数字知识更多向欠发达地区流动。具体而言,政府可以建设跨区域数字知识共享平台,推动发达地区的科研成果、创新技术和经验加速传递。通过共享平台,发达地区的企业、高校和科研机构可以及时向欠发达地区传递最新的数字技术,帮助其获取适合本地的创新资源。与此同时,应注重促进城市

内部知识的高效流动。政府可以鼓励中心区县向外围区县输出数字知识,通过设立知识辐射中心和区域共享实验室,以及完善基础设施研发建设、专业人才培养和技术服务体系,促进技术和经验在城市内部不同区域间的扩散,支持外围区县建设数字产业园区,增强其对中心区县技术外溢的承接能力,实现城市内区域协调发展。针对欠发达地区技术吸收能力较弱的问题,建议政府设立数字知识引导基金,专门用于支持企业和科研机构提升数字技术承接能力,包括研发基础设施建设、专业人才培养和技术服务体系的完善。此外,政府应制定明确的政策激励措施。例如,通过财政支持和奖励计划,鼓励发达地区企业在后发地区设立分支机构或联合实验室,推动创新技术的转移与落地,全面提升数字知识跨区域流动的质量与效率,为推动区域协调发展提供强有力的保障。为确保政策落实到位,可设立绩效评估机制,对跨区域知识流动的实际成效进行动态监测和评估,以便及时优化相关政策举措。

需要说明的是,虽然本文基于知识适配性、需求适配性和要素禀赋差距等相对宏观的视角探讨了数字知识流动如何促进区域协调发展,但该议题在微观和中观层面仍存在广阔探索空间,诸多关键问题有待深入挖掘。一方面,企业是区域经济的微观基石,其创新成果转化效率直接影响区域经济活力。未来可探究数字知识在企业内部的微观传播路径,例如,分析数字知识如何跨越企业部门界限以加速创新成果转化,进而为区域经济增长注入动力。另一方面,产业集群是区域经济发展的重要引擎,其创新系统涵盖企业、高校、科研机构和政府等多元主体。未来可研究数字知识在产业集群内流动和溢出的中观机制,例如,分析其如何促进产学研深度融合,进而提升区域创新能力。

〔参考文献〕

- [1]白俊红,王钺,蒋伏心,李婧.研发要素流动、空间知识溢出与经济增长[J].经济研究,2017,(7):109-123.
- [2]曹湛,戴靓,杨宇,彭震伟.基于“蜂鸣—管道”模型的中国城市知识合作模式及其对知识产出的影响[J].地理学报,2022,(4):960-975.
- [3]邓慧慧,周梦雯,程钰娇.数字经济与城市群协同发展:基于夜间灯光数据的研究[J].浙江大学学报(人文社会科学版),2022,(4):32-49.
- [4]郭峰,熊云军,石庆玲,王靖一.数字经济与行政边界地区经济发展再考察——来自卫星灯光数据的证据[J].管理世界,2023,(4):16-33.
- [5]郭继文,马述忠.目的国进口偏好差异化与中国跨境电子商务出口——兼论贸易演变的逻辑[J].经济研究,2022,(3):191-208.
- [6]洪俊杰,李研,杨曦.数字经济与收入差距:数字经济核心产业的视角[J].经济研究,2024,(5):116-131.
- [7]洪银兴.区域共同富裕和包容性发展[J].经济学动态,2022,(6):3-10.
- [8]蒋伏心,高丽娜.区际知识溢出不对称、产业区位与内生经济增长[J].财贸经济,2012,(7):118-125.
- [9]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.
- [10]荆文君,孙宝文.数字经济促进经济高质量发展:一个理论分析框架[J].经济学家,2019,(2):66-73.
- [11]刘修岩,李松林,陈子扬.多中心空间发展模式与地区收入差距[J].中国工业经济,2017,(10):25-43.
- [12]陆铭,李鹏飞.城乡和区域协调发展[J].经济研究,2022,(8):16-25.
- [13]马述忠,张道涵,陈逸凡.跨境电子商务与市域协调发展——兼论跨境电商的普惠逻辑[J].经济学动态,2024,(3):77-95.
- [14]彭刚,杨德林,杨琳.中国市域尺度共同富裕水平格局及其影响因素[J].经济地理,2023,(1):44-54.
- [15]姚常成,沈凯琦.要素流动视角下数字经济与区域经济的包容性增长效应[J].经济地理,2023,(4):10-19.

- [16]叶静怡,林佳,张鹏飞,曹思未.中国国有企业的独特作用:基于知识溢出的视角[J].经济研究,2019,(6):40-54.
- [17]叶堂林,王雪莹.数字经济对协调性均衡发展的影响——兼论共同富裕的实现路径[J].经济学动态,2023,(1):73-88.
- [18]张可云,冯晟,席强敏.东西部协作政策效应评估——基于要素流动的视角[J].中国工业经济,2023,(12):61-79.
- [19]张勋,乔坤元.中国区域间经济互动的来源:知识溢出还是技术扩散[J].经济学(季刊),2016,(4):1629-1652.
- [20]赵勇,白永秀.知识溢出:一个文献综述[J].经济研究,2009,(1):144-156.
- [21]朱丰毅,桂文林.粤港澳区域知识溢出与经济增长[J].数量经济技术经济研究,2022,(3):44-65.
- [22]Arrow, K. J. The Economic Implications of Learning by Doing[J]. Review of Economic Studies, 1962, 29(3): 155-173.
- [23]Brynjolfsson, E., Y. Hu, and D. Simester. Goodbye Pareto Principle, Hello Long Tail: The Effect of Search Costs on the Concentration of Product Sales[J]. Management Science, 2011, 57(8): 1373-1386.
- [24]Duranton, G., and M. A. Turner. Urban Growth and Transportation[J]. Review of Economic Studies, 2012, 79(4): 1407-1440.
- [25]Faulkner, P., and J. Runde. On the Identity of Technological Objects and User Innovations in Function[J]. Academy of Management Review, 2009, 34(3): 442-462.
- [26]Faulkner, P., and J. Runde. Theorizing the Digital Object[J]. MIS Quarterly, 2019, 43(4): 1279-1302.
- [27]Gentzkow, M., B. Kelly, and M. Taddy. Text as Data[J]. Journal of Economic Literature, 2019, 57(3): 535-574.
- [28]Gibson, J., S. Olivia, G. Boe-Gibson, and C. Li. Which Night Lights Data Should We Use in Economics, and Where[J]. Journal of Development Economics, <https://doi.org/10.1016/j.jdeveco.2020.102602>, 2021.
- [29]Guellec, D., and C. Paunov. Digital Innovation and the Distribution of Income[R]. NBER Working Paper, 2017.
- [30]Henderson, J. V., A. Storeygard, and D. N. Weil. Measuring Economic Growth from Outer Space[J]. American Economic Review, 2012, 102(2): 994-1028.
- [31]Hund, A., H. T. Wagner, D. Beimborn, and T. Weitzel. Digital Innovation: Review and Novel Perspective[J]. Journal of Strategic Information Systems, <https://doi.org/10.1016/j.jsis.2021.101695>, 2021.
- [32]Jaffe, A. B., M. Trajtenberg, and R. Henderson. Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations[J]. Quarterly Journal of Economics, 1993, 108(3): 577-598.
- [33]Lucas, R. E. Jr. On the Mechanics of Economic Development[J]. Journal of Monetary Economics, 1988, 22(1): 3-42.
- [34]Nunn, N., and N. Qian. US Food Aid and Civil Conflict[J]. American Economic Review, 2014, 104(6): 1630-1666.
- [35]Owen-Smith, J., and W. W. Powell. Knowledge Networks as Channels and Conduits: The Effects of Spillovers in the Boston Biotechnology Community[J]. Organization Science, 2004, 15(1): 5-21.
- [36]Romer, P. M. Increasing Returns and Long-Run Growth[J]. Journal of Political Economy, 1986, 94(5): 1002-1037.
- [37]Singh, J. Collaborative Networks as Determinants of Knowledge Diffusion Patterns[J]. Management Science, 2005, 51(5): 756-770.
- [38]Yoo, Y., O. Henfridsson, and K. Lyytinen. Research Commentary—The New Organizing Logic of Digital Innovation: An Agenda for Information Systems Research[J]. Information Systems Research, 2010, 21(4): 724-735.
- [39]Zhang, X., and K. H. Zhang. How Does Globalization Affect Regional Inequality within a Developing Country? Evidence from China[J]. Journal of Development Studies, 2003, 39(4): 47-67.
- [40]Zhong, X. Y., Q. W. Yan, and G. E. Li. Development of the Time Series Nighttime Light Dataset of China (2000—2020) [J]. Journal of Global Change Data & Discovery, 2022, 3: 416-424.

How Digital Knowledge Flows Promote Regional Coordinated Development: On the Dual Goals of Economic Growth and Balanced Development

MA Shu-zhong¹, ZHANG Dao-han¹, HU Zeng-xi²

(1. China Academy of Digital Trade, Zhejiang University;

2. School of International Trade and Economics, Central University of Finance and Economics)

Abstract: Regional coordinated development is necessary for China to achieve high-quality economic growth, requiring sustained economic growth and balanced development on a per capita basis. As a key element of the era of the digital economy, digital knowledge flows may significantly impact regional coordinated development. This study focuses on the relationship between digital knowledge flows and regional coordinated development, aiming to provide policy guidance for balancing equity and efficiency in the era of the digital economy.

Using patent text data from the China National Intellectual Property Administration, this study employs a large language model to identify digital technologies covered in patents. Digital knowledge flows are then characterized based on patent citations. The findings indicate that digital knowledge inflows from developed to less developed regions contribute to regional coordinated development. Mechanism analysis reveals that under the premise of industry compatibility to digital knowledge in less developed regions, enhancing supply-side compatibility to demand to mitigate the gap in factor endowments is a key pathway for digital knowledge flows to promote regional coordinated development. Heterogeneity analysis shows that digital knowledge flowing from developed cities in the eastern region to less developed cities in central and western regions, from large developed cities to small less developed cities, and between cities with similar development levels has a more significant impact on regional coordinated development. Extended analysis shows that within cities, digital knowledge flowing from central districts to peripheral districts significantly enhances intra-city coordinated development. Moreover, the positive effect of digital knowledge flows on regional coordinated development is sustained in the long run.

This study carries significant policy implications. The findings suggest that the government should strategically and selectively promote cross-regional flows of digital knowledge by optimizing regional digital industry structures and establishing flexible demand feedback mechanisms to maximize long-term positive effects. Furthermore, this study contributes to relative literature with its innovative approach and academic value. It is the first to systematically examine the impact of digital knowledge flows on regional coordinated development. By analyzing the dual compatibility within the context of regional gaps in factor endowments, this study delves deeply into the mechanisms through which digital knowledge flows influence regional coordinated development. Additionally, utilizing a large language model to precisely measure regional digital knowledge flows enriches the application of large language models in economics.

Keywords: digital knowledge flow; regional coordinated development; knowledge compatibility; demand compatibility; factor endowment gap

JEL Classification: O18 O33 R11

[责任编辑:崔志新]