

城市群视角下的产业共聚与产业空间治理： 机器学习算法的测度

陈 露， 刘修岩， 叶信岳， 胡汉辉

[摘要] 城市群是未来中国产业发展的主要空间载体与重要地理单元。本文在机器学习 Wasserstein 距离算法思想的基础上,运用矩阵扩张 Sinkhorn 算法和熵正则化约束法改进 Wasserstein 距离算法及其求解方式,通过假设检验与蒙特卡洛模拟构建产业共聚指数,测算中国城市群的产业共聚水平。测度结果发现,中国城市群同二位数行业内产业间的共聚水平大于跨二位数行业产业间共聚水平;经济相对落后的城市群表现出明显的产业共聚特征,同时技术密集型产业在上述城市群内部偏向共聚;2007—2013 年间中国城市群产业空间共聚水平总体呈下降态势,不同城市群的演化模式存在差异。本文还利用产业共聚指数的方向性,分析了不同城市群产业空间核心产业的差异,绘制出城市群产业空间结构。对产业共聚影响因素的实证研究发现,投入产出关联、技术关联与规模差距对产业间共聚的影响显著为正,行业内部竞争水平从共聚和被共聚两个方向抑制产业间共聚趋势;城市群层面诸如政府规模等区域特征均与产业间共聚水平负相关。据此,本文提出了实施城市群产业空间治理的相应政策建议。

[关键词] 产业共聚; 产业空间治理; 机器学习算法; 城市群

[中图分类号]J424 **[文献标识码]**A **[文章编号]**1006-480X(2020)05-0099-19

DOI:10.19581/j.cnki.ciejournal.2020.05.017

一、问题提出

2019 年 8 月,中央财经委员会第五次会议指出:“中国经济发展的空间结构正在发生深刻变化,中心城市和城市群正在成为承载发展要素的主要空间形式”。会议明确提出,要打好产业基础高级化、产业链现代化攻坚战,完善城市群空间治理,形成优势互补、高质量发展的区域经济布局。当前由多个产业共聚形成的产业集群综合体逐渐成为未来产业空间发展的重要形式,新时代下产业空间新形态正在不断突破传统行政边界,逐步形成以核心城市为枢纽、多城市产业协同发展的圈层

[收稿日期] 2019-09-05

[基金项目] 国家自然科学基金面上项目“以‘一带一路’创新牵引中国区域产业转移、转型并高质量发展的机理、路径与政策研究”(批准号 71873030);国家自然科学基金面上项目“新型城镇化约束与引导下的产业转移的模式、路径与组织研究”(批准号 71473037)。

[作者简介] 陈露,东南大学经济管理学院博士研究生,新泽西理工学院吴鹰计算机学院联合培养博士研究生;刘修岩,东南大学经济管理学院教授,博士生导师,经济学博士;叶信岳,新泽西理工学院吴鹰计算机学院副教授,博士生导师,地理学博士;胡汉辉,东南大学经济管理学院教授,博士生导师。通讯作者:刘修岩,电子邮箱:qyjj2010@163.com。感谢哈佛大学中国数据实验室鲍曙明教授、新泽西理工学院袁军博士、南京财经大学李松林博士和苏州科技大学陈志辉博士的建议与帮助,感谢匿名评审专家和编辑部的宝贵意见,当然文责自负。

化产业空间新格局。打造世界级城市群已上升为中长期国家战略,城市群是未来产业发展的主要空间载体与重要地理单元,城市群空间下的产业结构转型升级与产业高质量发展将是推动区域经济发展质量变革、效率变革、动力变革的重要抓手。城市群产业空间治理的难点在于,区域层面既要“谋全局”也要“谋一域”,产业层面既要“顾整体”也要“抓重点”。不少地方在产业体系调整和重构中出现了不同程度的产业工业区“大杂烩”与产业“孤岛”并存的问题,而城市群内部政府间也存在产业政策不协调以及区域低效、过度竞争等乱象。准确判断产业体系内不同产业的空间联结与互动关系,是解决上述问题、提高城市群产业空间治理水平的重要前提和基础性工作,科学合理地测度城市群产业共聚水平进而深入发掘产业空间分布客观规律,对于从城市群之间以及城市群内部把握产业体系空间重构调整方向、制定布局优化策略有着十分重要的理论价值与现实意义。

产业空间分布最显著的特征是集聚(Krugman, 1991),这一社会经济现象长期吸引着学者们的关注。企业为什么要在地理上临近?劳动力“蓄水池”、毗邻市场及知识溢出所带来的正外部性是三大核心驱动力。单个企业选择在地理位置上靠近某些特定企业,一定是希望可以从这一过程中获益(Ellison et al., 2010),而这些特定的企业可能来自具有相似产业性质或者与其有业务往来的产业,也可能来自看似毫无关联或者从无业务往来的产业。上述通过地理上临近多样化产业而产生的知识溢出与规模效应被称为雅各布(Jacobs)外部性,而这种多类产业的空间集聚被定义为产业集群(Cluster)现象(Porter, 1990)。特定的多个空间关联产业更愿意在地理空间上临近,产业空间共聚体或集群综合体对产业政策的制定与企业的区位选择十分重要(Porter, 1998)。值得注意的是,产业空间共聚(Coagglomeration)有别于产业空间集聚(Agglomeration),本质区别在于:产业空间共聚强调跨产业空间分布的依赖、联结与互动关系,而产业空间集聚侧重产业总体或者单个产业的空间分布形态。产业共聚能够更好地描绘某一区域内产业体系的空间结构关系,从这个角度看,产业空间共聚关系的研究是产业集群与产业结构探讨的核心内容。

基于上述理论与现实背景,本文试图解决两个方面的问题:一是如何科学准确地测度区域产业间的共聚关系,以及如何通过产业间共聚的方向性进一步刻画产业空间结构;二是从产业共聚的角度出发,政府可以从哪些方面入手实施产业空间治理?针对上述问题的解答,本文的边际贡献主要来自三个方面:一是利用企业地理微观数据,在机器学习 Wasserstein 距离算法思想的基础上,运用矩阵扩张 Sinkhorn 算法和熵正则化项约束法改进 Wasserstein 距离算法及其求解方式,通过假设检验与蒙特卡洛模拟获取反事实样本集进而构建产业共聚指数,在城市群尺度下测度产业间的共聚水平,该指数对进一步拓展产业经济学和区域经济学等相关学科的研究都有着十分重要的意义;二是基于产业共聚测度指数刻画城市群产业空间,一方面对产业空间的共聚水平、时间演化趋势和区域差异进行比较分析,另一方面运用该指数的方向性进行城市群产业空间核心产业与空间网络分析;三是针对不同城市群下产业共聚的影响因素做跨区域跨产业的实证检验,特别是从方向性特征的角度验证了投入产出关联与技术关联对共聚指数的影响,同时弥补了关于跨区域产业共聚实证研究的不足,为促进城市群区域产业间协同发展、产业结构优化升级等产业空间治理政策的实施提供决策依据。

二、文献综述

国外学者对于产业空间共聚的实证研究基本上沿着两个方向发展:一是在划定的地理范围和行业层次上不断创新和改进产业共聚的测度指标;二是基于不同指标对影响产业空间共聚的微观机制进行检验。

Delgado et al.(2016)将目前主流的产业共聚测度指标分为基于离散空间单元和连续空间单元指标两大类。第一类指标的代表是基于区位选择模型理论度量的EG指数(Ellison and Glaeser, 1997),该类指标存在的主要问题是,任何空间尺度共聚的产业均会导致所有共聚变量之间的虚假相关性,所选择的加总水平越高,问题就越严重,这一问题被定量地理学家称为可更改的地理单元问题(MAUP)。第二类指标以DO指数为代表。Duranton and Overman(2005)基于微观地理距离、高斯线性核密度函数与反事实样本构建产业集聚指数,并在此基础上拓展性地提出了产业间共聚指数(Duranton and Overman, 2008)。与EG指数相比,DO指数能有效克服MAUP问题,但是该类共聚指数的不足主要有:点集之间的规模差距导致的测量误差问题、距离标准的客观性不足、产业共聚测度时构建的DO指数置信区间存在统计意义问题、计算复杂度过高、不同空间尺度的DO指数不可比等。

上述两类指标所具有的共同缺陷是基于单一产业集聚测度思想向产业间共聚测度方面延伸,而不是针对产业共聚提出的测度指标,使得这两类指标均没有考虑产业间共聚的方向性特征。产业共聚与产业集聚的重要差别是产业共聚具有方向性,这一特性可以与生态学中群落的相互关系类比。生态学家指出,两个群落的互动如果是为了某种正向关系而彼此依赖,则被称为群落间的正相互关系,在群落正相互关系中有偏利共生(Commensalism)、原始共生(Protocooperation)和互惠共生(Mutualism)的区别。其中,偏利共生是指两个群落接触时一方有利,一方无影响,分开时一方有害,一方无影响;原始共生是指接触时彼此有利,分开时彼此无影响;互惠共生是指接触时彼此有利,分开时对彼此都有害(Odum, 2004),这种群落之间的共生方向性在总结产业群体行为特征时也有一定的适用性。在传统产业共聚研究中,由上述EG指数或者DO指数所做的产业共聚测度都认为两个产业是共同集聚的(陈国亮和陈建军, 2012),或者说是协同集聚的(陈曦等, 2018),这显然不符合产业存在单向共聚的实际情况。

鉴于上述两类指标存在的主要问题,Billings and Johnson(2016)提出了基于Wasserstein距离算法思想的产业共聚测度方法,这一指数的提出是近年来产业共聚研究的重要突破,在归类上属于第二类连续空间单元指标,但是算法理念与上述两类现存方法完全不同,因此,可以认为这是一个全新的产业共聚测度指标。不过,该方法被引入产业共聚测度领域时使用的是较为基础的线性规划算法,计算结果稳定性和精确性都存在显著问题。彼时机器学习这一研究领域尚未得到广泛重视。2017年康奈尔大学的Martin团队将Wasserstein距离引入机器学习领域的GAN训练(Martin et al., 2017),创立了轰动机器学习领域的WGAN算法,该算法问世后被广泛应用于人脸识别、图片分析等机器学习训练领域(郑昌金等, 2017; Carlsson et al., 2018)。经过几年的发展,Wasserstein距离算法的精确度得到广泛认可,并且在计算理念与算法技巧上取得了突飞猛进的发展,日趋成熟。本文在Billings and Johnson(2016)算法思想的基础上,借鉴计算机学科在机器学习领域对该方法的改进策略,提高计算精度并降低计算复杂度。

虽然基于机器学习算法的产业共聚指数可以在任意空间尺度下进行产业共聚指数测度,即Delgado et al.(2016)提出的“对任意空间分类变化的无偏性特征”,但是在实际测度过程中仍然需要选择一定的空间尺度来测量。在应用中国的企业地理信息测度产业空间分布的过程中,本文认为在城市群这一空间尺度下测度是较为合适的。过往学者在研究产业空间集聚或共聚问题时,空间尺度选择多为国家、省级和地级市层面(陈国亮和陈建军, 2012; 贺灿飞和肖晓俊, 2012; 袁海红等, 2014; 邵朝对等, 2018; 陈曦等, 2018),鲜见针对城市群层面的测度。在城市群建设被逐步提升至国家战略的过程中,应用产业共聚指数准确把脉城市群产业结构、厘清产业间互动关系和区域内部产

业合理分工协作具有重要的现实意义。值得一提的是,Billings and Johnson(2016)的产业集聚指数正是对“丹佛—奥罗拉—莱克伍德”大都市功能区进行测度的,而美国大都市功能区与中国城市群的概念和规模较为接近,基于城市群尺度的产业集聚指数的测度结果更具国际可比性。

针对产业集聚研究,另一个值得关心的问题是,如何通过实证方法对产业集聚的决定因素和机制做经验分析。国外学者通过实证研究证实,在美国、加拿大与欧洲投入产出联系、知识溢出效应和劳动池效应对单一产业集聚做出了积极贡献(Duranton and Overman,2005,2008;Ellison et al.,2010;Behrens and Bougna,2015)。对于跨产业空间集聚的实证检验主要是从马歇尔外部性和要素禀赋理论角度展开,结果证明,马歇尔效应对跨产业集聚具有正向影响(Ellison et al.,2010),同时针对美国丹佛大都市圈的实证研究也验证了交通基础设施的可及性、消费者的可及性和知识溢出对区域内的产业集聚具有很强的积极作用(Billings and Johnson,2016)。但是这些研究只测算了一个区域或一个国家整体的产业集聚程度,无法对跨区域的产业集聚影响因素予以验证。同时,过往研究没有针对产业集聚指数方向性特性展开足够多的实证研究(Billings and Johnson,2016)。为此,本文分别测度了中国21个城市群2007年和2013年的产业集聚指数,针对指数方向性,验证了投入产出关联和技术关联对产业集聚发生概率的影响,兼顾讨论产业集聚跨城市群的影响因素,为城市群产业协同发展与区域政府间科学协同决策提供理论依据。

三、测度方法与指数特性

1. 基于 Wasserstein 距离的产业分布距离测算

两个产业在空间上分布的集聚程度可以转化为一个数学问题:两个产业在空间上的分布相似性问题。Wasserstein 距离是度量两个概率分布之间的距离度量函数,在概率论和数理统计中被广泛应用,近年来由于其优异的数学特性,在计算机领域也得到了进一步拓展,被应用于人脸识别、图片分析等人工智能领域。

Wasserstein 距离的一般数学表述为,对于定义在复平面 R^n 上的概率分布 $\mu(x)$ 和 $\nu(y)$, $\Gamma(\mu, \nu)$ 是 $R^n \times R^n$ 上的联合分布,并且同时满足 μ 和 ν 是该联合分布的边缘分布。对于 $\forall p \geq 1$, μ 和 ν 之间的 Wasserstein 距离被定义为:

$$W_p(\mu(x), \nu(y)) := \left(\inf_{\gamma(x,y) \in \Gamma(\mu, \nu)} \int_{R^n \times R^n} d(x,y)^p \gamma(x,y) dx dy \right)^{1/p} \quad (1)$$

在求解 Wasserstein 距离时,原始的计算方法是使用线性规划求解,求解的核心是找出联合分布 $\Gamma(\mu, \nu)$ 中的距离最小方案。对于两个连续分布,其 Wasserstein 距离采用随机抽取相同规模样本方式来计算;而对于两个离散分布,则对分布采取归一化后再计算。但在实际计算过程中,两个分布的样本规模对计算复杂度的影响较大,采用线性规划求解的结果存在精度不高与计算复杂度过高的双重问题。为了在提高求解精度的情况下缩短求解时间,Cuturi(2013)通过证明提出了使用矩阵扩张算法中的 Sinkhorn (1964) 不动点迭代和增加熵正则化约束项改进求解方式^①,这一方法在 Carlsson et al.(2018)的应用研究中得到了完善,并由 Luise et al.(2018)进一步证明了使用 Sinkhorn 算法后 Wasserstein 距离结果的稳健性。增加熵约束项后的 Wasserstein 距离为:

① 这里的快速求解是相对于原求解方式的超长时间而言,对于两个规模超过 1000 的空间离散分布来说,即便在 Cuturi 的算法下,单个 Wasserstein 距离的求解复杂度也极高。本文对 21 个城市群一共进行了超过 5 亿次的 Wasserstein 距离计算,采用 GPU 和云虚拟机多进程等手段加速计算,总计算时间超过 1500 小时。

$$W_p(\mu(x), \nu(y)) := \left(\inf_{\gamma(x,y) \in \Gamma(\mu, \nu)} \int_{R^n \times R^n} d(x,y)^p \gamma(x,y) dx dy - \alpha \Omega(\gamma) \right)^{1/p} \quad (2)$$

其中, $\Omega(\gamma)$ 被定义为联合分布 $\Gamma(\mu, \nu)$ 的熵约束:

$$\Omega(\gamma) = - \int_{R^n \times R^n} \gamma(x,y) \ln \gamma(x,y) dx dy \quad (3)$$

通过熵约束项 $\Omega(\gamma)$ 让式(2)为严格凸函数, 因此, 式(2)必存在唯一的最优解。

进一步定义核密度(Kernel Density)函数: $\kappa(x,y) = e^{-td(x,y)^p}$, 对所有 $t > 0$ 的情况, 该函数是正定的, 将其代入式(2)中, 可得:

$$W_p(\mu(x), \nu(y))_k := \left(\inf_{\gamma(x,y) \in \Gamma(\mu, \nu)} \int_{R^n \times R^n} \gamma(x,y) [\alpha \ln(\gamma(x,y)) - \ln(\kappa(x,y))] t dx dy \right)^{1/p} \quad (4)$$

式(4)即为基于 Sinkhorn 算法的熵正则化 Wasserstein 距离计算公式, 在实际求解过程中一般令 $t=1/\alpha$, 本文中 α 的经验值为 1。具体来说, 在研究产业空间共聚的问题上, p 取值为 2。二维空间下, $W_{j,k}$ 的实际含义就是在空间上将产业 j 的空间分布改造为产业 k 的空间分布的最小成本。由于两个产业 j 和 k 的规模往往是不同的, 参照 Billings and Johnson(2016)的处理方式, 需要将两个产业的分布分别通过 $1/N_j$ 和 $1/N_k$ 加权进行归一化, N 代表了各个产业内的企业数量^①。

2. 依据统计意义的反事实样本集的产业共聚指数构建

为了判断两个产业在空间上是否存在统计意义上的共聚, 需要进一步通过反事实样本构建产业 j 与产业 k 之间的共聚指数。这里提出构建产业共聚指数的原假设 H_0 : 在产业 j 空间分布既定的情况下, 产业 j 和产业 k 之间不存在空间分布相似性。

基于这一假设, 采用蒙特卡洛模拟方法进行原假设条件集合的构建, 即在 R^2 上随机选取对于产业 j 的模拟产业区位集 \tilde{k} , 与 Duranton and Overman(2005)、Billings and Johnson(2016)构建反事实样本的要求类似, 产业区位集 \tilde{k} 必须满足两个条件: 一是产业 \tilde{k} 的企业数量与产业 k 的数量应当相同; 二是选取随机产业区位集策略是从空间上所有已知的企业所在点的集合中抽取, 这一策略的主要理由是为了进一步控制随机选取的产业区位点集受到不可开发土地以及其他不可观测因素的影响。因此, 产业随机模拟组 \tilde{k} 将通过从空间内所有产业的企业地点信息中抽取 N_k 个区位点产生。将产业 j 与模拟产业 \tilde{k} 的 Wasserstein 距离记为 $W_{j,\tilde{k}}$, 这一随机抽取过程将重复 1000 次, 进而得到 1000 个模拟产业对距离 $W_{j,\tilde{k}}$, 对于产业 j 来说, 其选址与随机生成的模拟产业区位集 \tilde{k} 应是无关的, 因此, 模拟 1000 次计算的 $W_{j,\tilde{k}}$ 距离在统计意义构成了原假设条件集。将产业 j 与实际产业 k 的 Wasserstein 的距离 $W_{j,k}$ 与这 1000 次反事实结果作比较, 把模拟结果中违背原假设 ($W_{j,\tilde{k}}$ 小于 $W_{j,k}$) 的发生次数记为 n , 进而定义产业共聚指数 $Coagg_{j,k} = n/1000$ 。例如, 如果 1000 次模拟中, 有 950 次是 $W_{j,\tilde{k}}$ 小于 $W_{j,k}$, 则可以理解为在 95% 的置信度下拒绝原假设, 接受备择假设 H_1 : 在产业 j 空间分布既定的情况下, 产业 j 和产业 k 之间存在显著的空间分布相似性, 这种空间相似性是由产业 j 向产业 k 的方向共聚产生的。此时, 共聚指数 $Coagg_{j,k}$ 被定义为 0.95, 即蒙特卡洛模拟时违背原假设的发生频率, 通过该频率来近似代表产业 j 向产业 k 的方向共聚发生的可能概率。

① 关于产业 Wasserstein 距离的进一步解释参见《中国工业经济》网站 (<http://www.ciejournal.org>) 附件。

对于产业 j 和产业 k 这一对产业来说, $W_{j,k}$ 和 $W_{k,j}$ 是相等的, $Coagg_{j,k}$ 与 $Coagg_{k,j}$ 却是不同的, 前者的随机产业集是以产业 k 的规模 N_k 为基准, 而后的随机产业集的规模基准是 j 行业规模 N_j 。因此, 在进行某一空间下 m 个产业间的两两产业共聚指数计算时, 需要计算 $m \times (m-1)$ 个结果, 每一对产业对会有两个共聚指数, 每个产业共聚方向上都会赋予一个共聚指数。

该指标主要优点: 一是可以跨越行政区域来计算, 克服了 MAUP 的问题; 二是算法理念上沿用 DO 指数的蒙特卡洛反事实模拟法, 放弃了基于每个距离单位构建置信区间计算指数的主观做法, 借助统计学假设检验方法赋予客观性与统计意义; 三是巧妙利用反事实样本赋予产业共聚指数矢量特性^①; 四是由于采用企业微观地理信息, 该产业共聚指数可以在任意空间尺度上使用。

四、基于产业共聚指数的城市群产业空间刻画

Hidalgo et al.(2007) 创造性地将基于产品邻近度构建的区域产品网络定义为“产品空间”(Product Space)。借鉴其思路, 本文将基于产业共聚指数所构建的产业间地理共聚关系网络定义为“产业空间”(Industry Space)。本文一方面从城市群的产业空间角度对产业共聚的总体水平和城市群间的产业共聚分布与演化差异进行分析、比较, 另一方面利用产业共聚的方向特性分析产业空间网络中的核心产业, 通过产业间互动关系准确刻画城市群产业空间结构。

1. 数据来源

计算产业共聚指数的数据基础是微观企业层面的空间地理信息, 这是产业在空间上的最小组成单元。本文使用的数据来自 2007—2013 年中国工业企业数据库, 该数据库样本范围为全部国有工业企业以及规模以上的非国有工业企业^②, 利用企业的基本信息从百度 API 平台中进行经纬度解码来获取企业微观地理信息^③。在将企业进行城市群分类的过程中, 参照已有研究使用中国工业企业数据库的做法, 剔除总资产、工业总产值和固定资产净值缺失、0 值和负值的样本以及剔除不符合会计准则的样本 (Brandt et al., 2012)。

中国城市群个数和具体划分范围在学术界尚未形成共识, 在综合考虑企业微观数据的观测年份、数据的可得性与测度的经济学意义的基础上, 本文在《国家新型城镇化规划(2014—2020 年)》的 20 个城市群的方案基础上, 结合已有研究对城市群处理方法的经验 (黄金川和陈守强, 2015; 赵娜等, 2017), 剔除了天山北坡城市群, 将长江中游城市群仍分别视为武汉城市群、长株潭城市群和环鄱阳湖城市群, 最终确定了 21 个城市群作为研究对象^④。

2. 测算结果与分析

(1) 中国城市群产业共聚总体水平分析。本文使用 2013 年数据, 测度了 21 个城市群空间范围内 3385 个三位数产业间的空间共聚关系^⑤。本文将不同二位数行业内的三位数产业间共聚定义为跨二位数产业共聚, 将同二位数产业下的三位数产业间的共聚定义为同二位数产业共聚。2013 年

① 关于矢量性的详细解释与说明参见《中国工业经济》网站 (<http://www.ciejjournal.org>) 附件。

② 感谢匿名评审专家提出的 2013 年中国工业企业数据库规模以上的标准改变带来的样本代表性问题, 与 2007 年测度的结果对比发现, 这两年的共聚水平结果基本一致且稳定。

③ 在这一过程中, 本文的匹配策略是按照“企业名称—企业详细地址—企业所在县政府所在地”的顺序匹配, 即优先通过企业名称来匹配, 若无法查询, 则将企业的详细地址进行匹配, 两者若都无法匹配, 则用企业所在县政府所在地经纬度替代。

④ 城市群的名单及划分情况参见《中国工业经济》网站 (<http://www.ciejjournal.org>) 附件。

⑤ 由于统计的全面性差异及城市群产业分布的差异, 不同城市群的三位数产业数量并不一致。

一共测度了 277846 对三位数产业间的共聚指数^①, 这其中包含 269846 对跨二位数产业对与 8000 对同二位数产业对。正如上文提到的, 基于 Wasserstein 方法计算的产业共聚指数存在方向性, 因此, 实际计算的产业共聚指数共有 555692 个结果, 其中最小值为 0, 出现频数为 15869 (占 2.86%), 最大值为 1, 出现频数为 13066 (占 2.35%), 在统计意义上显著共聚 ($Coagg_{j,k} \geq 0.95$) 的产业共聚指数共 60474 个 (占 10.88%), 其中有 10763 对产业双向共聚, 占全部显著共聚指数的比例为 35.60%。在跨二位数产业共聚的 539692 个结果中, 统计意义上显著共聚的有 57670 个 (占 10.69%), 这其中有 10069 对产业双向共聚, 占全部显著共聚指数的比例为 34.92%。同二位数产业共聚的 16000 个结果中, 统计意义上显著共聚的有 2804 个 (占 17.53%), 其中有 694 对产业双向共聚, 占全部显著共聚指数的 49.50%。基于以上数据, 总体上可以判断, 中国城市群层面同二位数产业共聚水平显著高于全部产业共聚水平与跨二位数产业共聚水平。

为了观察产业空间共聚随时间演变的趋势, 本文使用 2007 年的数据测度 21 个城市群的 3399 个三位数产业间空间共聚关系。这里共测度了 280407 对三位数产业间的共聚指数, 包含 272264 对跨二位数产业对与 8143 对同二位数产业对。全部的 560814 个结果中, 最小值为 0 的频数为 21537 个 (占 3.84%), 最大值为 1 的频数是 16179 (占 2.88%), 统计意义上显著共聚的个数为 66122 个 (占 12.10%), 其中双向共聚个数为 25442 个 (占 38.38%)。跨二位数产业共聚的 544528 个结果中, 统计意义上显著共聚的有 63190 个 (占 11.60%), 这其中有 11978 对产业双向共聚, 占全部显著共聚指数的比例为 37.91%。同二位数产业共聚的 16286 个结果中, 统计意义上显著共聚的有 2932 个 (占 18.00%), 这其中有 743 对产业双向共聚, 占全部显著共聚指数的 50.68%。

为了更直观地展现全部测度结果分布, 本文通过核密度函数分别构建了 2007 年和 2013 年全部产业共聚指数、同二位数产业共聚指数和跨二位数产业共聚结果的密度分布 (见图 1), 该密度分布横坐标越接近 1, 表明产业间在空间上是趋于共聚的, 越接近 0 则表明产业在空间上是趋于非共聚的 (或者说是分散的)。测度结果的总体分布上, 2007 年和 2013 年两年的产业共聚指数 $Coagg_{j,k}$ 在 0—0.1 和 0.9—1.0 附近的分布密度较为集中, 并且产业空间分散趋势强于产业空间集聚趋势, 这是符合常理的, 因为大部分产业之间不存在空间关联, 社会经济的分工细化使得总体细分产业在空间上趋于分散, 但是在局部细分产业中合作共生的趋势正在增强, 因此整体上呈现出左高右低的“马鞍型”分布。这种效应尤其体现在跨二位数产业共聚指数结果中, 其密度分布与总体产业的密度分布几乎是一致的, 但还是存在微小差别, 主要体现在更强的分散效应和更弱的共聚效应, 这些效应表明中国城市群下的跨二位数产业的空间共聚效应偏弱。相较之下, 同二位数产业内的产业共聚效应明显强于总体与跨二位数产业的共聚效应, 三位数行业间的分散效应远远小于其共聚效应, 密度分布呈现出左低右高的“马鞍型”分布, 这说明中国城市群下的同二位数产业内的细分行业合作是普遍的, 且呈现出非常明显的聚集趋势。从时间演变看, 2007—2013 年中国城市群层面下的产业空间共聚水平总体呈下降态势, 在显著共聚和双向共聚的数量与比例上均表现出明显的减少。从密度分布看, 2013 年较 2007 年产业间共聚水平的两极分化态势加强, 全部产业间和跨二位数行业间的共聚指数在 0—0.1 附近与 0.9—1.0 附近的密度变大, 中部密度降低, 这一两极分化趋势在同二位数行业间表现得尤为明显。

从产业空间共聚密度分布与演变趋势看, 针对城市群产业空间有效合理共聚的产业政策仍有较大的施展空间。一是城市群跨二位数产业显著共聚的数量较多, 说明应当破除传统二位数产业行业划分产生的隐形壁垒, 产业政策导向应当打破传统产业分类的思维定势, 从整体以及结构的角度

① 具体结果详见《中国工业经济》网站 (<http://www.ciejournal.org>) 附件。

把握产业间联系的客观规律,在城市群产业链、产业集群共同体的规划、升级及高质量发展过程中,应当将跨二位数行业的产业关联一并纳入考察;二是城市群同二位数产业间的共聚虽然呈现出微弱的下降态势,但是显著共聚维持在一定的比例,这说明过去一段时间的产业政策总体导向是正确的,在同二位数产业内部应当进一步优化其产业分类和产业结构,促进同二位数产业内部进一步扩大规模与创新升级;三是产业共聚指数密度分布介于 0.50—0.95 区间的产业仍占有相当的比例,并且密度分布显示,随着时间变化呈现出两极分化的态势,这部分产业间存在空间共聚的趋势,但是仍未显现出明显的产业共聚迹象,城市群决策群体应当通过营造良好的合作环境以及积极的产业政策引导使得相应产业在城市群内部流动以进一步促进产业间空间有效合理分布^①,降低交易成本,实现规模经济与协同创新以达到产业协同合作的效益最大化。

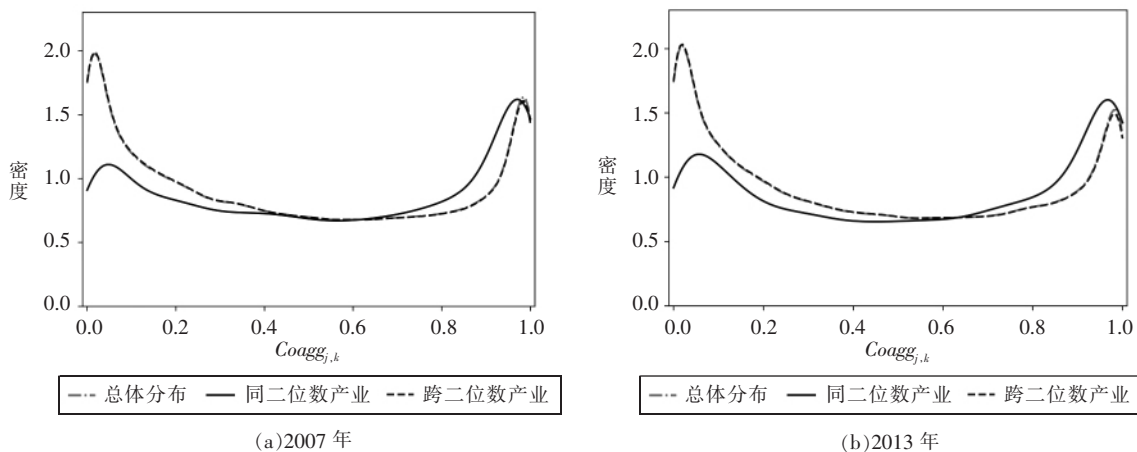


图1 2007年和2013年中国城市群产业共聚指数密度分布

注:2007年和2013年总体与跨二位数产业的共聚指数密度分布都几乎一致,图中二者的分布曲线高度重合。

(2)城市群产业共聚水平区域差异分析。通过对不同城市群的空间共聚指数的测度,可以进一步发掘中国产业空间分布的区域差异。本文对2007年和2013年21个城市群的产业共聚测度从全部产业、跨二位数产业和同二位数产业三个角度汇总(见表1)^②。总体看,2013年的城市群总体共聚指数均值与产业显著共聚数量低于2007年。其中,指数均值上升的城市群有京津冀城市群、辽中南城市群、长江三角洲城市群、长株潭城市群、黔中城市群、滇中城市群、关中—天水城市群;产业显著共聚数量降低最多的城市群是长江三角洲城市群。江淮城市群、环鄱阳湖城市群和北部湾城市群的时间演变态势非常特殊,其总体指数均值下降,产业共聚显著数量却在增多,表明该空间内部产业共聚两极分化趋势极为明显,即总体共聚态势减弱的同时局部产业共聚趋势加强。与之相对应的长江三角洲城市群却呈现出相反的演化趋势,其空间内部产业共聚总体态势增强,而局部产业共聚正在减弱。

2013年和2007年各个城市群的共聚指数均值最大的城市群都是晋中城市群,表明该区域内的产业在空间上呈现总体共聚的趋势,共聚指数均值最小的也都是长江三角洲城市群,表明该区域内的产业在空间上呈现总体分散的趋势。两年各城市群共聚指数方差稳定,介于0.3—0.4之间。从

① 本文同意匿名评审专家提出的观点:产业共聚作为一种多维因素共同决定的产业生态系统,究竟何种共聚程度比较合适是很难定论的。因此,在立场上,本文只提供产业共聚的现状描述,将产业共聚指数作为一种产业空间治理工具予以推荐。

② 跨二位数产业与同二位数产业内的共聚具体结果详见《中国工业经济》网站(<http://www.ciejjournal.org>)附件。

共聚指数显著数量和占比看,2013年产业共聚指数显著数量最多的是长江三角洲城市群(4562个,占13.12%),显著数量占比最大的是晋中城市群(3501个,占25.80%),而兰西城市群产业共聚显著指数不管从数量和占比上都是最少的(1005个,占6.48%)。2007年产业共聚指数显著数量最多的也是长江三角洲城市群(5847个),产业显著共聚数量占比最大的也是晋中城市群(占30.65%)。有所区别的是,2007年产业共聚显著指数数量最少的是宁夏沿黄城市群,产业显著共聚占比最低的是关中—天水城市群。在这些显著的产业共聚数量中,双向共聚是一个值得关注的重要特征,这表明产业空间共聚的共生关系。2013年产业双向显著共聚数量最多的是京津冀城市群,数量最少的是兰西城市群;双向共聚指数数量占产业共聚指数显著数量的比例最高的是黔中城市群(占46.09%),占比最低的是珠江三角洲城市群(占27.49%)。

值得注意的是,从总体的产业共聚指数的统计指标看,产业共聚指数显著数量最多的长江三角洲城市群共聚指数均值却最小,并且随着时间的变化,这一趋势在进一步加强。这表明,该城市群区域内的产业分工进一步细化,产业总体上趋于分散,同时局部产业共聚合作密切,这刻画了当前中国区域产业总体协作水平较低、局部精细化合作水平较高的产业结构特征。针对这一现象,各个城市群应当结合自身产业结构的特点与优势,在保持局部细分产业合作水平继续提高的同时,进一步扩大跨二位数产业间的协同合作。从2007年与2013年的对比结果看,黔中城市群与关中—天水城市群正在向这一趋势发展。

将2013年中国城市群三位数行业总体的空间共聚指数结果进一步分解为城市群跨二位数行

表1 2007年和2013年中国城市群三位数产业总体共聚水平

城市群名称	2007					2013				
	共聚指数均值	共聚指数显著数量	占比(%)	其中:双向显著数量	占显著产业对比例(%)	共聚指数均值	共聚指数显著数量	占比(%)	其中:双向显著数量	占显著产业对比例(%)
京津冀城市群	0.426	4546	13.35	2094	46.06	0.436	4438	12.90	1834	41.32
辽中南城市群	0.425	3598	10.46	1234	34.30	0.426	3426	10.40	1110	32.40
山东半岛城市群	0.397	3799	10.92	1382	36.38	0.389	3374	10.02	1128	33.43
长江三角洲城市群	0.340	5847	16.99	2128	36.39	0.380	4562	13.12	1398	30.64
珠江三角洲城市群	0.445	2740	8.50	734	26.79	0.438	2634	8.08	724	27.49
晋中城市群	0.678	4750	30.65	2640	55.58	0.669	3501	25.80	1536	43.87
呼包鄂榆城市群	0.500	1809	11.67	770	42.56	0.492	1802	9.81	670	37.18
哈长城市群	0.469	2458	8.46	1044	42.47	0.459	2306	8.22	844	36.60
江淮城市群	0.467	3068	10.08	1196	38.98	0.461	3166	10.52	1064	33.61
海峡西岸城市群	0.436	4731	14.36	2104	44.47	0.417	3755	12.47	1516	40.37
环鄱阳湖城市群	0.470	2371	8.06	784	33.07	0.465	2649	9.44	964	36.39
中原城市群	0.448	2628	8.53	834	31.74	0.437	2196	7.21	642	29.23
长江中游城市群	0.508	3638	11.04	1230	33.81	0.480	3035	9.42	922	30.38
长株潭城市群	0.469	4192	13.31	1442	34.40	0.480	4061	12.60	1434	35.31
北部湾城市群	0.488	2055	10.26	720	35.04	0.478	2100	10.34	838	39.90
成渝城市群	0.475	4537	13.47	1720	37.91	0.422	3856	11.09	1322	34.28
黔中城市群	0.532	2307	13.14	938	40.66	0.584	3137	16.59	1446	46.09
滇中城市群	0.558	2832	15.42	1014	35.81	0.577	2310	13.16	778	33.68
关中—天水城市群	0.461	1876	7.10	552	29.42	0.482	2118	8.02	620	29.27
兰西城市群	0.516	1225	7.42	376	30.69	0.510	1005	6.48	314	31.24
宁夏沿黄城市群	0.538	1115	10.82	506	45.38	0.517	1043	9.74	422	40.46
合计	0.464	66122	12.10	25442	38.48	0.468	60474	10.88	21526	35.60

业共聚与同二位数行业共聚。可以发现,2013年各个城市群的跨二位数行业共聚测度结果与总体的测度结果相比,均值与共聚指数的显著数量都明显下降,这一特征在各个城市群的表现趋势一致,只是程度有所不同,而这一部分的下降的原因是各个城市群内同二位数产业内的空间共聚都保持较高的水平。从同二位数行业共聚的指数测度结果中可以发现,各个城市群同二位数产业共聚指数均值最大值为0.708(晋中城市群),共聚指数均值最小值为0.477(中原城市群)。从共聚指数显著数量和占比看,同二位数产业共聚指数显著数量最多的为长江三角洲城市群的247个(占25.41%),数量最少的是宁夏沿黄城市群的47个(占11.99%);显著数量占比最高的是晋中城市群的159个(占32.45%),占比最低的是兰西城市群49个(占9.76%)。但是从双向共聚产业对占总体显著产业对的比重看,除京津冀城市群与宁夏沿黄城市群的占比较高外,其他城市群的同二位数产业双向共聚产业对均较少,而珠江三角洲城市群全部2634个显著共聚结果中,只有362对双向共聚产业,这说明珠三角的产业体系中存在若干非常强势的支柱产业(如计算机制造业),使得其他产业都产生向其单向共聚的趋势。这对珠三角城市群的产业政策制定者来说应当予以重视,特别是近些年珠三角地区实施“腾笼换鸟”的产业升级政策过程中应当结合产业结构中的支柱产业来做出有效决策、精准决策和科学决策。

为了更直观地展现不同城市群在总体产业共聚指数分布与同二位数产业共聚分布之间的差异性,这里绘制了2013年代表性城市群产业共聚指数的核密度分布(见图2)。依据总体产业共聚指数的核密度分布表现特征,把21个城市群归为三类:第一类左高右低“马鞍型”分布,这表明产业间的空间分散趋势大于共聚趋势,如图2(a)所示。这一类的城市群有京津冀城市群、辽中南城市群、山东半岛城市群、长江三角洲城市群、珠江三角洲城市群、海峡西岸城市群、环鄱阳湖城市群、中原城市群和成渝城市群。第二类为左右持平“平稳型”分布,这表明在这些区域内产业间的空间分散与共聚趋势相当,如图2(b)所示。这一类的城市群有呼包鄂榆城市群、哈长城市群、江淮城市群、长江中游城市群、长株潭城市群、北部湾城市群、关中—天水城市群和兰西城市群。第三类为左低右高“马鞍型分布”,如图2(c)所示。这类特殊的城市群分别是晋中城市群、黔中城市群、滇中城市群和宁夏沿黄城市群,在这个四个城市群中出现了显著的产业空间共聚趋势大于空间分散趋势的情况,尤以晋中城市群为甚。通过产业分布特征的分类比较,可以得出这样一个结论:经济总体水平较高、产业体系较为完备的城市群的产业分散趋势更强,而经济相对落后、城市群发展起步较晚的地区的产业共聚的趋势更强,这一现象背后可能包含着区域规模经济集聚与竞争效应分散力多重作用的结果,类似于一些学者发现的中国产业集聚在2004年前后呈现“先集聚后分散”的特征(文东伟和洗国明,2014),产业共聚可能也存在着“先共聚后分散”的周期特征。本文同时还绘制了21个城市群同二位数产业内共聚的核密度分布,如图2(d)所示。可以发现,21个城市群的同二位数产业间与跨二位数产业间表现出的总体趋势存在非常明显的差异,同二位数产业内的各个细分产业间在各个城市群均表现出明显的共聚效应强于分散效应的趋势。以上结果进一步说明中国当前产业结构区域多样性与差异性的事实,不同城市群的产业分工受到自然禀赋、要素与消费市场等多方面的影响,城市群建设者在产业空间布局上应当因地制宜;区域的产业结构受到产业分工、历史政策等多方面影响,因而产业结构调整与升级上应当精准发力,“对症下药”。

(3)城市群二位数产业内部共聚水平差异分析。不同产业内部的共聚联系程度在城市群之间也表现出了非常明显的差异,这里通过计算同二位数产业内不同三位数产业间的共聚指数均值,进而对不同城市群内二位数产业做横向比较^①,此处着重讨论纺织业、食品制造业、医药制造业和汽车制

^① 具体结果详见《中国工业经济》网站(<http://www.ciejournal.org>)附件。

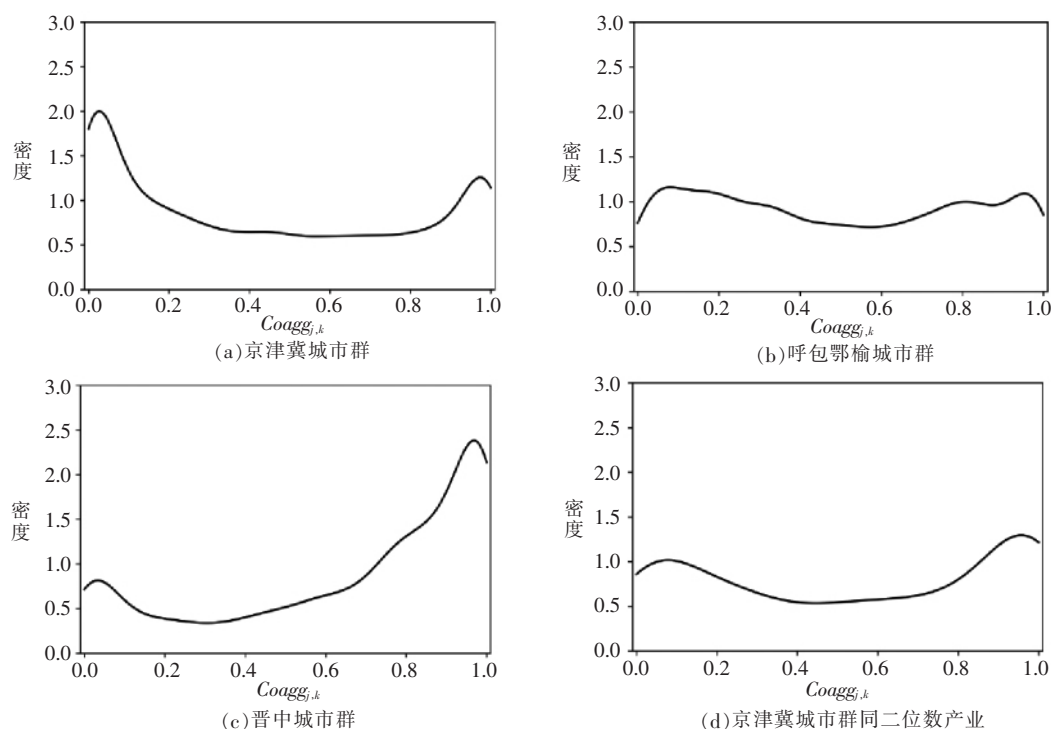


图2 2013年代表性城市群产业共聚指数密度分布

注:每个城市群密度分布参见《中国工业经济》网站(<http://www.ciejjournal.org>)附件。

制造业这四个代表性行业的城市群差异(见图3)。从行业的总体差异看,劳动力密集型产业(纺织业和食品制造业)内部共聚水平明显高于技术密集型产业(医药制造业和汽车制造业),这与中国劳动力密集型产业体系的完备是分不开的。中国纺织业与食品制造业的产业链与产业生态已经趋于成熟,产业在各个城市群的布局已经完成且已经形成相对稳定的产业生态,同时由于技术门槛较低,对劳动技能的要求不高,使得对劳动力市场的需求趋于一致,进而更有可能在地理空间上趋近,产业内部共聚水平较高是合理的;与之相比较,中国的技术密集型产业起步晚,发展比较失序,例如,较为典型的医药行业,其“小、散、乱”的发展特征使得产业生态不稳定,同时由于对高素质技术人才的需求较高,使得产业内部的分化较为明显,医药制造业内部的化学原料药、中成药与医疗器械产业之间有明显的行业壁垒,人才需求差异较大,难以形成统一的劳动力需求市场,加之产业布局受资本的影响,地方政府竞争性的工业园区招商引资使得高技术产业在空间上趋于分散。从城市群差异角度看,城市群劳动力密集型产业的共聚程度与地区经济水平关联性较低,技术密集型产业却在一些经济欠发达的城市群产生高度内部共聚,这一现象可能与经济欠发达地区在发展技术密集型产业时采用“集中力量办大事”的策略有关。当然,技术密集型产业间共聚能否显著提高彼此产业的创新能力和生产效率尚需验证,但是从马歇尔外部性的理论与“硅谷”这样成功的多产业共聚实践结果看,技术密集型产业间加强跨产业的协同合作应当是利大于弊的。

(4)城市群产业空间网络差异分析。产业空间网络可以划分出“核心—边缘”产业,产业空间核心产业为网络结构性主导产业,指的是在一定区域内被其他产业共聚最多的产业,这类产业不管从地理位置、劳动力池与资本市场等多方面均占据优势,这种优势一方面使得一些有业务关系或者生产关系的产业依附于该产业,在空间上向其聚集;另一方面由于产生了资本市场与劳动力市场的“虹吸效应”,使得一些与其有相似资本、劳动力与创新需求的产业也在其附近聚集。这样的产业从

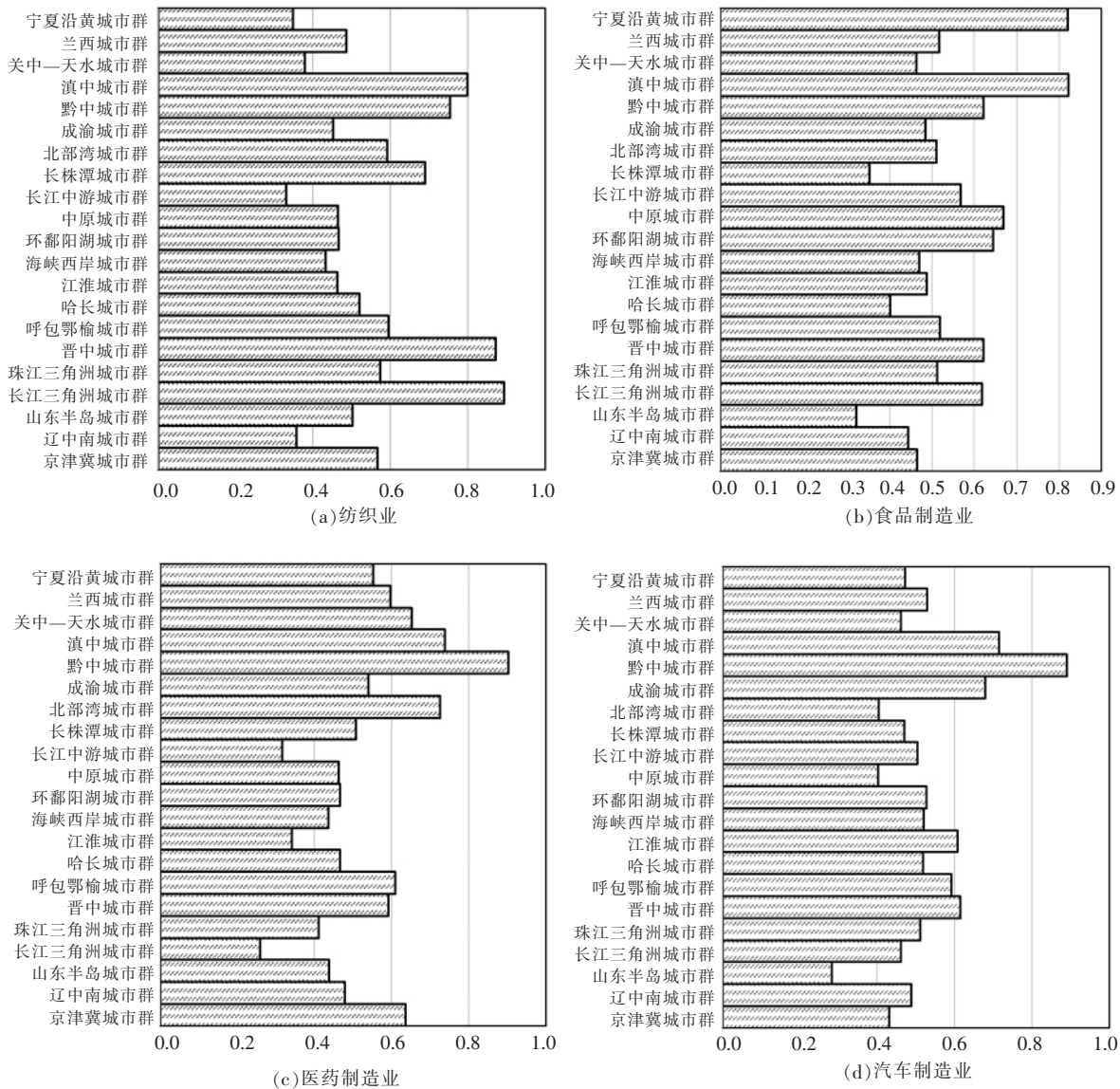


图3 2013年城市群部分二位数产业内部共聚水平比较

产业结构的角度看处于居中地位,与其他产业互动关系最为频繁,可谓“牵一发而动全身”,从区域政策制定的角度来说这些产业是最值得关注的产业。

本文将各个城市群下的显著产业共聚进行了汇总统计,计算各个三位数产业被共聚的频数,进一步加总得到所属二位数产业的频数,进而得到各个城市群前三位的产业空间核心产业^①。可以看出,中国城市群层面的网络核心产业主要有金属制品业,计算机、通信和其他电子设备制造业,通用设备制造业等行业。其中,金属制品业、通用设备制造业和专用设备制造业等行业成为区域性结构性支柱产业的主要原因是其产业链相对较长,生产工序相对复杂,与之关联的行业也更多,其跨产业联系更为密切。同时,各个城市群的网络核心产业存在明显的地区特点,以京津冀、长江三角洲与珠江三角洲这三大城市群为例(见表2),除了计算机、通信和其他电子设备制造业在这三个城

① 全部 21 个城市群网络空间核心产业结果详见《中国工业经济》网站(<http://www.ciejournal.org>)附件。

市群占据结构性主导行业前三位以外,每个城市群都有各自特色的网络核心产业,京津冀城市群的化学原料和化学制品制造业、长江三角洲城市群的纺织业与珠江三角洲城市群的电气机械和器材制造业都有鲜明的地域产业特色。通过产业空间核心产业的比较与分析,可以清楚地了解各个城市群的主导产业以及主导产业影响的潜在产业。

表 2 2013 年三大城市群前三位产业空间核心产业

城市群名称	第一位	第二位	第三位
京津冀城市群	金属制品业	计算机、通信和其他电子设备制造业	化学原料和化学制品制造业
长江三角洲城市群	纺织业	计算机、通信和其他电子设备制造业	通用设备制造业
珠江三角洲城市群	通用设备制造业	电气机械和器材制造业	计算机、通信和其他电子设备制造业

在产业空间的视角下,区域产业间的互动联系将由具有方向性特征的产业共聚指数构建。如图 4 所示,本文从产业共聚的视角绘制了长江三角洲城市群三大网络核心产业二位数产业,即纺织业(17),计算机、通信和其他电子设备制造业(39),通用设备制造业(34)的产业空间结构。从图中可以直观地看出,毛纺织及染整精加工业(172),电子器件制造业(397),文化、办公用机械制造业(347)在三大产业中的核心地位,与其他的三位数产业之间的共聚关系非常密切。从所属的二位数的角度可以看出各个二位数产业内部关联性与外部的互动关系。从内部关联性看,纺织业(17)的内部关联性最强,劳动密集型产业的产业特征使得其共享相同的劳动力池,并且也会享有相同的上下游产业;而关联性最弱的产业是通用设备制造业(34),通用设备制造业的细分子产业有明显细分产业差别,内部的产业关联度弱,但是与其他产业的互动却是很强的,这也是其产业特征所引致的;从二位数产业间的互动关系看,通用设备制造业产业内部联系弱,却与其他两个产业互动关系最强,如物料搬运设备制造业(343)与从属的二位数产业均没有关联,却与计算机、通信和其他电子设备制造业(39)的内部子产业互动非常密切。

五、城市群产业共聚影响因素分析

产业间地理临近关系的产生、巩固与消失是一个动态发展的过程,这一过程受到多种因素的影响。关于产业共聚的已有实证研究均从马歇尔外部性和雅各布外部性的理论基础上展开验证(Duranton and Overman, 2005, 2008; Ellison et al., 2010; Behrens and Bougna, 2015; Billings and Jonhson, 2016),主要局限是仅仅对理论机制做检验,未考虑区域的差异影响,同时未充分考虑产业共聚特有的方向性,即便是在最初提出这一共聚指数的研究中, Billings and Jonhson(2016)囿于数据限制,也只是讨论了具备方向性的投入产出指标。本文在前人研究的基础上,针对产业共聚的方向性特性,重点探讨中国城市群层面上产业共聚所受到的投入产出与技术关联影响,同时兼顾讨论产业间规模差距、竞争效应以及区域因素的作用。

1. 计量模型的设定和变量说明

本文将产业共聚 $Coagg_{j,k}$ 指数作为被解释变量,针对产业共聚指数的矢量性特征,重点构建同样具备矢量特征的投入产出关联与技术关联指标,产业规模差距、行业竞争和区域政策差异角度选取解释变量,研究产业共聚的驱动因素及地区性差异,建立如下计量模型:

$$Coagg_{j,k,i} = \alpha_0 + \alpha_1 Inputoutput_{jki} + \alpha_2 TechCorrelation_{jk} + \alpha_3 ScaleDissimilarity_{jki} + \alpha_4 Competition_{Level_{ji}} + \alpha_5 Competition_{Level_{ki}} + \alpha_6 Economy_i + \alpha_7 Government_i + \alpha_8 Transport_i + \alpha_9 Open_i + \varepsilon_{j,k} + \xi_{jki} \quad (5)$$

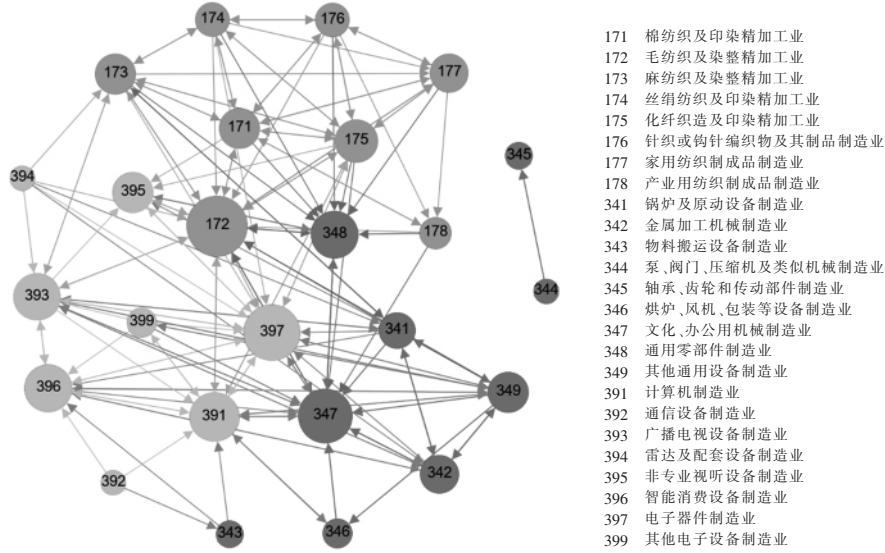


图 4 2013 年基于产业共聚指数的长江三角洲城市群局部产业空间结构

注:图中节点的大小代表其被其他产业依附的程度,带有箭头的连线代表了其共聚的方向。2013 年中国工业企业数据库中长江三角洲城市群没有企业统计归类为电子元件及电子专用材料制造业(398)。

其中, j 和 k 为行业代码, i 代表城市群, $\varepsilon_{j,k}$ 代表一系列固定效应, 具体包括三位数行业 j 的固定效应、三位数行业 k 的固定效应, 以及二位数行业 j 的固定效应与二位数行业 k 的固定效应的交乘项。引入这些固定效应, 可以控制三位数行业内以及二位数行业内影响产业共聚的不可观测因素, 同时将被解释变量的变异来源限制在每个二位数行业内进行比较。 ξ_{jki} 为随机扰动项。本文采用滞后变量消除内生性带来的影响, 除了投入产出变量为 2012 年的数据, 其他变量均选取 2010 年的数据。以上指标来自中国工业企业数据库(2010)、历年《中国城市统计年鉴》、《中国投入产出表》(2012)和夜间灯光遥感数据。其他变量说明如下:

$Input-output_{j,k}$ 表示产业间投入产出关系。产业间投入产出关系是衡量产业关联的重要指标之一, 这一指标表明了产业间的上下游关系, 具有方向性。验证产业间投入产出关系对产业空间共聚的影响, 可以进一步解释前文所定性分析的产业空间单向共聚的成因。参照 Billings and Johnson (2016) 和 Ellison et al.(2010) 的做法, 使用产业间投入产出表对产业间投入产出关系进行判别, $Input-output_{j,k}$ 表示产业 j 的产出中由产业 k 投入贡献的比例。本文使用 2012 年中国 139 行业投入产出表, 通过行业匹配构建该指标。该指数说明了产业 j 对产业 k 的投入产出关系。

$TechCorrelation_{j,k}$ 表示产业间技术关联。本文的技术关联是衡量产业间技术依赖的指标, 具有方向性。产业间的技术关联内在动力来自于知识溢出, 是指临近的企业能够从彼此的技术外部性中获益。技术外部性不仅表现在研发层面, 还可能通过教育培训体现。技术外部性程度与产业间技术关联间的方向性通常不易测量, 现有研究大多以专利数量或引用数近似代替, 但是这种方法只涵盖了有形的技术层面, 而且数据较难获得。Hidalgo et al.(2007)使用产品临近度指标度量“产品空间”, Guo and He(2017)、贺灿飞和胡绪千(2019)对该指标进行拓展和引申, 将其用于度量产业间的技术关联, 可以理解为“产品临近”是由于产业间的技术关联所致。本文借鉴上述研究的思路和方法, 进一步赋予该指标方向性, 用来测度行业间的技术依赖水平。本文定义 $TechCorrelation_{j,k}$ 指标为产业 j

对产业 k 的技术依赖,具体计算公式如下:

$$TechCorrelation_{j,k} = P(RCA_{cj} > 1 | RCA_{ck} > 1)$$

$$RCA_{cj} = Employment_{cj} / \sum_j Employment_{cj} / (\sum_c Employment_{cj} / \sum_{c,j} Employment_{cj}) \quad (6)$$

其中, RCA_{cj} 为 c 国家(或城市)产业 j 的显性比较优势。该技术关联指标表示为产业 k 具备比较优势的条件下产业 j 具备比较优势的条件概率,反映了产业 j 对产业 k 技术依赖的可能性。

Scale Dissimilarity 表示规模差距,本文使用工业总产值代表产业规模,使用产业间规模差的绝对值与规模总和的比值衡量产业间规模差距。*Competition Level* 表示产业内的竞争水平,为了考察竞争效应对产业空间共聚的影响,本文分别引入 j 行业与 k 行业的企业数量作为行业内部竞争能力的指标。*Government* 表示政府规模,由政府主导的产业政策在产业集聚差异方面的影响和作用不容忽视,有学者研究表明政府规模与产业间的协同集聚水平有正向关系(陈曦等,2018),本文用政府财政支出衡量政府规模。*Economy* 表示经济水平,为避免与其他区域指标的共线性,使用城市群夜间灯光数据来度量区域间的经济发展水平,并对灯光数据取对数处理。*Transport* 表示运输成本,选取城市群人均道路面积来度量运输成本的大小。*Open* 表示开放程度,用外商直接投资占 GDP 的比重考察城市群的对外开放程度。

2. 回归结果及分析

表3中报告了城市群产业共聚的影响结果,第(1)列报告了不控制行业固定效应的基准回归结果,第(2)列报告了不控制区域差异的基准回归结果,第(3)、(4)列报告了同时控制三位数行业固定效应与二位数固定效应的结果,需要指出的是,本文进行行业固定效应控制时同时控制了行业 j 与行业 k 的交叉固定效应。

从行业层面实证结果可以发现,投入产出关联对产业空间共聚有正向影响,说明两个产业间的投入产出关联越密切越有可能发生共聚,从指标的实际意义上可以得出产业越容易向对其投入比例高的相关产业地理临近的结论;产业间的技术依赖正向影响产业空间共聚的发生概率,产业 j 对产业 k 技术依赖性越强,产业 j 则越有可能向产业 k 空间临近,从而表现出共聚特征。

实证结果还验证了规模差距对产业共聚显著的正向影响,说明规模差距越大的两个产业越易于发生共聚,规模差距越小的两个产业越不易发生共聚。这是符合经济学常识的,即具备规模经济的产业越易于与规模不经济的产业共聚,而两个同时规模经济或同时规模不经济的产业是不会通过共聚的方式来实现规模经济的,这可以总结为规模经济的“学习效应”。从竞争效应的验证结果看,不管是共聚产业还是被共聚产业,其产业内部的竞争力都会对其共聚趋势产生负向影响,产业间的竞争主要集中于资源的竞争,这包括人力资源、生产资源、市场份额等一系列产业间关联资源,围绕这些资源的争夺越激烈,则越不利于产业间彼此的方向性共聚。

从区域层面实证结果看,政府的规模越大对越不利于产业的共聚,这与陈曦等(2018)的结论不一致。本文的解释是,政府的产业政策往往不具有稳定性,其空间的区位选择是随机的,空间分布上表现为分散化,因此,政府对产业政策的干预越强,产业在空间上越发分散。由于陈曦等(2018)采用的是 EG 指数的简化测度方法,且在地级市层面测度,由 MAUP 问题导致的测度结果不准确也可能是导致这一差异的原因。与上文的产业共聚指数在不同的城市群的密度分布模式的分析结果相对应,本文发现经济发展水平与产业共聚间存在显著的负相关关系。一方面,从产业生命周期的角度可以将其解释为,落后地区的产业都位于城镇地区,在地理空间上都集中在狭小的区域内,因此,从空间分布上表现为显著的共聚效应。另一方面,由于中国工业企业数据库只统计规模以上的企业,经济发展较为落后的城市群规模以上企业以国有企业为主,一般都聚集在城市群的特定经济功能

表 3 城市群产业共聚影响因素回归结果

变量	<i>Coagg_{j,k}</i>					
	(1) 基准回归	(2)	(3)	(4)	(5) 跨二位数行业	(6) 同二位数行业
<i>Input-output_{j,k}</i>	0.1925*** (0.0154)	0.1971*** (0.0154)	0.1224*** (0.0297)	0.1232*** (0.0298)	0.2276*** (0.0452)	0.0080 (0.0837)
<i>TechCorrelation_{j,k}</i>	0.0384*** (0.0042)	0.0497*** (0.0042)	0.0704*** (0.0073)	0.0722*** (0.0073)	0.0694*** (0.0074)	0.0587 (0.0613)
<i>ScaleDissimilarity</i>	0.0111*** (0.0018)	0.0136*** (0.0018)	0.0235*** (0.0019)	0.0278*** (0.0019)	0.0238*** (0.0019)	0.0246** (0.0119)
<i>Competition Level_j</i>	-0.0001*** (0.0000)	-0.0001*** (0.0000)	-0.0001*** (0.0000)	-0.0001*** (0.0000)	-0.0001*** (0.0000)	-0.0000 (0.0000)
<i>Competition Level_k</i>	-0.0000*** (0.0000)	-0.0001*** (0.0000)	-0.0001*** (0.0000)	-0.0001*** (0.0000)	-0.0001*** (0.0000)	-0.0000 (0.0000)
<i>Government</i>	-0.0424*** (0.0023)		-0.0291*** (0.0023)		-0.0294*** (0.0024)	-0.0157 (0.0145)
<i>Economy</i>	-0.0018* (0.0009)		-0.0019** (0.0009)		-0.0021** (0.0009)	0.0059 (0.0052)
<i>Transport</i>	-0.0016*** (0.0001)		-0.0015*** (0.0001)		-0.0015*** (0.0001)	-0.0002 (0.0008)
<i>Open</i>	-0.0141*** (0.0005)		-0.0133*** (0.0005)		-0.0132*** (0.0005)	-0.0185*** (0.0029)
<i>Constant</i>	0.5100*** (0.0027)	0.4455*** (0.0017)	0.4933*** (0.0030)	0.4365*** (0.0023)	0.4919*** (0.0030)	0.5318*** (0.0269)
三位数行业固定效应			是	是	是	是
二位数行业固定效应			是	是	是	是
行业间交叉固定效应			是	是	是	是
Observations	375662	375662	375636	375636	364034	11602
R-squared	0.0150	0.0089	0.1160	0.1118	0.1161	0.0926

注:***、**、* 分别对应 1%、5%和 10%的显著性水平。括号内为标准差。

区内,因此,造成多产业空间分布集中的现象;而发达地区特别是东部沿海城市群市场经济发展迅速,非国有规模以上企业较多,在空间上表现为更分散。城市群交通基础设施对产业间共聚的影响显著为负,说明交通基础设施越发达,运输成本下降,受到的地理空间制约越弱,进而在空间上表现为产业间的分散化。对外开放程度越高,产业的共聚效应越弱,这与很多关于单一产业集聚的实证研究结论不一致。对此的解释是,产业集聚有别于产业共聚,单一产业集聚并不意味着与其他产业共聚,对外部投资规模依赖程度较高的产业其本身对资本的单一依赖使得其不需要通过与其他产业共聚实现利益最大化。

第(5)、(6)列分别报告了跨行业与同行业的分组实证结果,在行业与区域的效应上,跨行业共聚表现出了更强的效应,这与上文的密度分布分析是一致的,跨行业共聚目的性更强,受行业与区域的效应影响更为明显。与之相对应的是,同行业的共聚实证结果发生较大变化,投入产出关联与技术关联对同二位数行业内部共聚的影响没有得到验证。不过,同二位数行业内规模差距引致的

“学习效应”影响更为强烈。此外,同二位数行业共聚不受经济发展水平、政府规模与交通基础设施的影响,却显著地受到对外开放程度的影响,影响方向与总体产业共聚趋势是一致的。这些都说明与跨二位数行业共聚相比,同二位数行业间产业共聚的影响模式更为单一。

六、结论与政策建议

以城市群为空间尺度的产业集群综合体将是区域经济发展新的增长极,而城市群产业空间治理的关键是广域空间下以城市群为支撑平台的产业空间与产业链整合的治理(陈建军等,2018)。本文基于产业共聚指数描绘的产业空间具备地理空间与统计分布特征,从静态与动态特征全面描绘区域内部产业空间,并进一步分析产业空间共聚的影响因素,为城市群决策者开展产业空间治理提供了一个分析框架。本文研究发现:从静态角度看,中国城市群层面同二位数产业共聚水平显著高于全部产业共聚水平与跨二位数产业共聚水平,不同城市群的产业共聚分布存在显著差异,产业空间核心产业各具特色,初步形成了全国范围内的产业分工格局。从动态角度看,中国城市群产业空间共聚水平总体呈下降态势,不同城市群之间表现出具有产业共聚总体两极分化与局部产业合作纵深发展等特征的差异化演化模式。从产业共聚与产业空间的影响因素看,产业层面上投入产出关联与技术关联对产业间的共聚发生的影响显著,产业间规模差距产生了跨产业的学习效应;在城市群层面上的政府规模、经济水平、交通基础设施与开放程度都不同程度地与产业共聚负相关。

结合以上分析框架与研究结果,本文的政策启示在于:①建立与完善以城市群为主导的产业空间治理体系,将产业互动融合与产业链高效整合作为城市群产业空间治理的重点与方向。各城市群在寻求产业结构升级并高质量发展过程中,要充分考虑产业间的内在联结关系,特别是产业间的单方向关联,对这一客观规律的准确认知有利于区域产业结构调整的高效、合理与科学决策。②产业空间治理存在较大的政策施展空间。一方面,城市群跨二位数产业间的共聚水平仍有较大的发展空间,利用跨二位数产业的事实共聚突破传统产业划分壁垒,进而实施高效产业空间结构调整是未来城市群产业治理的重要抓手;另一方面,产业性质相似的产业间存在空间依赖,城市群产业政策需要充分调动同二位数产业的共享生产要素,合理配置产业上下游资源高效共聚以形成良性的产业发展生态圈。③鉴于当前中国各个城市群的产业共聚程度与模式存在较为明显的不同,城市群治理群体需要制定差异化的产业治理策略。决策者可结合本区域内现有产业空间分布特点,摸清内部产业空间结构,促使具备投入产出依赖与技术依赖等共聚条件产业间共聚的“从无到有”,提升高效能、高产能产业间共聚的“从有向优”。④面对产业高质量发展与战略性新兴产业发展的客观需求,地方政府应当进一步为高技术产业的有效共聚提供良好的外部条件。一方面,要尊重投入产出关系与技术依赖对产业共聚影响的客观规律,在建立高新技术园区时充分考虑园区内部产业细分技术差距与特点,切勿“生拉硬搬”强行上马一些不符合产业规律的项目;另一方面,经济落后地区应当进一步发挥市场的自发力量(如马歇尔外部性、规模经济等),虽然这些地区通过税收优惠、补贴等政策吸引部分技术密集型产业集中在某个区域,但是产业间充分发挥劳动力、技术与市场的正外部性,仍然需要进一步借助市场力量优化当前技术密集型产业的空间布局。⑤对于区域内部诸如外资投资、交通基础设施等多个经济因素对产业空间分布的冲击效应,各个区域应当从实际出发,统筹规划,形成相互合作、良性竞争的产业空间新稳态。

产业共聚指数的内涵十分丰富,本文仅从两年的测算结果进行分析,结合投入产出表对产业单向共聚与双向共聚的模式进行分类,挖掘产业链、产业生态与产业空间分布的内在机理是未来的一个重要研究拓展方向;同时,仅从政府规模的角度讨论产业空间共聚的影响因素是不充分的,中国

特色社会主义市场经济中的政府产业政策引导对产业空间分布的影响也十分重要,未来的研究中可以进一步拓展政府主导产业政策等因素对产业空间共聚的影响。此外,基于这一测度从城市经济学角度进一步验证雅各布外部性效应也是一个重要的研究拓展方向。当然,这一测度思想还可以扩展延伸至城市群空间治理的各个领域,如土地管理和生态治理等带有空间地理特征的主题,均可以运用这一方法开展空间结构层面的讨论。

〔参考文献〕

- [1]陈国亮,陈建军. 产业关联、空间地理与二三产业共同集聚——来自中国 212 个城市的经验考察[J]. 管理世界, 2012,(4):82-100.
- [2]陈建军,陈菁菁,陈怀锦. 我国大都市群产业——城市协同治理研究[J]. 浙江大学学报(人文社会科学版), 2018,(5):166-176.
- [3]陈曦,朱建华,李国平. 中国制造业产业间协同集聚的区域差异及其影响因素[J]. 经济地理, 2018,(12):104-110.
- [4]贺灿飞,肖晓俊. 产业集聚、产业共聚与中国制造业生产率[J]. 哈尔滨工业大学学报(社会科学版), 2012,(1):111-120.
- [5]贺灿飞,胡绪千. 1978 年改革开放以来中国工业地理格局演变[J]. 地理学报, 2019,(10):1962-1979.
- [6]黄金川,陈守强. 中国城市群等级类型综合划分[J]. 地理科学进展. 2015,(3):290-301.
- [7]邵朝对,苏丹妮,李坤望. 跨越边界的集聚:空间特征与驱动因素[J]. 财贸经济, 2018,(4):99-113.
- [8]文东伟,冼国明. 中国制造业产业集聚的程度及其演变趋势:1998—2009 年[J]. 世界经济, 2014,(3):3-31.
- [9]袁海红,张华,曾洪勇. 产业集聚的测度及其动态变化——基于北京企业微观数据的研究[J]. 中国工业经济, 2014,(9):38-50.
- [10]赵娜,王博,刘燕. 城市群、集聚效应与“投资潮涌”——基于中国 20 个城市群的实证研究[J]. 中国工业经济, 2017,(11):81-99.
- [11]郑昌金,章登义,苏科华,武小平,洪程. 基于 LBP 特征和熵正则化 Wasserstein 距离的人脸表情识别[J]. 计算机与数字工程, 2017,(2):242-246.
- [12]Behrens, K., and T. Bougna. An Anatomy of the Geographical Concentration of Canadian Manufacturing Industries[J]. Regional Science and Urban Economics, 2015,(51):47-69.
- [13]Billings, S. B., and E. B. Johnson. Agglomeration within an Urban Area [J]. Journal of Urban Economics. 2016,(91):13-25.
- [14]Brandt, L., J. Van Biesebroeck, and Y. Zhang. Creative Accounting or Creative Destruction? Firm-level Productivity Growth in Chinese Manufacturing[J]. Journal of Development Economics. 2012,97(2):339-351.
- [15]Carlsson, J. G., M. Behroozi, and K. Mihic. Wasserstein Distance and the Distributionally Robust TSP[J]. Operations Research, 2018,66(6):1603-1624.
- [16]Cuturi, M. Sinkhorn Distances: Lightspeed Computation of Optimal Transportation Distances [J]. Advances in Neural Information Processing Systems, 2013,(26):2292-2300.
- [17]Delgado, M., M. E. Porter, and S. Stern. Defining Clusters of Related Industries [J]. Journal of Economic Geography, 2016,16(1):1-38.
- [18]Duranton, G., and H. G. Overman. Testing for Localization Using Micro-Geographic Data [J]. Review of Economic Studies. 2005,72(4):1077-1106.
- [19]Duranton, G., and H. G. Overman. Exploring the Detailed Location Patterns of UK Manufacturing Industries Using Microgeographic Data[J]. Journal of Regional Science, 2008,48(1):213-243.
- [20]Ellison, G., and E. L. Glaeser. Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach[J]. Journal of Political Economy, 1997,105(5):889-927.
- [21]Ellison, G., E. L. Glaeser, and W. R. Kerr. What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns[J]. American Economic Review, 2010,100(3):1195-1213.

- [22]Guo, Q., and C. He. Production Space and Regional Industrial Evolution in China [J]. *GeoJournal*, 2017,82 (2):379–396.
- [23]Hidalgo, C. A., B. Klinger, A. L. Barabási, and R. Hausmann. The Product Space Conditions the Development of Nations[J]. *Science*, 2007,317(5837):482–487.
- [24]Krugman, P. History and Industry Location——The Case of the Manufacturing Belt [J]. *American Economic Review*, 1991,81(2):80–83.
- [25]Luise, G., A. Rudi, M. Pontil, and C. Ciliberto. Differential Properties of Sinkhorn Approximation for Learning with Wasserstein Distance[R]. 32nd Conference on Neural Information Processing Systems, 2018.
- [26]Martin, A., S. Chintala, and L. Bottou. Wasserstein Generative Adversarial Networks [R]. International Conference on Machine Learning, 2017.
- [27]Odum, E. P. *Fundamentals of Ecology*[M]. Stanford, CT: Cengage Learning, 2004.
- [28]Porter, M. E. *The Competitive Advantage of Nations*[M]. New York: The Free Press, 1990.
- [29]Porter, M. E. Clusters and the New Economics of Competition[J]. *Harvard Business Review*, 1998,76(6):77.
- [30]Sinkhorn, R. A Relationship between Arbitrary Positive Matrices and Doubly Stochastic Matrices [J]. *Annals of Mathematical Statistics*, 1964,18(2):876–879.

Industrial Coagglomeration and Industrial Space Governance from the Perspective of Urban Clusters: Measurement of Machine Learning Algorithm

CHEN Lu^{1,2}, LIU Xiu-yan¹, YE Xin-yue², HU Han-hui¹

(1. School of Economics and Management, Southeast University, Nanjing 211189, China;

2. Ying Wu College of Computing, New Jersey Institute of Technology, Newark 06102, America)

Abstract: Urban clusters will be the main carriers and important geographical units of China's industrial spatial governance in the future. Based on the Wasserstein distance algorithm of machine learning, this paper uses micro-enterprise geographic data and applies matrix expansion Sinkhorn algorithm and entropy regularization constraint method to improve the solution method. We construct industrial coagglomeration index through hypothesis testing and Monte-Carlo simulation. This paper calculates the industrial coagglomeration level of China's urban clusters. Based on the results, it is found that the level of industrial coagglomeration within two-digit industries is higher than that cross two-digit industries in China's urban clusters. The relatively economically backward urban clusters show more obvious characteristics of industrial coagglomeration, meanwhile, the technology-intensive industries are more inclined to coagglomeration in those regions. From 2007 to 2013, the overall level of industrial coagglomeration was reduced and the evolution patterns of urban clusters are different among China's urban clusters. Based on the directivity of the industrial coagglomeration index, this paper analyzes the differences of industrial spatial core industries of different urban clusters and draws the structure chart of industrial space. Then, this paper studies the influencing factors of industrial coagglomeration and finds that input-output correlation, technical correlation and scale gap have significant positive effects on industrial coagglomeration, while level of internal competition in the industry inhibits the trend of industrial coagglomeration from two directions: coagglomerate and being coagglomerated. The regional characteristics of urban clusters, such as government size, are negatively correlated with the level of industrial coagglomeration. Accordingly, this paper puts forward the corresponding policy suggestions for the implementation of urban clusters industrial space governance.

Key Words: industrial coagglomeration; industrial space governance; machine learning algorithm; urban cluster

JEL Classification: C63 L60 R12

[责任编辑:章毅]