
基于街道网络的城市分区对 制造业生产效率的影响

赵晶^{1,2} 韩清¹¹

(1. 上海社会科学院 经济研究所, 上海 200235;

2. 济宁学院 数学系, 山东 曲阜 273155)

【摘要】: 经济学的研究通常以行政区划为对象, 探讨变量之间的相关关系, 然而, 已有文献表明, 城市的行政区划缺乏灵活性且容易忽略空间拓扑关系, 造成拟合结果的偏差。在空间句法理论中, 街道被认为是城市发展与演化的重要属性, 可以用于城市区块的划分, 而在街道基础上形成的城市分区也反映了更真实的城市状况。基于此, 以上海市杨浦区为例, 运用不同的社区检测技术对街道进行分类, 构建了基于街道网络的新的城市分区, 并比较了不同分区在产业间协同集聚对制造业生产效率影响上的拟合效果优劣。旨在为经济学研究提供一种新的思路和方法, 以探索更为真实的变量关系, 对城市规划和管理具有重要的意义。

【关键词】: 街道网络 城市分区 社区检测 制造业效率

【中图分类号】: F062.9; F062.2 **【文献标识码】:** A **【文章编号】:** 1671-4407(2021)06-074-07

最近的研究表明, 人们的“视觉感知地图”可能比行政区划边界更有意义^[1], 更能反映局部区域之间的经济与社会联系。如果仍然将传统的行政区划作为研究单位, 能否真正反映变量之间的相关关系呢?而在此基础上进行的资源配置又能否实现利益最大化呢?

在我国, 对于城市内部的划分通常是基于行政区划, 是国家为了进行分级管理而实行的区域划分, 但是这一级别的分区在内部特征和地理细节上存在较大的差别^[2], 缺乏一定的可比性, 而且也无法捕捉行政边界随时间发生的变化。Hillier & Hanson^[3]等提出的空间句法理论较好地解决了这一问题, 该理论将空间看作是城市结构的基本单位, 空间之间通过“街道”相互连接, 最终形成复杂的城市系统。Hillier 认为, 人类不可能像鸟一样, 瞬间从一个空间移动到另一个空间, 当人们以某种方式移动时, 必须穿过街道; 而由于视觉和感知的变化, 影响运动流的关键因素不再依赖于距离, 而是依赖于街道的“连接值(连接的街道数量)”和“深度值(到达其他街道所需改变的方向)”, 即空间之间的内在逻辑。在人们探索环境的运动中, 视觉和感知会发生动态变化^[4], 这种变化最终决定了人们前进的方向; 沿着街道的运动带动了土地的使用及周边商业的增长, 这反过来又鼓励人们向这些地区迁移, 最终形成城市的中心和边界, 推动了城市的发展和演化。

空间句法理论运用轴线图来表示街道的拓扑结构, 同时也使其具备了网络的基本特征, 因此可以用网络的研究方法来分析

作者简介: 赵晶, 博士研究生, 研究方向为区域经济学与城市经济学。E-mail:m15853700795@163.com

韩清, 博士, 研究员, 博士研究生导师, 研究方向为数量经济学、统计学、金融计量学。E-mail:qhan@sass.org.cn

基金项目: 国家自然科学基金青年项目“处理效应模型的非参数估计方法及其拓展应用”(71803134)

街道的“小世界”属性。由于社区检测算法具有定义子图的特性，因此可以用于网络分区。Law^[1]指出，运用社区检测算法检测出的局部区域内的街道之间具有最相似的特征和最紧密的联系，更符合不同区域之间比较的共同假设，能更好地捕捉城市发展进程中的区域效应，探索变量之间的相关关系。由于街道网络的节点数多且结构复杂，因此不能用传统的社区检测算法进行分区，只能选用适合大型网络的基于模块度的算法，并根据模块度的数值选择最好的分区结果，用于经济模型的研究。

由于街道在空间句法理论中被映射为点，因此，可以运用社区检测技术对其进行划分，将最相似的街道划分为一组，形成一个新的区块。由于这种分区技术考虑了街道的拓扑结构和城市的空间逻辑，因此，每个区块中所包含的信息也是最相似和最匹配的，确保了研究变量在分析之前的一致趋势，能更真实地反映变量之间的相关关系。

本文的主要目的是提供基于街道网络的新的城市分区方法，并比较新的分区与行政区划分区在模型拟合效果上的优劣。为此，本文以产业间协同集聚对制造业生产效率的影响为例，分别运用混合 OLS 模型、固定效应模型和随机效应模型对不同方法的分区进行比较，结果均表明新的分区方法在模型拟合上具有更好的效果。

与已有的研究相比，本文的贡献主要体现在三个方面：一是将空间句法理论应用于城市分区的研究上，不仅考虑了视觉感知对运动流的影响，还关注了街道网络的空间拓扑结构，提供了微观数据区域化的新思路；二是基于街道网络构建了局部区域单元(SLA)，并将其重新聚合形成新的城市分区，解决了传统行政区划内部特征(如空间结构、经济与社会特征等)不一致性的问题；三是运用检测性能较好的基于模块度的算法来划分街道网络，确保所构建的分区在街道层面上具有内部的最大联系和外部的最小联系；四是比较了不同社区检测技术分区结果的优劣，验证了社区检测模块度与街道网络划分结果之间的关系。

1 文献综述

无论是研究数据还是研究单位，国家统计的方式大多是基于行政区划，但是越来越多的研究表明，这种划分的详细程度和灵活性比较有限，也不一定能够产生一致的实证结果。有的行政区划甚至会横穿街道和建筑物，根本无法反映行政区划内部的真实状况以及行政区划之间的相互联系，甚至会对次级产业市场产生影响。Leishman^[4]指出，这种影响已经扩散到次级房地产市场。而 Goodman & Thibodeau^[5]对纽黑文市街区和人口普查区，Bourassa 等^[6]对墨尔本和悉尼住宅数据分组区和人口普查区以及 Collins & Kaplan^[7]对田纳西州学区变化的研究均表明，不同的分区会产生完全不同的结果。随着空间句法理论的发展，学者们开始研究一种更符合人们视觉感知行为的分区方法，以更好地拟合模型，探索不同变量之间的相关关系。

1984 年，Hillier & Hanson^[3]出版了 *The Social Logic of Space* 一书，该书奠定了空间句法的理论基础，创设性地提出用“轴线图”的方法来分析城市的空间形态。该方法运用街道为节点，交叉点为边缘的对偶图来表示街道网络系统，形成覆盖街道网络的最少线路图，并进一步运用“深度”和“整合度”等变量来反映街道的连通属性。相比于原始图(交叉点为节点、街道为边缘)，对偶图忽略街道的公制长度，直接将街道映射为点，交叉点映射为边缘，这意味着每条街道都可以连接无数个其他街道，有利于与其他生物、社会网络对比分析。不仅如此，由于街道之间的关系在对偶图中被反映为点对点的关系，因此在运用社区检测方法聚类时速度更快，精度更高。然而，由于对偶图放弃了对街道的距离度量，因此也一直被认为是主观的轴向映射构建过程^[8]，与现实生活存在较大差距。针对这一质疑，Porta 等^[9]进一步对两种图示方法进行了比较，结果表明，对偶图四个中心性指标的累积分布基本表现出和原始图一致的趋势。

Jiang & Claramunt^[8]对美国 40 个城市的研究发现，街道系统表现出明显的网络特征——“小世界性”与“无标度性”，即在长度大于平均数的街道中，仅有 1%能构成街道网络的主干，因此要揭示街道网络的拓扑结构和固有模式就必须将单个街道合并为有意义的街道片段，而社区检测定义子图的特性恰好可以用来解决合并问题。传统的社区检测技术包括部分聚类、层次聚类、光谱聚类以及图聚类，这些算法的时间复杂度较高，不适用于超过几百个节点的大型网络，而且如果前期处理不好，还会影响分区的效果。

2002年, Newman^[10]首次将模块度(Q-Modularity)的概念用于社区检测,为比较社区划分结果的优劣提供了理论基础。模块度本身是质量函数,取值在-1~1之间,越高表明社区划分效果越好。Newman在模块度的基础上开发了Fastgreedy(贪婪算法),这是目前应用最为广泛的社区检测技术。由于在求解问题的过程中,该算法始终选择当前情况下的最优解,因此可以在大型网络中迅速奔跑。2008年,Blondel等^[11]在Fastfreedy算法的基础上进行优化,开发了Louvain算法,不仅在分区的过程对节点排序,而且首次引入分层的概念,将每次迭代形成的社区作为新的超节点重复迭代,大大缩短了运算时间。随着算法技术的不断成熟,考虑节点随时间变化的动态算法,如随机游走、自旋模型等逐渐发展起来,但是由于网络复杂性的提高,运算精度和运算效率均不是很高,且容易出现无社区归属的单独节点。

以往的文献主要研究了不同分区在拟合结果上是否具有不同,并没有比较不同分区方法的优劣。Law^[1]首次将社区检测技术中的Louvain算法应用于街道网络的对偶图,构造了英国大伦敦地区的局部区域单元,并指出局部区域单元对房地产市场的影响要明显优于其他的行政区划单位。然而, Law^[1]并没有进一步验证运用不同社区检测技术对街道网络划分的优劣,也就是说,具有高模块度的社区检测技术在具体的拟合过程中是否要优于低模块度的社区检测技术。

为此,本文分别运用Fastfreedy算法和Louvain算法进行了不同的划分,并将新的分区分别与传统行政区划的拟合效果进行对比,以验证新分区方法的可靠性。

2 理论框架

2.1 以街道网络为基础的城市分区

空间句法理论将城市定义为建筑物的集合体,而建筑物之间通过街道相互连接,街道之间形成的网络构成了城市的空间逻辑。在空间句法理论中,如果一条街道所连接的其他街道的数量较多,那么该街道被认为是高度整合的,而整合度高的街道通常具有较强的吸引力,可以吸引人流和车流在此聚集,从而形成城市的分中心,而整合度较低的街道或者成为主干道的附属街道,或者与主干道之间形成一定的边界,这样,城市就自然地被划分成多个区块,这些区块内部的街道之间具有最相似的特征和最紧密的联系,而区块之间的街道则差别较大。

在图论中,街道网络是具有复杂联系关系的平面图。为了简化网络关系,空间句法理论给出了街道网络对偶图的构建方法,即将街道映射为节点,而将街道之间的连接映射为边缘,具体的模型为:

具体可以表示为: $DG(u, e)$, 其中, u 是节点(街道), e 是边(街道之间的连接)。

2.2 基于街道网络对偶图的Fastgreedy和Louvain算法

社区检测技术的最终目的是获得内部结构最为相似的子图,对本研究来说,就是在模型(1)的基础上找到最优的 $SG_k(k=1, 2, \dots, K)$, 其中, SG_k 是 DG 的子图, K 是子图的个数。Fastgreedy和Louvain均是基于模块度优化的算法,在社区检测领域以速度快和精度高而著称。模块度本身是质量函数,用于比较不同检测算法的准确性,模块度越高表示社区划分越准确,具体的计算公式如下:

$$Q = \frac{1}{2m} \sum_{ij} \left[A_{ij} - \frac{d_i d_j}{2m} \right] \delta(c_i, c_j) \quad (1)$$

式中: Q 为模块度; m 为网络中边的总数; d_i 和 d_j 分别表示节点 i 和节点 j 的度; A_{ij} 表示邻接矩阵,当 i 和 j 之间有一条边

时, $A_{ij}=1$, 否则, $A_{ij}=0$; 而当 i 和 j 属于同一个社区时, 函数 $\delta(c_i, c_j)=1$, 否则, $\delta(c_i, c_j)=0$ 。

Fastgreedy 算法基于树状图的思想, 重复地将成对节点中使得 Q 值增加最多(或减少最少)的两个节点连接在一起, 形成一个社区, 当最后只剩下一个社区时, 算法终止, 该树状图就代表了网络在各个层次上的分解结构; 由于没有公共边的社区连接不会增加 Q 值, 因此, 该算法在计算过程中只存储一条边或多条边连接的成对社区的 ΔQ 值, 故大大缩短了运行时间。

Louvain 算法主要分为两个步骤, 第一阶段的过程与 Fastgreedy 算法类似, 只是在算法伊始对每个节点进行排序, 以节约运行时间^[11]; 在第二阶段, 以第一阶段形成的社区为超节点构造一个新的网络, 继续进行聚类, 通过这种方式可以减少最终形成的社区数量, 提高运算速度。具体的实现步骤如图 1 所示。

2.3 验证不同分区在产业间协同集聚对制造业生产效率影响上的不同

Fastgreedy 算法和 Louvain 算法得到的街道子图结构仍然为对偶图形式, 需要将对应的节点映射为街道, 连接映射为交叉点, 在原始图中实现分区。具体的映射模型为:

具体可以表示为: $PG(u, e)$, 其中: u 是街道(节点), e 是交叉点(边缘)。

为了比较不同分区的拟合效果, 本文参考胡玫等^[12]的做法, 将工业生产总产值作为产出指标, 将制造业从业人数、流动资产年平均余额以及固定资产年平均余额作为投入指标, 测算出制造业的 Malmquist 指数, 并将其作为制造业生产效率的替代变量, 构建了如下回归模型:

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \beta_2 X_{cmt} + \varepsilon_{it} \quad (2)$$

式中: Y_{it} 表示测算出的制造业 Malmquist 指数; i 表示城市区块的序号; t 表示时间; X_{cmt} 表示一系列控制变量, 参照郝书辰等^[13]对工业企业效率指标的分析, 共选取 7 个控制变量, 分别为人均工业总产值、人均资产合计、人均负债合计、人均利润总额、人均销售收入、人均应交增值税以及人均出口交货值, 其中, 由于 2004 年后的工业企业数据库中没有销售收入的数据, 用营业收入代替; β_0 是常数项, β_1 表示产业间协同集聚对制造业生产效率的影响; β_2 表示控制变量对制造业生产效率的影响; ε_{it} 是随机变量; X_{it} 表示制造业产业间协同集聚指数, 即产业间协同集聚指数大于 0 的组合占有所有产业组合的比例。产业间协同集聚指数采用 Glenn 等^[14]提出的测度指标来计算, 具体公式为:

$$EG_{ij} = \frac{\sum_{r=1}^R (s_{ir} - x_r)(s_{jr} - x_r)}{1 - \sum_{r=1}^R x_r^2} \quad (3)$$

式中: EG_{ij} 表示行业 i 和行业 j 的产业间协同集聚指数; s_{ir} 表示区域 r 行业 i 的就业份额; s_{jr} 表示区域 r 行业 j 的就业份额; x_r 表示区域 r 所有行业就业份额的平均值; R 为所有区域的总数。

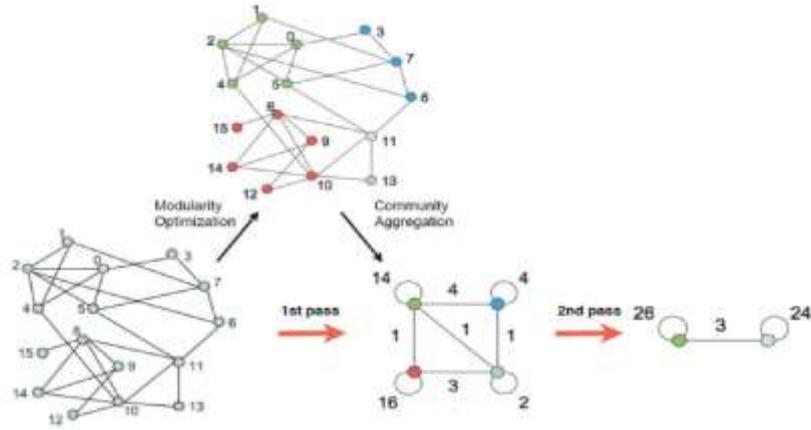


图 1 Louvain 算法主要步骤

注：该图来源于文献[11]。

考虑到不同分区可能存在固定效应和时间效应，本文又构造了如公式(4)和公式(5)所示的个体固定效应和双向固定效应模型：

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \beta_2 X_{c,it} + \mu_i + \varepsilon_{it} \quad (4)$$

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \beta_2 X_{c,it} + \mu_i + \lambda_t + \varepsilon_{it} \quad (5)$$

式中：各变量的含义均与模型(2)相同， μ_i 和 λ_t 分别表示个体固定效应和时间固定效应。

由于3种分区豪斯曼检验的结果并不一致，因此，本文又构造了如公式(6)和公式(7)所示的个体随机效应和双向随机效应模型：

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \beta_2 X_{c,it} + \mu_i + \varepsilon_{it} \quad (6)$$

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \beta_2 X_{c,it} + \mu_i + \lambda_t + \varepsilon_{it} \quad (7)$$

式中：各变量的含义均与模型(2)相同， μ_i 和 λ_t 分别表示个体随机效应和时间随机效应。

3 数据说明与实证分析

在前面的分析方法上，本文以上海市杨浦区为例，基于杨浦区街道网络图应用空间句法和社区检测相关理论实现对杨浦区的重新划分，具体的数据和方法如下。

3.1 数据与指标选取说明

本文最主要的数据源是杨浦区的街道网络图，经过整合与筛选，共确定了351条道路，其中包括内环、中环、军工和逸仙4条高架路的部分路段、6条隧道、23条一级道路、19条二级道路、24条高架、隧道连接路以及其他道路等。经过矫正以后，共

形成 1817 对不重复的连接关系。

第二种数据源是杨浦区的行政规划区域图，该图共包含 12 个区域，按从北向南、由西向东的顺序依次为新江湾城街道、殷行街道、五角场街道、五角场镇、四平街道、控江街道、延吉新村街道、长白新村街道、江浦街道、平凉街道、大桥街道和定海街道。

第三种数据来源于工业企业数据库。由于 2009 年以后的数据存在较大的错误与匹配问题，且缺乏重要的财务指标(流动资产年平均余额、固定资产年平均余额等)，因此本文只选取了 1998—2008 年 11 年的数据。

3.2 实证方法

3.2.1 构建杨浦区街道网络对偶图

由于街道网络图为平面结构的拓扑图，无法体现高架路和隧道与其他街道的连接关系，因此本文首先将杨浦区街道网络图中隶属于同一条街道的片段整合，剔除高架路与隧道构造对偶图，然后根据百度全景地图提取高架路和隧道的连接关系，添加到对偶图中，最后形成具有 351 个节点，1817 条边的街道网络图。原始图与对偶图分别如图 2 中的 a、b 所示。

3.2.2 分别运用 Fastgreedy 和 Louvain 算法进行街道网络分区

首先，根据杨浦区街道网络对偶图可以得到街道之间的连接关系表。然后分别基于 Fastgreedy 和 Louvain 的算法原理形成街道网络分区，其中，Fastgreedy 算法得到 8 个分区，Q(模块度)值为 0.63，Louvain 算法得到 10 个分区，Q 值为 0.67。单纯从 Q 值上来看，Louvain 算法的分区结果确实优于 Fastgreedy 算法。最后，将分区结果映射回原始图，勾勒出杨浦区的新分区，分区结果如图 3 所示。

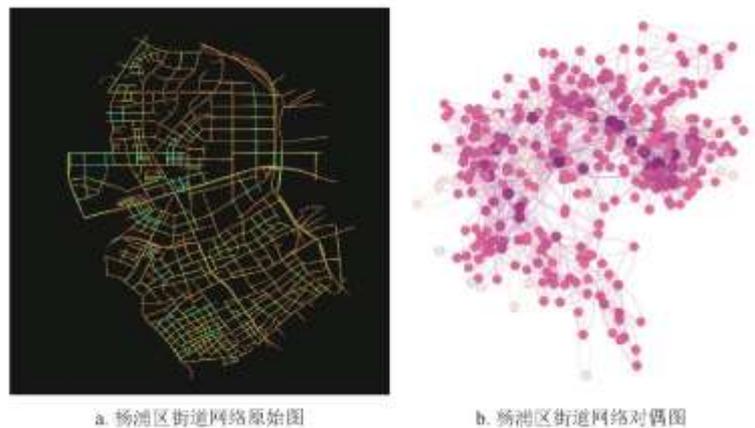


图 2 杨浦区街道网络原始图与对偶图

注：a 图中道路颜色越深代表该道路的连接值越大，也即连通性更好；b 图中节点颜色深浅代表节点度的大小，颜色越深表示节点度越大。



图3 杨浦区行政区划图与社区检测算法分区图

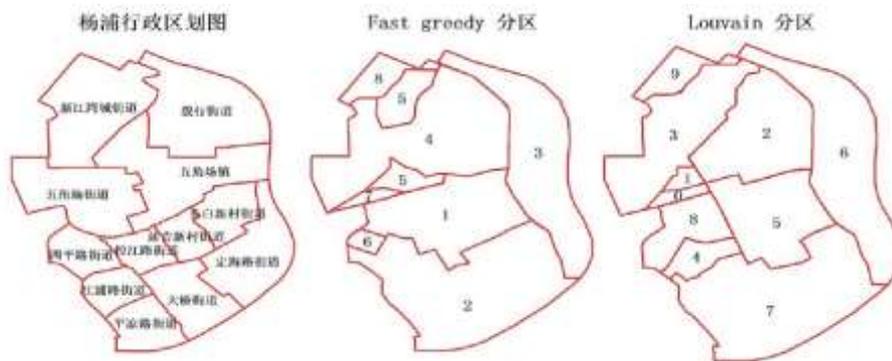


图4 杨浦区行政区划图与社区检测算法分区编号图

注：图中黑色边界表示分区，编号或者街道名称表示城市内部的区块。Fastgreedy 分区实际为 8 个区块，有两个编号为 5 的区块。

由图 3 可以明显地看出三者之间的区别，Fastgreedy 和 Louvain 两种社区检测算法均将杨浦区东部沿海以及南部沿海地区分别看作一个独立的整体，这表明东部沿海、南部沿海地区的街道拓扑结构最为相似，但是杨浦区行政区划图并没有考虑到这一因素，而是将东部沿海地区分割到殷行、五角场镇、长白新村与定海 4 个街道区域，而将南部沿海地区划分为江浦、平凉、大桥、定海 4 个街道区域。单纯从地图上看，东部沿海的街道地形特征比较明显，即街道数量少且比较分散，具有相似的表象特征，因此，将东部沿海划分为一个整体的区域可能会更加合理。

3.2.3 其他变量信息

(1) 企业位置的确定。

首先，分别将确定的 3 种分区图加载到地图上，并为分区图中的每一区块编号，分区结果如图 4 所示，其中，杨浦区行政区划分为 12 个区块，Fastgreedy 算法分区图分为 8 个区块，Louvain 算法分区图分为 10 个区块；然后，提取工业企业数据库中的法人信息并编号，按照编号所确定的顺序将法人的地址信息转换成经纬度并定位在分区图上；最后，确定分区图中每一区

块所包含的法人编号。

(2) 数据的填入。

首先，按照步骤(1)中确定的区块编号与法人编号的对应关系统计相应的变量信息；然后按照前述方法计算每一区块的 Malmquist 指数和产业间协同集聚指数，并统计每一区块内的控制变量信息；最后，分别运用模型(2)、模型(4)和模型(5)进行回归拟合。

3.2.4 实证分析结果

首先运用模型(2)对3种分区分别进行回归，并比较回归结果。经检验，本文的数据不存在显著的异方差和序列相关，因此只采用稳健的标准误进行修正，回归结果如表1所示。

表1 3种分区的混合 OLS 结果

变量	杨浦区行政区划	Fastgreedy 分区	Louvain 分区
β_1	0.374**	1.087***	1.033***
β_2	-4.92e-07	1.16e-06***	1.41e-06**
β_0	0.584***	0.038**	0.033***
R^2	0.0619	0.825	0.792
样本量	132	88	110

(1) 混合 OLS 回归结果。

由表1的结果可以看出，3种分区所得到的结果具有较大的差异。从模型的拟合性上看，Fastgreedy 和 Louvain 算法的 R^2 远大于杨浦区行政区划的结果，比行政区划分区 R^2 的10倍还要大。这表明，基于街道网络所构造的分区在探讨产业间协同集聚对制造业生产效率影响的问题上具有更高的可信性，街道网络的拓扑结构确实对制造业产业集聚产生了影响，这与 StephenLaw 的研究结果也是相同的。

从参数显著性上看，无论是 Fastgreedy 算法还是 Louvain 算法，产业间协同集聚水平在1%的水平上均显著，而基于行政区划的结果只在5%的水平上显著。从系数的大小上看，两种社区检测算法所得到的分区结果较为相似，均略大于1，而在行政区划分区上的结果只为0.4左右，低估了产业间协同集聚对制造业生产效率的影响。

(2) 个体固定效应与双向固定效应结果。

考虑到区块可能存在的固定效应，对3种分区分别进行豪斯曼检验，由于3种分区的结果对于选择固定效应还是随机效应的结果并不一致，因此先运用个体固定效应(模型4)和时间个体双向固定效应模型(模型5)进行估计，回归结果如表2所示。

表2 3种分区的个体固定效应与双向固定效应结果

变量	杨浦区行政区划		Fastgreedy 分区	Louvain 分区		
	个体	双向	个体	双向	个体	双向
β_1	-0.343	-0.411	0.349	0.492	0.925***	0.901***
β_2	-2.39e-07	3.90e-08	4.65e-07**	5.05e-07	3.17e-07	4.14e-07
β_0	1.058***	1.103***	0.523**	0.373	0.146**	0.146***
μ_i	Y	Y	Y	Y	Y	Y
λ_t		Y		Y		Y
R^2	0.028	0.119	0.059	0.182	0.628	0.694
样本量	132	132	88	88	110	110

从表 2 可以看出，首先，时间个体双向固定效应的 R^2 要比个体固定效应模型的 R^2 略高，且时间虚拟变量中约有一半是显著的，这表明，产业间协同集聚对制造业生产效率的影响在不同年份是不同的。其次，Fastgreedy 算法和 Louvain 算法所得到的分区进行回归后的 R^2 均比行政区划分区的 R^2 要大，这再一次表明，街道网络的拓扑结构在分析制造业产业集聚问题上的重要作用，也说明正确分区更能发现经济问题中的变量关系。

在参数显著性上，Louvain 分区的结果不只在 1% 的水平上显著，且接近混合 OLS 的系数值，近似为 1。而 Fastgreedy 分区与行政区划结果的分区的 R^2 虽然都不显著，但是就系数的符号来看，Fastgreedy 分区均为正，但是行政区划的结果均为负，也就是说，Fastgreedy 分区的结果更符合经济学意义。

(3) 随机效应结果。

分别运用个体随机效应[模型(6)]和时间个体双向随机效应模型[模型(7)]对 3 种分区进行估计，回归结果如表 3 所示。由表 3 可以看出，随机效应 R^2 的结果与固定效应结果呈现相同的结论，时间随机效应的结果略大于随机效应的结果，且运用社区检测算法得到的两种分区的 R^2 均比行政区划分区的 R^2 要大，也就是说，不同分区之间的内在特征存在较大的差异，这也进一步表明正确分区的重要性。

在参数的显著性上，行政区划依然为负且不显著，Louvain 分区的结果依然接近于 1 且在 1% 的水平上显著，唯一不同的是，采用随机效应模型估计后，Fastgreedy 分区的结果与 Louvain 分区接近，在 1 附近，且在 1% 的水平上也变得显著，这说明，基于社区检测技术在街道网络基础上的分区在分析经济变量之间的相关关系时，表现出更好的性能和可信性。

(4) 不同算法的比较。

由前述分析可知，Louvain 算法的 Q 值略高于 Fastgreedy 算法的 Q 值，也就是说，单纯从社区检测的角度来说，Louvain 算法的性能要好于 Fastgreedy 算法。

在表 1 中，基于 Fastgreedy 分区和 Louvain 分区的 R^2 均在 0.8 左右，但是根据表 2 和表 3 的结果，Louvain 分区的 R^2 明显高于 Fastgreedy 分区的 R^2 ，而且，无论是运用固定效应模型还是随机效应模型，Louvain 分区的系数均在 1% 的水平上显著，而 Fastgreedy 分区只在随机效应模型下显著，也就是说，Louvain 分区在实际的模型估计中也表现出了比 Fastgreedy 分区更好的

性能。因此，我们可以将社区检测算法的 Q 值与模型拟合优劣对应起来，推测 Q 值高的算法得到的分区具有更高的模型拟合度，而 Q 值低的算法得到的分区模型拟合度可能会更低。

4 结论与政策启示

(1) 与行政区划相比，两种基于街道网络的分区在进行模型回归时均表现出比行政区划更好的结果。这是因为，街道环境的微妙变化决定了人们认知感觉控制的运动流，进而影响了人们的经济与社会生活，最终表现为与街道网络相适应的经济现象，如在连通性更强的街道上集聚更多的劳动力和商业。

(2) 街道网络的拓扑强化了不同区域之间经济和社会特征的相似性和差异性，有利于在同样的基础(内部特征等)上分析变量之间的相关关系，得到更准确的分析结果。正是由于这一原因，两种基于街道网络的分区在模型拟合度上才得到远好于行政区划的值，且表现出更好的参数显著性。

(3) 在模型的估计上，无论是混合 OLS 模型、固定效应模型还是随机效应模型，当产业间协同集聚的系数显著时，值均在 1 左右，表现出一定的稳定性。这说明，只要分区是合理可靠的，得到的结论就是一致的，也就是说，在分析经济问题时，只有选择了合适的分区，才能捕捉到正确的变量相关关系，而这是单纯利用行政区划分区所不能得到的。

(4) 不同的社区检测方法得到的分区结构与行政区划具有明显的不同，而且 Q 值高的算法得到的分区要比 Q 值低的算法得到的分区模型拟合度更高，因此，可以将社区检测技术与城市内部分区研究相结合。

本文的结论对城市规划和资源配置具有重要的实际意义。首先，证实了街道网络在决定城市结构框架上的重要作用。本文的街道网络不仅包含一级、二级和其他一些旁路分支，还包括高架路和隧道，这些道路均会对行业产生影响，因此，相关部门在规划道路时应充分考虑新道路带来的街道网络变化。

其次，验证了不同分区结构对制造业的不同影响，证明了城市分区的必要性和重要性。我国的行政规划区域自划分以来就很少变动，所有的资源配置与责任界定均是按照行政区划的结构，这可能与街道网络的结构有较大差异，从而降低了资源的利用效率，造成了全要素生产率的下降。

最后，分析了区位差异与空间结构之间的关联关系，对分布较为广泛的服务行业及其他行业的位置选择与空间拓展具有重要的启示意义。随着城市的快速发展，街道的遍布性越来越广，所达之处在带去人流的同时，也带去了土地与商业的开发，进而吸引了服务性行业的进入，因此，当对城市进行规划时，应充分考虑街道的空间结构，为服务业的发展提供坚实的土壤。

表 3 三种分区的随机效应结果

变量	杨浦区行政区划		Fastgreedy 分区		Louvain 分区	
	个体	双向	个体	双向	个体	双向
β_1	-0.167	-0.210	1.075***	1.093***	0.964***	0.932***
β_2	-3.49e-07	-9.25e-08	1.11e-06***	1.19e-06***	1.19e-06	8.99e-07
β_0	0.944***	0.975***	0.049	0.001	0.101*	0.207***
μ_i	N	N	Y	Y	Y	Y

λ_t		N		Y		Y
R^2	0.022	0.111	0.058	0.154	0.620	0.687
样本量	132	132	88	88	110	110

参考文献:

[1]Law S.Defining street-based local area and measuring its effect on house price using a hedonic price approach:The case study of Metropolitan London[J]. Cities, 2017, 60:166-179.

[2]Lebel A, Pampalon R, Villeneuve P Y. A multi-perspective approach for defining neighbourhood units in the context of a study on health inequalities in the Quebec City region[J]. International Journal of Health Geographics, 2007. DOI: <https://doi.org/10.1186/1476-072X-6-27>.

[3]Hillier B, Hanson J. The social logic of space[M]. Cambridge:Cambridge University Press, 1989.

[4]Leishman C. Spatial change and the structure of urban housing sub-markets[J]. Housing Studies, 2009, 24(5) :563-585.

[5]Goodman A C, Thibodeau T G. Housing market segmentation and hedonic prediction accuracy[J]. Journal of Housing Economics, 2003, 12(3) :181-201.

[6]Bourassa S C, Hamelink F, Hoesli M, et al. Defining housing submarkets[J]. Journal of Housing Economics, 1999, 8(2) :160-183.

[7]Collins C A, Kaplan E. Capitalization of school quality in housing prices:Evidence from boundary changes in Shelby County, Tennessee[J]. American Economic Review, 2017, 107(5) :628-632.

[8]Jiang B, Claramunt C. Integration of space syntax into GIS:New perspectives for urban morphology[J]. Transactions in GIS, 2002, 6(3) :295-309.

[9]Porta S, Crucitti P, Latora V. The network analysis of urban streets:A dual approach[J]. Physica A:Statistical Mechanics and Its Applications, 2006, 369(2) :853-866.

[10]Newman M E J. Fast algorithm for detecting community structure in networks[J]. Physical Review E, 2004. DOI: 10.1103/PhysRevE.69.066133.

[11]Blondel V D, Guillaume J L, Lambiotte R, et al. Fast unfolding of communities in large networks[J]. Journal of Statistical Mechanics:Theory and Experiment, 2008. DOI:10.1088/1742-5468/2008/10/p10008.

[12]胡玫, 刘春生, 陈飞. 产业集聚对中国企业全要素生产率的影响——基于广东省制造业的实证研究[J]. 经济问题, 2015(4) : 78-82.

[13]郝书辰, 田金方, 陶虎. 国有工业企业效率的行业检验[J]. 中国工业经济, 2012(12): 57-69.

[14]Glenn E, Glaeser E L, Kerr W R. What causes industry agglomeration? Evidence from coagglomeration patterns[J]. American Economic Review, 2010, 100(3): 1195-1213.