

新经济地理视角下中国大陆区域 经济差异影响因素研究

庄赞 曾五一¹

【摘要】通过莫兰散点图和莫兰指数对中国大陆 2019 年省域经济空间相关性进行分析,表明中国大陆省域人均 GDP 存在显著的空间正相关性,经济发展水平接近的省份呈现明显的空间聚集,从而导致区域经济空间分布不均。在索洛增长模型包含物资资本和人力资本两种要素的基础上,结合新经济地理视角下区域经济增长的影响因素,根据经济理论隐含的空间相关变量的数据生成过程(DGP)选择空间杜宾模型(SDM),并利用 2003-2016 年中国大陆省域面板数据进行实证研究,通过模型所揭示的变量间复杂的数量关系,估计所研究的影响因素对经济增长的直接效应和溢出效应,从而为采取更有效的缩小区域经济差异的对策建议提供实证依据。

【关键词】经济地理 区域差异 影响因素 空间杜宾模型(SDM) 溢出效应

【中图分类号】:F061.5 **【文献标识码】**:A **【文章编号】**:1009-2382(2020)12-0008-09

一、引言

改革开放四十多年来,我国国民经济得到迅速发展的同时,区域经济发展空间差异的影响因素也成为关注的焦点。如有学者从理论层面分析了资本、劳动力、制度变革和结构变化等要素对地区经济差距变化的作用(王小鲁和樊纲,2004);有学者提出了影响地区经济发展差异的三大地理要素,并运用 OLS 估计得出这三大要素对中国的发展具有十分重要的影响(刘清春和王铮,2009);有学者运用 OLS 估计发现不同地方政府的行为倾向和经济政策选择存在着显著差异进而导致经济增长绩效的区域差异(张鹏飞等,2019);有研究运用空间计量 SLM 模型对中国 31 个省区市 1978-2015 经济发展水平进行分析,认为中国经济飞速发展越来越依赖于技术进步、人力资本提升、制度改革和市场化推进等因素的增强(程名望等,2019);另有研究通过选取省际面板数据并选用 SAC 模型用于溢出效应分析,认为区域金融化程度加深会提高本地区和邻近地区的实体经济发展(吴金燕等,2020);也有研究运用 SLM 探究全球化、市场化、分权化和本土化因素对长三角县域经济差异的作用机制(赵丹等,2020)。

从文献中关于区域经济差异影响因素的研究方法来看,所构建的计量模型由传统的线性回归模型向空间计量模型转变。区域经济学的前提是地理位置和距离是影响市场经济活动的重要力量,该理论依赖于空间相互作用、位置层次和空间溢出等概念。大量文献基于新经济地理学的角度研究了技术革新的空间溢出效应对区域经济增长的影响,为区域经济空间相关性提供了理论依据。利用空间经济数据研究经济变量的关系时,由于变量之间空间自相关的出现,将违背高斯马尔可科夫条件下计量经济模型的适用假定,传统线性回归模型的 OLS 估计量可能是有偏不一致的,从而无法正确解释区域经济增长过程。国外学者介绍了一系列处理空间样本数据的方法来构建空间计量模型(Anselin, L., 1988),不同的空间计量模型隐含不同的理论依据及数据生成过程(Data Generating Processes, DGP)。例如在新古典增长模型中已有一些拓展的理论研究为在生产函数模型中加入空间滞后变量作为解释变量提供理论依据,空间滞后变量是由相邻区域观测值的加权平均所构造,这些理论假设资本和劳动存在外部性及区域之间存在技术的相互依赖,可以推导出在模型中包含被解释变量的空间滞后项,空间计量模型方法可以检验模型中是否应该加入

¹作者简介:庄赞,集美大学财经学院副教授(厦门 361021);

曾五一,吉林财经大学统计学院特聘教授(长春 130117),上海对外经贸大学统计与信息学院特聘教授(上海 201620)。

基金项目:教育部青年基金项目“空间聚集与区域经济差异的统计测度及因素分析”(编号:16YJC90008);福建省高等学校新世纪优秀人才支持计划项目“福建省区域经济差异及其影响因素研究”(编号:JYXSJ201801)。

空间滞后项以及进行相应的统计推断来验证模型中所隐含的理论假设是否与实际相一致。

本研究对中国大陆区域经济差异影响因素的分析主要基于新经济地理学框架进行,并根据经济理论隐含的变量空间相关的数据生成过程(DGP)选择合适的空间计量模型进行实证研究。如何正确解读空间计量模型所揭示的变量间复杂的相互作用,准确把握影响中国区域经济增长的重要因素,这对我国实施区域均衡发展战略和进行经济空间结构的调整与优化具有重要作用。

二、新经济地理视角下区域经济差异的影响因素

2009年,世界银行发布的《世界发展报告——重塑世界经济地理》中以克鲁格曼“新经济地理”框架为基础,强调空间因素在经济活动中的作用,提出用密度(Density)、距离(Distance)、分割(Division)的“3D”框架实现经济地理的格局变迁(世界银行,2009)。

密度(Density)是指单位土地面积的经济产出或总购买力,用于描述经济活动的空间聚集程度,是新经济地理学中规模报酬递增导致经济活动集聚的表现。密度会从两个方面影响区域经济发展差异。在初始的均衡状态,一旦一个偶然的因素改变这种状态,由于市场力量,随着密度的提高,更多厂商开始在某个区域集聚,并带来更多的资源,进而形成所谓的发达地区,集聚会显著促进发达地区的经济发展,只要厂商在集聚的过程中获得的外部收益大于拥挤效应带来的损失时,就会有源源不断的厂商集聚在此,从而拉大区域间经济差距;但这种差距不会一直扩大下去,因为当密度过高时,即所有的厂商都在该区域聚集,会带来拥挤效应,厂商竞争力加大,并会使产品的生产成本提高,使得外部收益小于损失,此时集聚区的产业会向外转移,从而带动周边地区的经济增长,缩小集聚区与周边地区的经济差距。

距离(Distance)是指商品、服务等穿越地理空间所耗费的成本。一方面,距离的缩短降低转移成本,实现生产和消费在地理空间上的分离,可以促进区域间的贸易、交流与合作,进而加大发达地区对周边地区的正向溢出效应,周边地区能够以更低成本获得发达地区的先进技术,增强自主创新能力,提升发展速度,缩小区域间经济差异;另一方面,邻近地区对外贸易距离的缩短会增强该邻近地区对生产要素的吸引力,进而使得生产能力提高,促进经济发展,拉大与本地经济发展水平的差距。

分割(Division)是指国家或者地区间限制要素、信息和技术等流动的因素,即阻碍地区经济活动的壁垒(齐美虎等,2012)。分割也会从两个方面影响区域经济发展差异。在经济发展初期,分割有助于保护本地产业发展,促进本地区经济增长,进而缩小地区间的经济差距;但从长远来看,分割会拉大经济差距,这是因为随着经济的增长,分割会使地区信息闭塞,忽视与外界的交流,创新不足,进而导致区域间差异增大。

世界银行的《世界发展报告——重塑世界经济地理》认为,提高密度、缩短距离、降低分割程度是国家或地区经济发展的基本条件。下文用经济密度、经济距离和经济分割表示这三要素。

三、中国大陆省域经济空间相关性分析

本文选取人均GDP(per capita gdp)来衡量经济发展水平,省域人均GDP的空间相互作用结构由权重矩阵 W 来刻画,这是Moran

引入的一个概念(Harry Kelejian, 2017)。 $n \times n$ 空间权重矩阵 $W = \begin{bmatrix} w_{11} & \cdots & w_{1n} \\ \vdots & \ddots & \vdots \\ w_{n1} & \cdots & w_{nn} \end{bmatrix}$ 是分析空间相关性的重要组成部分, w_{ij} 为一阶空间权重矩阵 W 第 i 行第 j 列的元素(用来度量地区 i 与地区 j 之间的空间相邻性), 如果地区 j 与 i 相邻, 则 $w_{ij} > 0$, 否则 $w_{ij} = 0$ 。对于多边形的区域单元, 空间权重的构造通常基于地理上相邻(contiguity)的概念, 本研究采用应用最广泛的Rook邻接标准定义空间权重矩阵(Rook Contiguity Weights), 即若省份 i 和省份 j 有公共边界, 则两省为邻接单元, 空间权重矩阵中的元素 $w_{ij} = 1$, 否则 $w_{ij} = 0$ 。按照惯例地区自身不作为自身的一阶邻接单元, $w_{ii} = 0$ 。在实际应用中, 通常将空间权重矩阵进行归一化处理, 即将矩阵中的

每个元素除以其所在行的元素之和,使每行的元素之和为 1。对于归一化后的 $n \times n$ 权重矩阵 W , 定义 W 与 $n \times 1$ 向量 y 的乘积 Wy 为 y 的空间滞后向量 (Spatially Lagged vector), 对于观测值 y_i 对应的空间滞后项则为该变量所有一阶相邻地区观测值的加权平均数, 即

$$[Wy]_i = W_{i1}y_1 + W_{i2}y_2 + \cdots + W_{in}y_n \quad (1)$$

利用 y 和 Wy 绘制的莫兰散点图 (Moran scatterplots) 可以直观体现变量观测值与相邻地区观测值空间相关的程度。在本研究中, 以 31 个省、直辖市、自治区 2019 年国民经济和社会发展统计公报发布的 GDP 数据进行标准化后的得分 z 为观测向量, 利用软件 stata15 绘制 z 和 Wz 的莫兰散点图 (见图 1), 以此观察各省人均 GDP 与周边相邻省份人均 GDP 的空间相关性。

各省人均 GDP 通过标准化变换后的 z 和 Wz 落在以零为中心的四个象限内, 分别代表自身和相邻省份经济发展水平与全国平均水平比较的不同关系。其中位于第 I 象限 (高-高) 的单元有北京、天津、上海、江苏、浙江、福建、山东等 7 个省市, 代表这 7 个省市的经济发展水平高于所有省份平均水平, 同时其周边相邻省份的平均发展水平也高于全国平均值, 这些省市主要集中在东部沿海和北京周边, 可以看出富裕地区和富裕地区明显地聚集在一起, 经济聚集程度较高; 有 17 个省市落在第 III 象限 (低-低), 代表这些省市的经济发展水平低于所有省份平均水平, 同时其周边相邻省份的平均发展水平也低于全国平均值, 这些省份也连成一片, 主要为东北和中西部地区; 广东、湖北、重庆 3 个省市落在第 IV 象限 (高-低), 代表这 3 个地区的经济发展水平高于所有省份平均水平, 但其周边相邻省份的平均发展水平仍低于全国平均值; 位于第 II 象限 (低-高) 的地区有河北、安徽等 4 个省份, 自身经济发展水平低于全国平均水平, 但相邻省份经济发展水平普遍较高。可以看出大部分省市落在 I、III 象限, 显示省域经济与周边经济发展水平存在较高的空间关联性, 并且落在第三象限的省份占中国大陆省份数量的一半以上, 说明中国大陆地区较多省份经济发展低于平均水平, 可以看出中国大陆省域经济发展呈两极分化状态, 总体空间的不均衡分布仍然明显。

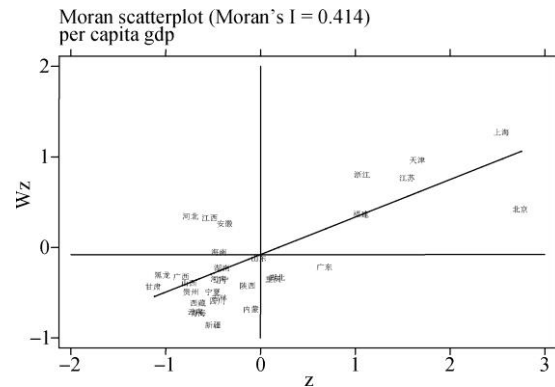


图 1 2019 年中国大陆省域人均 GDP 的 Moran 散点图

省域人均 GDP 的空间相关性可以通过统计量莫兰系数 (Moran's I) 更严格地进行空间自相关检验, 即检验一个变量在一个位置的观测值是否独立于该变量在邻近位置的值。正空间自相关表明相似的值在空间上彼此接近或聚类; 负空间自相关表明邻近的值是不相似的, 或者说相似的值是分散的; 无空间自相关表示空间模式是随机的。

$$I = \frac{n}{\sum_{i=1}^n \sum_{j=1}^n w_{ij}} \cdot$$

莫兰系数定义为:

$$\frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (1)$$

其中 y_i 是所考虑地区 i 的变量值, 在此为人均 GDP, $y - \bar{y}$ 表示变量在所有地区的平均值, n 是地区的总数。Moran 系数采用了自相关系数的经典形式, 即计算 y 和 Wy 的相关系数。莫兰系数 I 的标准化统计量 I_z 服从渐近标准正态分布, 则变量空间相关性的统计推论正是基于 Moran's I 的渐近标准正态分布进行的 (Luc Anselin, 2014)。

图 1 利用 stata15 绘制莫兰散点图的同时报告了莫兰系数值 $Moran's I = 0.414$, 在 0.1% 的水平下显著, 说明中国大陆人均 GDP 存在显著的空间正相关性, 省域经济空间依赖度高, 经济发展水平接近的省份呈现一定程度的空间聚集, 从而导致区域经济差异的产生。

四、区域经济影响因素分析的模型方法

以上对中国大陆省域经济空间相关性的分析证实了经济发展存在一定的空间依赖, 结合图 1 可以看出处于第 I、IV 象限的高收入省份集中在东部沿海一带和首都经济圈以及重庆、湖北二地, 呈较明显的聚集态势, 但其周边相邻内陆省份经济处于较低水平, 特别是江西和安徽两省被高收入地区所包围, 但其人均 GDP 仍低于平均水平。因此, 缩小区域经济发展差异, 需要正确解释经济增长的过程, 找到主要影响因素, 才能普遍提高经济发展水平。同时, 要提高高收入地区的辐射效应, 如落在第 II 象限的四个低收入省份都与 I、IV 象限的高收入省份相邻, 如何更快地带动这四个地区的经济发展, 形成更大的经济辐射效应, 再进一步拉动第 III 象限更多省份经济的增长, 都需要建立合适的模型正确估计经济增长因素产生的影响效应。

对空间关系进行建模常常会忽略潜在因素的空间影响, 而区域经济空间相关的原因之一正是受潜在的具有空间相关性的因素的影响, 可能会有一些地理上的区域共享的未被观察到的因素导致空间溢出, 如技术革新作为经济增长的主要因素, 不少研究表明技术革新呈现出空间关联性, 一个地区的技术创新变化会对其他地区的技术水平产生影响从而对经济产出产生空间溢出效应。但技术创新水平很难准确地衡量并被纳入模型中, 同时与资本、劳动力等其他可观测的影响因素不无关系, 这时简单地用 OLS 估计线性回归模型就会产生遗漏变量偏误从而影响我们对资本、劳动等因素对经济增长影响效应的正确推断。如何用合适的模型来减轻当存在与解释变量相关且自身有空间相关性的遗漏变量时产生的估计偏误, 是进行实证分析的前提。

为了解决上述类型的问题, Anselin (1988) 推导了当真实的数据生成过程反映了自变量与扰动项相关且扰动项存在空间相关性时的空间杜宾模型 SDM, 式 (3) 给出了 SDM 的一般表达式:

$$y = \rho Wy + X\beta + WX\theta + \epsilon \quad (3)$$

式 (3) 中, y 为 $n \times 1$ 因变量向量, X 为 $n \times k$ 解释变量矩阵, W 为 $n \times n$ 权重矩阵, ρ 为待估空间自回归系数, 描述了观测样本中空间依赖的平均强度, 如果 ρ 显著为零表示未被观测到的潜变量不存在空间依赖性, β 、 θ 为 $k \times 1$ 待估系数向量, 可以通过极大似然估计法得到。

空间杜宾模型 (SDM) 通过在传统的线性回归模型中引入被解释变量的空间滞后项 Wy 来刻画空间自回归过程, 同时通过引入解释变量的空间滞后项 Wx 考虑了样本中单个地区解释变量的变化带来的溢出效应, 这从式 (3) 隐含的数据生成过程 (DGP) 可以得出, 由式 (3) 可得: $y - \rho Wy = X\beta + WX\theta + \epsilon$, 则有:

$$y = (I_n - \rho W)^{-1} X (\beta + W\theta) + (I_n - \rho W)^{-1} \varepsilon = \\ (I_n + \rho W + \rho^2 W^2 + \rho^3 W^3 + \dots) X (\beta + W\theta) + \\ (I_n - \rho W)^{-1} \varepsilon = X (\beta + W\theta) + (\rho W + \rho^2 W^2 + \rho^3 W^3 + \dots) \\ X (\beta + W\theta) + (I_n - \rho W)^{-1} \varepsilon \quad (4)$$

式(4)中, $(\beta + W\theta)$ 为系数向量, W^2 代表二阶空间权重矩阵, 表示与一阶相邻地区相邻的地区的关系, W^3 为三阶空间权重矩阵, 以此类推。| ρ | < 1, 因此有:

$$(I_n - \rho W)^{-1} = I_n + \rho W + \rho^2 W^2 + \rho^3 W^3 + \dots$$

从式(4)中可以看出, y 在任意位置 i 上的值不仅是 X 在该位置上的值的函数($X(\beta + W\theta)$), 还是一阶相邻位置上 X 的函数($\rho WX(\beta + W\theta)$), 以及通过 $\rho^2 W^2$ 、 $\rho^3 W^3 \dots$ 与高阶相邻地区的 X 值产生联系。换句话说, 任何地区的被解释变量值都取决于所有位置的解释变量, 同时任何位置的解释变量的变动都会对所有地区被解释变量产生影响, 因 | ρ | < 1, 高阶空间相关地区解释变量溢出效应随着距离而呈几何级数衰减, 衰减的程度取决于空间自回归系数的大小。

综上, 本文拟利用式(3)的空间杜宾模型 SDM 来进行区域经济影响因素研究。因变量向量 y 为地区人均收入水平, 解释变量矩阵 X 为影响经济增长的特征, 包括人力和物质资本以及新经济地理视角下区域经济影响因素, 则收入水平的区域差异被建模为依赖于由空间滞后向量 Wy 体现的邻近地区的收入水平, 以及依赖于由 WX 所表示的邻近地区的变量特征。

五、基于 SDM 模型的变量影响效应分析

由于空间杜宾模型 (SDM) 利用了地区观测值之间复杂的依赖结构, 因此, 参数估计量包含大量关于地区观测值之间相互关系的信息, 与任何给定的解释变量相关的单个地区观测值的变化将直接影响该地区本身, 并可能间接影响所有其他地区。为了明确 SDM 模型中变量的数量关系, 将式(3)的 SDM 模型 $y = \rho Wy + X\beta + WX\theta + \varepsilon$ 改写成为式(5)的表达形式:

$$y = (I_n - \rho W)^{-1} X (\beta + W\theta) + (I_n - \rho W)^{-1} \varepsilon \quad (5)$$

式(5)中, $n \times 1$ 被解释变量向量 $y = (y_1, y_2, \dots, y_n)'$, $X = (x_{11}, x_{12}, \dots, x_{1r}, \dots, x_{1k})$ 为 $n \times k$ 解释变量矩阵, 其中 $x_r = (x_{1r}, x_{2r}, \dots, x_{nr})'$ 为第 r 个解释变量向量, $\beta = (\beta_1, \beta_2, \dots, \beta_r, \dots, \beta_k)'$ 和 $\theta = (\theta_1, \theta_2, \dots, \theta_r, \dots, \theta_k)'$ 均为 $k \times 1$ 参数向量, W 为 $n \times n$ 权重矩阵, I_n 为单位矩阵。则式(5)可表示成式(6):

$$y = \sum_{r=1}^k (I_n - \rho W)^{-1} (I_n \beta_r + W \theta_r) x_r + (I_n - \rho W)^{-1} \varepsilon \quad (6)$$

为了便于简化表达式, 记: $V(W) = (I_n - \rho W)^{-1} = I_n + \rho W + \rho^2 W^2 + \rho^3 W^3 + \dots$, $S_r(W) = V(W) (I_n \beta_r + W \theta_r)$

$S_r(W)$ 为第 r 个解释变量对应的系数矩阵, 则式(6)可简写为:

$$y = \sum_{r=1}^k S_r(W) x_r + V(W) \epsilon \quad (7)$$

展开得：

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \sum_{r=1}^k \begin{pmatrix} S_r(W)_{11} & S_r(W)_{12} & \dots & S_r(W)_{1n} \\ S_r(W)_{21} & S_r(W)_{22} & \dots & S_r(W)_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ S_r(W)_{n1} & S_r(W)_{n2} & \dots & S_r(W)_{nn} \end{pmatrix} \begin{pmatrix} x_{1r} \\ x_{2r} \\ \vdots \\ x_{nr} \end{pmatrix} + \begin{pmatrix} V(W)_{11} & V(W)_{12} & \dots & V(W)_{1n} \\ V(W)_{21} & V(W)_{22} & \dots & V(W)_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ V(W)_{n1} & V(W)_{n2} & \dots & V(W)_{nn} \end{pmatrix} \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix} \quad (8)$$

式(8)中 $S_r(W)_{ij}$ 和 $V(W)_{ij}$ 分别是矩阵 $S_r(W)$ 和 $V(W)$ 第 i 行第 j 列的元素, 反映了地区 i 的被解释变量和地区 j 的第 r 个解释变量及误差项之间的数量关系, 有:

$$y_i = \sum_{r=1}^k [S_r(W)_{i1} x_{1r} + S_r(W)_{i2} x_{2r} \dots + S_r(W)_{ij} x_{jr} \dots + S_r(W)_{in} x_{nr}] + \sum_{j=1}^n V(W)_{ij} \epsilon_j \quad (9)$$

从式(9)可以更清楚地看出, 解释变量 x_r 在任何一个地区 j 的取值 x_{jr} 的变化都会通过 $S_r(W)_{ij}$ 影响地区 i 的被解释变量 y_i 的取值。可以看出, $\partial y_i / \partial x_{ir} = S_r(W)_{ii}$ 为 $n \times n$ 矩阵 $S_r(W)$ 的主对角线元素, 表示地区 i 解释变量 x_r 的变化对本地的被解释变量 y_i 所产生的所有影响, $S_r(W)_{ii} \neq 0$, 是广义的直接影响, 这种影响包括反馈回路的影响, 即单个地区 i 观测值的变化将直接影响该地区本身, 并可能间接影响所有其它地区, 而其它地区又反过来影响地区 i , 这是因为高阶空间相关定义中地区 i 被认为是其相邻地区的邻居, 所以通过邻近区域的影响会对地区 i 本身产生反馈影响。 $S_r(W)$ 非对角元素则代表地区 j 变量 x_r 的变化对其它地区的 y_i 所产生的间接影响, $\partial y_i / \partial x_{jr} = S_r(W)_{ij}$, $i \neq j$ 。

从 $S_r(W)$ 中的元素可知, 解释变量变化的影响在所有地区 i 中都是不同的, 如何衡量解释变量的影响效应, LeSage 和 Pace (2009) 提出了基于对矩阵 $S_r(W)$ 的行(或列)的影响系数的总和, 对所有地区取平均值来体现影响效应的一般水平。 x_r 的变化对 y 直接影响的一般水平, 即地区 i 的解释变量 x_r (我们称其为 x_{ir}) 的变化对本地 y_i 的影响, 可以通过计算 $S_r(W)$ 主对角线元素 $S_r(W)_{ii}$ 的平均值

来代表, 称作 x_r 的平均直接影响效应 (Average Direct Effect), 表示为:

$$Effect_{direct} = \frac{\sum_{i=1}^n S_r(W)_{ii}}{n} \quad (10)$$

$S_r(W)$ 的第 j 列元素求和表示 x_r 在地区 j 观测值的变化对所有地区 y 产生的总影响, 各列和的平均数则表示单个地区 x_r 变化对所有地区 y 的平均总影响效应 (Average Total Effect):

$$Effect_{total} = \frac{\sum_{j=1}^n \sum_{i=1}^n S_r(W)_{ij}}{n} \quad (11)$$

则单个地区 x_r 变化产生的平均间接影响效应 (Average Indirect Effect) 为平均总影响与平均直接影响之差:

$$Effect_{indirect} = Effect_{total} - Effect_{direct} \quad (12)$$

平均间接影响效应也是 x_r 在单个地区观测值的变化对其他所有地区产生的溢出效应, 是区域经济分析中一个重要的统计量。

以往有一些研究误解了关于 SDM 模型式 (3) 中回归系数的含义, 将 β 作为直接效应的估值, 将解释变量空间滞后项 (WX) 的系数 θ 解释为间接影响效应, 并通过该系数的显著性检验来得出是否存在空间溢出的结论。从以上分析可以看出, 变量 x_r 的直接效应和间接效应均取决于矩阵 $S_r(W)$ 所包含的变量间复杂的关系, 除了 β_r 和 θ_r , 还受空间自回归系数 ρ 、空间权重矩阵 W 以及 $\rho^2 W^2$ 、 $\rho^3 W^3 \dots$ 所代表的高阶空间相关关系的影响。利用上述公式对 x_r 的影响效应进行显著性检验, 需要知道构成这三种影响效应的统计量 ρ 、 W 、 β_r 和 θ_r 的分布, 可以利用高效的快速模拟方法从最大似然多元正态分布中构造大量的模拟参数来得到 ρ 、 W 、 β_r 和 θ_r 的经验分布。另外, 用贝叶斯马尔可夫链蒙特卡洛 (MCMC) 估计方法来估计影响效应的离散程度, 也可以对直接影响效应和总影响效应产生有效的推断 (LeSage 和 Pace, 2009)。

我们利用 SDM 构建区域经济增长模型, 主要目的就是为了正确掌握所研究的影响因素对经济增长的直接效应和溢出效应, 从而了解区域经济差异产生的原因, 为采取更有效的促进区域经济增长的对策建议提供实证依据。

六、中国大陆省域经济差异影响因素的实证分析

1. 变量说明和数据来源

本文在索洛增长模型包含物资资本和人力资本两种要素的基础上, 结合新经济地理视角下影响区域经济增长的三大因素对中国大陆省域经济差异的影响因素进行实证分析。为了更准确估计传统生产要素和新经济地理因素对地区经济差异的影响, 本文还引入地方政府支出和制度因素两个控制变量。虽然政府加大支出力度会引起私人消费或投资降低, 即产生“挤出效应”, 但是只要地方未达到充分就业前, 扩张性的财政政策都会通过刺激劳动供给或吸引追加投资刺激地方经济增长。另外, 新制度经济学认为任何经济增长过程都是在一定的制度环境和制度安排下进行的, 没有制度背景, 经济增长不可能独立存在。合理的制度通过保护人们的努力, 为专业化发展提供机会, 允许市场有最大限度的活动自由来促进经济增长并通过发挥榜样作用, 促进邻近地区

经济发展,进而缩小地区间经济发展差距。基于以上分析,因变量为省域经济发展水平,选取资本投入量、人力资本存量、经济密度、经济距离、经济分割、地方政府支出和制度因素作为自变量进行区域经济差异影响因素分析。

(1)被解释变量(Y)为省域经济发展水平,用各省实际人均国内生产总值(per capita gdp)来表示。

(2)解释变量。人力资本存量(H):人力资本存量采用人均受教育年限法来衡量。受教育程度划分为5种:未上过学、小学、初中、高中、大专及以上,并且分别以0年、6年、9年、12年、16年表示不同受教育程度人口的教育年限。人均受教育年限法是指按相应人口数为权重加权平均推算的地区人均受教育年限。人均受教育年限越长,地区人力资本越高。

资本投入量(K):由于无法直接获得各省市历年的物质资本存量数据的统计数据,本文采用永续盘存法估算物质资本存量。估算公式为 $K_{i(t+1)} = (1 - \sigma) K_{i,t} + I_{i(t+1)} / P_{i(t+1)} \times 100$,其中*i*,*t*分别代表第*i*个省份和第*t*年,*K*表示物质资本存量,*I*为当年的名义总投资,*P*为固定资产投资价格指数, σ 为物质资本经济折旧率。名义总投资指标采用当年固定资产投资总额,物质资本经济折旧率取张军等(2004)计算得到的 σ 为9.6%,并以其估算得到的1990年中国各省市物质资本存量作为初始资本存量,同时选取1990年为基期的固定资产投资价格指数代入上述估算公式,得到2003-2016年各省市物质资本存量估算值,并以此作为资本投入量数据。

经济密度(density):根据经济密度的内涵,并参考已有的研究方法(张改素等,2013),从人口密度、地均经济产出两个层面来反映。人口密度使用单位面积人口数衡量,体现地区经济发展对产业、人口聚集的效果,地均经济产出选取单位面积上二、三产业增加值之和来衡量,反映单位面积建成区上的经济强度,二者都是经济密度最终的表现,对这两个指标赋予相同的权重,由于一地的人口和经济活动的集中会促使该地区经济密度的提高,因此综合后的数值越大,说明经济密度越高。

经济距离(distance):选择省内铁路里程数和公路里程数之和与各省面积之比,以及每百人使用因特网的人数,再对二者赋予相同的权重,计算得出的结果作为衡量省内经济距离的指标。铁路里程数与公路里程数在一定程度上代表一地的交通基础设施建设水平,交通基础设施建设越发达,其商品和服务穿越地理空间越容易,所耗费的成本就越低,根据定义,经济距离越短;每百人使用因特网的人数代表了地方通信水平的高低,使用因特网的人数越高,通信水平越高,信息流通越通畅,区域间的交流与合作越便利,经济距离越短。

经济分割(division):选取以当年折算为人民币的外商直接投资占GDP的比重作为经济分割指标,外商投资水平衡量国际层面的分割程度,外商直接投资数额占比越高,意味着该地区与外界、特别是其它国家的衔接更为密切,也就是说投资等经济活动进入本区域市场遭遇的壁垒和受到的阻力较小,即分割程度较低。

地方政府支出(czzc):本文选择地方财政支出变量作为控制变量之一。

制度因素(institution):通过对市场化指数和城镇化水平赋予相同的权重计算得出的数据作为衡量制度变量的指标。市场化指数越高,体现我国市场经济体制改革成效越明显,市场的主导作用越强,市场交易成本越低,市场经济的运行效率越高,制度所释放的红利越强,则制度安排越合理。城镇化水平越高,说明我国户籍制度改革、教育改革等方面收获的成效越高,农村人口涌入城市生产部门,释放人口红利,制度改革越合理。本文采用学者已计算得出的市场化指数(王小鲁等,2019),该指数综合了政府与市场关系、中介组织与法律、市场条件(产品市场、要素市场)、非国有经济等多个方面指标。选择各省区市城镇人口占年末常住总人口的比重反映城镇化水平。

研究的样本地区包括中国大陆31个省(自治区、直辖市),数据来源于2003-2016年《中国统计年鉴》《中国城市统计年鉴》以及各省、自治区、直辖市国民经济和社会发展统计公报。各变量取对数后的描述性统计见表1。

表1 变量描述性统计

变量 (Variable)	观测数 (Obs)	均值 (Mean)	标准差 (Std. Dev.)	(最小值) Min	(最大值) Max
LN(percapitagdp)	434	10.157	0.708	8.322	11.68
LN(H)	434	2.128	0.158	1.319	2.51
LN(K)	434	9.068	1.094	5.443	11.347
LN(density)	434	2.323	0.567	0.692	3.869
LN(distance)	434	7.821	1.023	5.144	9.305
LN(division)	434	11.894	1.485	6.38	14.459
LN(institution)	434	3.285	0.304	2.162	3.905
LN(czzc)	434	16.636	0.971	13.868	18.717

2. 实证模型

根据前文的影响因素分析, 本文基于传统柯布-道格拉斯生产函数得出扩展的经济增长函数形式:

$$y = A H^{\beta_1} K^{\beta_2} \text{density}^{\beta_3} \text{distance}^{\beta_4} \text{division}^{\beta_5} \text{institution}^{\beta_6} \text{czzc}^{\beta_7} \quad (3)$$

在此基础上构建式 (3) 的空间杜宾模型 (SDM) $y = \rho W y + X \beta + W X \theta + \varepsilon$ 对中国大陆省域经济差异的影响因素进行实证分析。其中: $X = (\ln H, \ln K, \ln \text{density}, \ln \text{distance}, \ln \text{division}, \ln \text{institution}, \ln \text{czzc})$; $\beta = (\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6, \beta_7)$; $\theta = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6, \theta_7)$, w_{ij} 为空间权重矩阵 W 的 (i, j) 元素, 仍然采用前文 Rook 邻接定义的空间权重矩阵, 若两省 i 和 j 有公共边界, 则 $w_{ij}=1$, 否则 $w_{ij}=0$, 同时将矩阵中的每个元素除以其所在行的元素的和, 使得每行的元素之和为 1。

本文使用省域面板数据, 在经济增长过程中, 很难假设个体效应与模型中的解释变量无关, 因此本文利用固定效应面板的空间杜宾模型进行实证分析, 得到实证分析模型如下:

$$\begin{aligned} \ln Y_{it} = & \alpha_i + \beta_1 \ln K_{it} + \beta_2 \ln H_{it} + \beta_3 \ln \text{density}_{it} + \\ & \beta_4 \ln \text{distance}_{it} + \beta_5 \ln \text{division}_{it} + \beta_6 \ln \text{institution}_{it} + \\ & \beta_7 \ln \text{czzc}_{it} + \rho \sum_{j=1}^n w_{ij} Y_{jt} + \theta_1 \sum_{j=1}^n w_{ij} \ln K_{jt} + \theta_2 \sum_{j=1}^n w_{ij} \ln H_{jt} + \\ & \theta_3 \sum_{j=1}^n w_{ij} \ln \text{density}_{jt} + \theta_4 \sum_{j=1}^n w_{ij} \ln \text{distance}_{jt} + \\ & \theta_5 \sum_{j=1}^n w_{ij} \ln \text{division}_{jt} + \theta_6 \sum_{j=1}^n w_{ij} \ln \text{institution}_{jt} + \\ & \theta_7 \sum_{j=1}^n w_{ij} \ln \text{czzc}_{jt} + \varepsilon_{it} \end{aligned} \quad (4)$$

式(14)中, $\alpha_i = \ln A_i$ 为个体固定效应, 是资本投入量 K 、人力资本存量 H 、经济密度、经济距离、经济分割、地方财政支出和制度变量等因素之外的解释经济增长的余量, 可以作为各省潜在的技术创新水平的度量。

3. 实证结果与分析

运用 Stata15 软件, 利用 2003–2016 年中国大陆地区 31 个省份的数据对模型式(14)进行估计, 回归结果见表 2。

表 2 SDM 模型的回归系数估计结果

变量	回归系数	t 统计量
$\ln H$	0.2727	1.5361
$\ln K$	0.3109***	7.0066
$\ln \text{density}$	0.1383***	3.6279
$\ln \text{distance}$	0.0928***	3.7423
$\ln \text{division}$	0.0021	0.1766
$\ln \text{institution}$	0.3461***	3.0071
$\ln \text{czzc}$	0.2853***	8.5739
ρ	0.5137***	7.7868
$W \cdot \ln H$	-0.1082	-0.6233

(续表)

变量	回归系数	t 统计量
$W \cdot \ln K$	-0.2508***	-4.2120
$W \cdot \ln \text{density}$	0.0448	1.0852
$W \cdot \ln \text{distance}$	-0.0640**	-2.1058
$W \cdot \ln \text{division}$	0.0224	0.6488
$W \cdot \ln \text{institution}$	-0.0550	-0.2494
$W \cdot \ln \text{czzc}$	-0.1733***	-2.6273
固定效应均值 α_i	0.3183	
对数似然函数 Log-L	740.7308	

空间杜宾模型回归结果表明, 空间自回归系数 ρ 在 1% 的水平下显著为正, 进一步验证了前文所述我国省域之间经济发展水平

存在正空间相关性的结论。关于变量的回归系数,由于包含空间滞后变量,参数的解释变得复杂,无法直接从模型回归系数直接推断变量的影响效应。需要根据前文中关于 SDM 模型变量影响效应的分析方法,正确估计各变量对人均 GDP 产生影响的总效应、直接效应和间接效应。运用 Stata15 估计的各变量影响效应系数值及检验统计量见表 3。

表 3 中直接效应表示单个省某个解释变量的变化对本省人均 GDP 所产生的影响,总效应表示该省某个解释变量的变化对所有省份人均 GDP 产生的总影响,某个变量变化对本省以外地区人均 GDP 产生的间接效应等于总效应与直接效应之差,用来衡量该变量空间溢出的程度。

比较变量的回归系数和影响效应系数,可以看到表 3 中变量的直接效应系数值与表 2 中报告的变量自身的回归系数估计值,二者并不相等,如变量 $\ln H$ 、 $\ln density$ 和 $\ln institution$ 对应的直接效应系数均大于回归系数,表示这三个变量对本地人均 GDP 产生的影响包含通过间接影响其它地区而其它地区又反过来影响本地区所产生的反馈效应,是变量变化最终对本地产生的总影响,但由于回归系数与直接效应系数值之间的差异非常小,可见反馈效应不大,模型中变量自身的回归系数基本体现了变量对本地经济增长的直接影响。再比较 SDM 模型中滞后变量的回归系数与间接效应系数估计值,二者有很大的差异。例如,表 3 中 $\ln density$ 的间接影响效应是 0.2167,表 2 中报告的空间滞后变量 $W \cdot \ln density$ 的系数估计值为 0.0448,且不显著,如果我们错误地将关于经济密度空间滞后项 $W \cdot \ln density$ 的回归系数看作是间接影响的反映,就会得出经济密度变量对区域经济发展没有带动作用的推断,然而,真正的影响效应估计指出了变量 $\ln density$ 除对本地经济产生显著的直接影响外,总影响效应也很显著,说明其产生了空间溢出效应从而对区域经济增长产生积极的影响。

表 3 SDM 模型各变量影响效应系数估计结果

变量	直接效应	总效应	间接效应
$\ln H$	0.2850 (1.5999)	0.3531 (1.3636)	0.0681 (0.3042)
$\ln K$	0.2948*** (6.7562)	0.1229 (1.1335)	-0.1718* (-1.8258)
$\ln density$	0.1599*** (4.1063)	0.3766*** (3.9929)	0.2167*** (3.0707)
$\ln distance$	0.0892*** (3.9757)	0.0606 (1.2735)	-0.0286 (-0.5907)
$\ln division$	0.0051 (0.3441)	0.0470 (0.5951)	0.0419 (0.6183)
$\ln institution$	0.3679*** (3.2444)	0.6023 (1.4012)	0.2344 (0.6095)
$\ln czzc$	0.2801*** (8.6986)	0.2308* (1.8598)	-0.0493 (-0.4250)

进一步比较表 3 中各变量的影响效应大小,先看直接效应系数,除了 $\ln H$ 和 $\ln division$ 不显著外,其余变量的直接效应均显著为正,说明这些因素对本地经济增长均有促进作用。其中直接影响效应最大的三个变量分别是 $\ln K$ 、 $\ln institution$ 和 $\ln czzc$,说明资本投入量、制度因素和地方政府财政支出是促进本省经济增长的最主要变量,也是导致地区间经济发展水平差异的最主要原因。再看总效应和对应的间接效应, $\ln K$ 、 $\ln distance$ 和 $\ln czzc$ 对经济增长产生的总效应小于直接效应,相应的间接效应为负,说明二者在促进本省经济增长的同时对周边地区并没有带动作用,反而在一定程度上抑制或削弱周边地区经济增长的潜力。而变量 $\ln density$ 的总效应系数值显著为正,且明显大于其它变量的总影响效应,对应的间接影响效应系数也较大,说明经济密度所产生的空间溢出效应明显,对本省和周边省份的经济增长均有促进作用。经济密度变量 $\ln density$ 的直接效应系数 0.1599 在所有变量中并不算大,仅为 $\ln K$ 、 $\ln institution$ 和 $\ln czzc$ 直接效应系数的 50%左右,但其间接效应为 0.2167,显著大于直接效应,是所有变量中间接效应最显著的一个,说明经济密度变量的变化对周边省份经济增长产生的累计影响较大,空间溢出效应最明显,

是带动区域经济增长的重要因素。正是因为某些影响因素存在空间溢出效应,才会使相邻省份经济表现出空间关联从而产生一定程度的聚集,同时溢出效应随着空间距离增加而递减,发达地区经济辐射的带动作用有限,也能解释区域经济差异出现的原因。

七、研究结论与建议

以上对中国大陆省域经济空间相关性分析表明经济增长存在显著的空间依赖,区域间经济发展两极分化,区域经济差异明显。实证分析的结果显示,物质资本存量、制度相关的市场化和城镇化水平及地方政府财政支出是促进各省经济增长直接效应最大的主要因素。在其他条件不变的情况下,各省市物质资本存量、制度因素和财政支出是形成各省经济差距的主要原因。实证结果也证实了新经济地理因素对本地经济增长存在显著影响,各省的经济密度、经济距离变量对经济增长有正向的促进作用,也是造成中国省域经济非均衡发展的直接原因。同时实证结果表明经济密度具有明显的空间溢出效应,在显著促进本地区经济发展的同时对周边省份经济的发展产生积极的间接影响,促进区域经济增长的总效应显著,对缩小区域经济差异具有积极作用。

因此,为提高中国大陆整体经济发展水平,缩小区域经济差异,对于广大中西部经济发展落后的地区,要加大发挥直接影响效应大的因素的作用,大力提高本地的经济发展水平。东部沿海发达地区要充分发挥溢出效应大的影响因素的作用,进一步带动周边省份经济增长成为经济发达地区,形成更大的溢出效应,逐步扩大发达经济带,缩小区域经济发展差距。为此,特提出以下几点建议:

落后省市要加大财政支出力度,完善财政支出结构。财政支出的直接效应水平显著为正,说明财政支出的经济增长效应明显。所以,今后应努力加大财政支出力度从而提升本地经济发展水平;此外,还应注意,在财政支出结构中,不同种类的财政支出对经济增长的作用力度及作用方式可能存在不同,因此应优化财政支出结构和进度,以期更好地服务于长期经济增长,缩小与发达地区的经济差异。

形成城市群并建立适当的经济中心点。要实现经济落后区域的较快发展,应该集中整合财力与物力,向某个中心城市集中倾斜,给予较强的支持力度,提高该中心城市的经济密度,让其发展成为整个区域的核心。通过集聚生产要素显著促进该地经济发展,并对周边地区产生辐射作用带动整个城市群的协调发展。东部沿海发达地区要稳定持续发展二三产业,充分发挥高经济密度的溢出效应,辐射带动中西部地区的经济发展。

重视制度因素在经济增长中的作用,构建完善的制度环境。中西部地区要积极推进城镇化进程,但也不要盲目追求城镇化数量以及一味追求成为特大城市。要高度重视推进新型城镇化,促进人才在城镇的集聚,通过城镇化释放的人口红利充分发挥城镇在促进区域经济发展中的作用。中西部地区也要不断完善市场经济体制,牢固市场机制在区域资源配置中的基础性地位。加快国有经济改革步伐,同时为非国有经济提供充足的自由发展空间,从而提高经济增长的速度,实现区域经济协调发展。

加强落后省市人才培育与增加资本投入,确立人力资本发展战略。虽然回归结果显示人力资本变量的回归系数与影响效应系数都不显著,这是由于本文采用当地人均受教育年限衡量人力资本水平,但由于人才流动的原因,参与计算受教育年限的人口可能并没有在当地从事生产工作,因此并没有为当地的经济增长作出贡献。但人力资本对于经济增长的作用毋庸置疑,随着我国正逐步从劳动密集型发展模式向知识密集型模式过渡,人口素质对本地经济增长的促进作用日益凸显。同时地区人力资本的提高为该地区创新产业的发展提供了施展空间,有助于该城市创新能力的提升和创新的扩散,进而带动周边地区的经济发展。因此,在加大地方物质资本投入的前提条件下,完善公共教育服务,培养高素质人才,并出台相关人才引进政策,建立健全人才制度体系,改善当地的劳动力质量,从而提升本地经济发展水平,缩小与发达地区的经济差异。

最后,值得一提的是,内生性问题是构建计量模型进行经济变量之间关系分析时必须考虑的重点难点,因此本文在构建模型时利用面板数据进行固定效应的SDM模型估计,以解决解释变量与不可观测的个体效应相关引起的内生性问题,此外,SDM模型本身也能较有效克服存在相关性遗漏变量时估计偏误的问题,但本文的模型方法是否完全解决了内生性问题,是后续研究中非常值

得进一步探讨的研究课题。

参考文献:

- [1]. Anselin, L. Spatial Econometrics: Methods and Models. Dordrecht: Kluwer Academic Publishers, 1988.
- [2]. Anselin, Luc. Modern Spatial Econometrics in Practice. GeoDa Press, 2014.
- [3]. Kelejian, Harry. Spatial Econometrics. Academic Press, 2017.
- [4]. LeSage, James and R. Kelly Pace Introduction to Spatial Econometrics. CRC Press, 2009.
- [5]. 程名望、贾晓佳、仇焕广:《中国经济增长(1978—2015):灵感还是汗水》,《经济研究》2019年第7期。
- [6]. 刘青春、王铮:《中国区域经济差异形成的三次地理要素》,《地理研究》2009年第2期。
- [7]. 齐美虎、张林:《云南3D经济地理特征及其影响》,《思想战线》2012年第6期。
- [8]. 世界银行编著,胡光宇等译:《2009年世界发展报告:重塑世界经济地理》,清华大学出版社2009年版。
- [9]. 吴金燕、滕建州:《经济金融化对实体经济影响的区域差异研究——基于省级面板数据的空间计量研究》,《经济问题探索》2020年第7期。
- [10]. 王小鲁、樊纲:《中国地区差距的变动趋势和影响因素》,《经济研究》2004年第1期。
- [11]. 王小鲁、樊纲、胡李鹏著:《中国分省份市场化指数报告(2018)》,社会科学文献出版社2019年版。
- [12]. 张改素、丁志伟、王发曾:《我国中部地区经济密度的时空分异研究》,《经济地理》2013年第5期。
- [13]. 张鹏飞、李国强等:《区域经济增长差异的再反思:历史起因与演化逻辑》,《经济学(季刊)》2019年第1期。
- [14]. 赵丹、孙东琪、陈明星:《长三角县域经济增长的时空差异与影响因素》,《经济经纬》2020年第4期。
- [15]. 张军、吴桂英、张吉鹏:《中国省际物质资本存量估算:1952—2000》,《经济研究》2004年第10期。