

# 基于空间面板数据模型的区域经济增长研究\*

## Application of Spatial Panel Data Model on the Analysis of Regional Economic Growth

李春红, 文利霞, 黄登香

LI Chun-hong, WEN Li-xia, HUANG Deng-xiang

(广西大学数学与信息科学学院, 广西南宁 530004)

(College of Mathematics and Information Science, Guangxi University, Nanning, Guangxi, 530004, China)

**摘要:**为分析我国省际及各时段经济增长的影响因素,在空间面板数据模型基础上加入时间和空间固定效应,建立无固定效应模型、空间(地区)固定效应模型、时间固定效应模型和时空(时间空间)混合固定效应模型等4个不同的模型,来拟合分析2001~2010年中国31个省市的经济数据。发现,时空混合效应面板数据模型的拟合度最高,中国各地区的经济增长的时空依赖关系是显著的,全国各地区的经济增长不仅具有地区差异性,还呈现出阶段性特征。

**关键词:**空间面板数据模型 固定效应 区域经济增长

中图分类号:O123.9 文献标识码:A 文章编号:1002-7378(2013)04-0265-04

**Abstract:**To analyzing the influence factors of economic growth among 31 provinces and cities of China in each period, four models, based on spatial panel data model with spatial and time fixed effects, are built to estimate and analyze the economic data of 31 provinces and cities during 2001~2010 in China. Comparing all the models, it is found that the spatial panel data model with spatial and time fixed effects is the best with the highest fitting degree, and regional spatial dependence of economic growth is statistically significant. Therefore, the results show that the regional economic growth is not only with spatial dependence but also time correlation.

**Key words:**spatial panel data model, fixed effect, regional economic growth

面板数据是指对不同时刻的截面个体进行连续观测所得到的多维时间序列数据,能同时反映时间和横截面两个方向上的变化规律。1988年,Anselin<sup>[1]</sup>在传统的面板数据模型中考虑空间相关性,提出空间面板数据模型。之后空间面板数据模型得到广泛关注和应用。Elhorst<sup>[2]</sup>对空间面板数据模型的分类作了详细的介绍,并初步讨论这些模型的估计和检验方法。胡洪胜<sup>[3]</sup>、季民河等<sup>[4]</sup>、张志强<sup>[5]</sup>分别对空间模型的检验问题、设定问题、参数估计的小样本特性作深入的研究。在经济增长的实证

分析方面,梁亚民等<sup>[6]</sup>运用传统面板数据模型,对苏中、苏南、苏北的区域间经济增长影响差异进行实证分析。苏良军等<sup>[7]</sup>选用空间面板数据模型比较长三角和珠三角地区经济增长情况,在加入时间和空间固定效应后,得出“这两个地区都存在经济增长的 $\beta$ 收敛”的结论(即经济基础较差的地区趋于更快的增长速度)。

采用空间面板数据模型研究经济增长问题,得到的结果比横截面方法更精确,因为空间面板数据模型不仅考虑空间相关,也将时间相关纳入其中。若再结合横截面和时间序列方法,则能更准确地反映区域经济增长的本质。现有文献大多还是采用横截面方法或时间序列方法研究经济增长问题,空间面板数据模型使用较少,有些人选用空间面板数据模型时,也只是估计出回归系数和空间相关系数,并

收稿日期:2013-07-16

修回日期:2013-09-10

作者简介:李春红(1971-),女,副教授,主要从事应用统计研究。

\*广西自然科学基金(No. 2012GXNSFBA053010),广西大学校基金(No. 20110626)资助。

未研究时间和空间固定效应参数的估计。本文在空间面板数据模型基础上加入了时间和空间固定效应,建立无固定效应模型、空间(地区)固定效应模型、时间固定效应模型和时空(时间空间)混合固定效应模型等4个不同的模型,分析我国31个省市2001~2010年的经济增长问题。发现,时空混合效应面板数据模型的拟合度最高,我国各地区的经济增长的时空间依赖关系是显著的。

### 1 空间面板数据模型

#### 1.1 模型

空间面板数据模型分为空间滞后面板数据模型(1)和空间误差面板数据模型(2):

$$y = \rho(I_T \otimes W_N)y + X\beta + (1_T \otimes I_N)\eta + (I_T \otimes 1_N)\delta + \epsilon, \tag{1}$$

$$\begin{cases} y = X\beta + (1_T \otimes I_N)\eta + (I_T \otimes 1_N)\delta + v, \\ v = \lambda(I_T \otimes W_N)v + \epsilon, \end{cases} \tag{2}$$

其中 $\otimes$ 为矩阵的kronecker乘积, $N$ 为空间维度(即所研究的地区总数), $T$ 为时间维度(即研究的时期数), $I_n$ 为 $n \times n$ 单位矩阵, $n=N$ , $1_n$ 为 $n \times 1$ 全1向量, $y(NT \times 1)$ 为被解释向量, $X(NT \times k)$ 为设计阵, $\beta(k \times 1)$ 为回归系数向量, $\rho$ 和 $\lambda$ 分别为空间相关系数和残差相关系数, $W(N \times N)$ 为空间权重矩阵, $\epsilon(NT \times 1)$ 为误差向量, $\eta(N \times 1)$ 为空间固定效应向量, $\delta(T \times 1)$ 为时间固定效应向量,且 $\eta \sim N(0, \sigma_\eta^2 I_N)$ , $\delta \sim N(0, \sigma_\delta^2 I_T)$ , $\epsilon \sim N(0, \sigma_\epsilon^2 I_{N \times T})$ 。其中空间权重矩阵一般是依据空间相邻来设定,其对角元素都为0,若地区*i*和地区*j*相邻,则其(*i*,*j*)元设为1,否则为0。实践中采用标准化空间权重矩阵,即*W*的每行元素之和为1,以保证 $\rho$ 的真值位于 $\pm 1$ 之间。

#### 1.2 参数估计

对空间面板数据模型,用极大似然方法作参数估计。先将数据作如下转换: $y_{it}^* = y_{it} - \frac{1}{T} \sum_{t=1}^T y_{it}$ ,

$x_{it}^* = x_{it} - \frac{1}{T} \sum_{t=1}^T x_{it}$ ,那么对模型(2),其对数似然函数为

$$\log L = -\frac{NT}{2} \log(2\pi\sigma^2) + T \log |I_N - \lambda W| - \frac{1}{2\sigma^2} \sum_{i=1}^N \sum_{t=1}^T \{y_{it}^* - \lambda [\sum_{j=1}^N w_{ij} y_{jt}^*] - (x_{it}^* - \lambda [\sum_{j=1}^N w_{ij} x_{jt}^*])\beta\}^2. \tag{3}$$

对给定的 $\lambda$ ,可以得出 $\beta$ 和 $\sigma^2$ 的极大似然估计

$$\hat{\beta} = ([X^* - \lambda(I_T \otimes W)X^*]' [X^* - \lambda(I_T \otimes W)X^*])^{-1} [X^* - \lambda(I_T \otimes W)X^*]' [Y^* - \lambda(I_T \otimes W)Y^*], \tag{4}$$

$$\hat{\sigma}^2 = e(\lambda)'e(\lambda)/NT, \tag{5}$$

其中 $e(\lambda) = Y^* - \lambda(I_T \otimes W)Y^* - [X^* - \lambda(I_T \otimes W)X^*]\beta$ 。关于 $\lambda$ 的集中似然函数为

$$\log L = -\frac{NT}{2} \log[e(\lambda)'e(\lambda)] + T \log |I_N - \lambda W|. \tag{6}$$

求使得以上似然函数最大的 $\lambda$ 即为其极大似然估计 $\hat{\lambda}$ 。类似地,可以得出模型(1)各参数的极大似然估计。

在模型选择的问题上,目前通行的方法是先用最小二乘法估计不考虑空间相关性的受约束模型,然后进行空间相关性检验,如果统计量LMsar(或LMsem)比LMsem(或LMsar)更显著,则认为空间滞后模型(或空间误差模型)是比较合适的模型。

### 2 基于空间数据面板模型的区域经济增长分析

#### 2.1 数据来源和变量选取

选取2001~2010年中国31个省市的经济增长数据(来自国家统计局网站上的2001~2011年《中国统计年鉴》<sup>[8]</sup>),共310个观测值。

因变量选取为人均GDP增长率(%),自变量分别为 $x_1$ :滞后一期人均GDP(元)的自然对数; $x_2$ :年末总人口(万人)的自然对数; $x_3$ :人口自然增长率(%); $x_4$ :全社会固定资产投资占GDP的比例(%); $x_5$ :政府消费占GDP的比例(%); $x_6$ :公路密度(km/km<sup>2</sup>)。

#### 2.2 模型建立

首先用最小二乘虚拟变量法估计传统的固定效应模型,其中DW值为1.3391,说明存在自相关。因此,对数据进行空间自相关检验,检验结果见表1。

表1 空间自相关检验结果

检验方法	样本数	检验值	P值
LMlag	310	54.1489	0.000
LMerror	310	63.1151	0.000
Robust LMlag	310	5.1419	0.023
Robust LMerror	310	14.1080	0.000

从表1结果看,检验的P值都小于0.05,说明

存在空间自相关。LMerror 检验值大于 LMlag 检验值, Robust LMerror 检验值大于 RobustLMlag 检验值, 则认为空间误差面板数据模型较为合适。因此, 本文选择空间误差面板数据模型进行分析。

由于本文是对中国 31 个主要城市的经济增长的分析, 所研究的截面单位是特定的。所以, 选择空间固定效应模型更为合适。因此建立 4 个空间误差面板数据模型: 无固定效应模型、空间(地区)固定效应模型、时间固定效应模型和时空(时间空间)混合固定效应模型。4 个模型的分量形式如下:

(1) 无固定效应模型

$$y_{it} = x_{1,it}\beta_1 + x_{2,it}\beta_2 + x_{3,it}\beta_3 + x_{4,it}\beta_4 + x_{5,it}\beta_5 + x_{6,it}\beta_6 + u_{it}, u_{it} = v_{it}, v_{it} = \lambda W_{31}v + \epsilon_{it},$$

(2) 空间固定效应模型

$$y_{it} = x_{1,it}\beta_1 + x_{2,it}\beta_2 + x_{3,it}\beta_3 + x_{4,it}\beta_4 + x_{5,it}\beta_5 + x_{6,it}\beta_6 + u_{it}, u_{it} = \eta_i + v_{it}, v_{it} = \lambda W_{31}v + \epsilon_{it},$$

(3) 时间固定效应模型

$$y_{it} = x_{1,it}\beta_1 + x_{2,it}\beta_2 + x_{3,it}\beta_3 + x_{4,it}\beta_4 + x_{5,it}\beta_5 + x_{6,it}\beta_6 + u_{it}, u_{it} = \delta_t + v_{it}, v_{it} = \lambda W_{31}v + \epsilon_{it},$$

(4) 时空混合固定效应模型

$$y_{it} = x_{1,it}\beta_1 + x_{2,it}\beta_2 + x_{3,it}\beta_3 + x_{4,it}\beta_4 + x_{5,it}\beta_5 + x_{6,it}\beta_6 + u_{it}, u_{it} = \eta_i + \delta_t + v_{it}, v_{it} = \lambda W_{31}v + \epsilon_{it},$$

其中  $\eta_i$  为第  $i$  个地区的固定效应参数,  $\delta_t$  为第  $t$  个时段的固定效应参数,  $i = 1, 2, \dots, 31; t = 1, 2, \dots, 10$ 。

### 2.3 分析结果

采用 Matlab 软件和 Spatial econometrics 模块对以上 4 个模型进行参数估计, 回归系数、空间固定效应参数和时间固定效应参数的估计结果分别见表 2、表 3 和表 4。由表 2~3 的结果得出如下结论:

(i) 各个模型的  $\lambda$  值都显著为正, 说明中国区域经济的经济增长存在正的空间相关性, 即集聚效应, 富裕的地区可以带动周边地区的经济发展。特别地, 在同时加入空间和时间固定效应后, 相对前两个模型,  $\lambda$  值显著减少到 0.171, 这说明加入空间和时间固定

表 2 模型回归系数估计及相关结果

系数	无固定效应模型	空间固定效应模型	时间固定效应模型	时空固定效应模型
$\beta_1$	-0.027(0.746)	0.201(0.282)	-0.164(0.039)	-0.685(0.000)
$\beta_2$	0.029(0.653)	-0.513(0.653)	-0.047(0.480)	-2.028(0.053)
$\beta_3$	0.304(0.000)	0.102(0.276)	0.239(0.000)	0.079(0.384)
$\beta_4$	0.014(0.871)	0.126(0.120)	-0.214(0.071)	-0.190(0.146)
$\beta_5$	-0.168(0.012)	-0.262(0.017)	-0.223(0.000)	-0.360(0.000)
$\beta_6$	-0.131(0.133)	-0.039(0.786)	-0.112(0.119)	-0.080(0.510)
$\lambda$	0.492(0.000)	0.471(0.000)	0.143(0.049)	0.171(0.023)
$R^2$	0.1175	0.2587	0.3856	0.4957
L	-388.952	-365.162	-361.292	-329.992

效应有助于消除空间自相关性。

(ii) 从  $\beta$  收敛的估计结果看, 在空间固定效应模型中  $\beta_1$  的符号为正, 其他 3 个模型为负, 且在空间固定效应模型和时空固定效应模型中显著为负, 表明我国经济增长不存在绝对  $\beta$  收敛。仅有在控制了时间固定效应和时空固定效应后, 期初人均 GDP 的系数才显著为负, 即为条件  $\beta$  收敛, 表明经济基础好的地区经济增长较慢, 反之则增长越快。随着时间的推移和经济的发展, 长时间内我国各地区的经济增长会维持各自的稳定水平。

(iii) 在建立的 4 个模型中, 时空混合固定效应模型拟合优度最高, 解释力度达到了 0.4957, Log-likelihood 值较其他 3 个模型也有较大的改进。这表明在同时控制了时间和空间固定效应后, 得到的模型较好, 说明各省份的经济增长确实存在地区性差异和阶段性特征。

(iv) 在时空混合固定效应模型的估计结果中,  $\beta_2$ 、 $\beta_5$  的估计显著为负, 即人口总数抑制该地区的经济发展, 政府消费的减少可提高该地区的经济增长速度, 政府消费每减少 1% 经济增长速度随之提高 0.36%。因此, 经济增长速度与期初人均 GDP、年末人口总数和政府消费占 GDP 比重息息相关。

(v) 从空间固定效应的估计结果看, 不同的地区, 其固定影响参数存在很大的差异性, 地区经济发展水平越高的地区, 对经济增长的影响就越大, 如上海、江苏、浙江、山东、广东等沿海经济发达地区, 其影响参数分别达到了 1.252、3.366、2.273、3.279、3.325, 其次是河北、辽宁、河南、湖北、湖南等中东北部地区, 西部地区及偏远地区如海南、贵州、西藏、甘肃、青海、宁夏、新疆最低, 这与我国的各地区经济发展现状相符。

(vi) 时间固定效应参数的估计结果表明我国经济增长呈现阶段性特征。从 2004 年起我国经济稳步增长, 尤其是 2008 年, 由于北京奥运会的举办, 使得我国第三产业迅速发展, 也带动其他产业的增长。

表3 空间固定效应参数估计结果

地区	$\eta_i$	地区	$\eta_i$	地区	$\eta_i$	地区	$\eta_i$
北京	1.251	上海	1.252	湖北	0.955	云南	-0.756
天津	-0.170	江苏	3.366	湖南	1.101	西藏	-5.629
河北	1.589	浙江	2.273	广东	3.325	陕西	-0.132
山西	-0.005	安徽	0.277	广西	0.050	甘肃	-1.927
内蒙古	0.623	福建	0.661	海南	-3.892	青海	-4.192
辽宁	1.251	江西	-0.410	重庆	-0.434	宁夏	-3.980
吉林	-0.296	山东	3.279	四川	1.104	新疆	-1.142
黑龙江	0.269	河南	2.062	贵州	-1.596		

表4 时间固定效应参数估计结果

年份	$\delta_t$	年份	$\delta_t$	年份	$\delta_t$	年份	$\delta_t$
2001	-2.571	2002	-1.845	2003	-1.144	2004	0.167
2005	0.255	2006	0.242	2007	0.836	2008	1.537
2009	0.624	2010	1.901				

### 3 结论

本文分别选用无固定效应模型、空间固定效应模型、时间固定效应模型和时空混合固定效应模型,分析我国31个省市2001~2010年的经济增长问题,不仅得出了回归系数的估计,还进一步做了空间和时间固定效应参数的估计,同时综合考虑了空间相关性和时间相关性,得到我国经济发展差异的影响因素。发现,我国区域经济增长呈现出明显的区域特征和阶段性特征。表明经济增长不仅依赖于地区自身条件,还与相邻地区的经济增长和经济周期息息相关。期初人均GDP、年末人口总数和政府消

费占GDP比重对我国经济增长的影响较为显著。

此外,由于空间权重矩阵的选取对研究结果有很大的影响,采用经济权重矩阵或更精确的权重矩阵是否可以得出更准确的结果,还有待于进一步研究。

#### 参考文献:

- [1] Anselin L. Spatial economics: methods and models [M]. Dordrecht :Kluwer Academic,1988.
- [2] Elhorst J P. Specification and estimation of spatial panel data models[J]. International Regional Science Review,2003,26(3):244-268.
- [3] 胡洪胜. Two-way 空间自回归随机效应面板数据模型的检验[J]. 中央民族大学学报:自然科学报,2012,21(1):84-87.
- [4] 季民河,武占云,姜磊. 空间面板数据模型设定问题分析[J]. 统计与信息论坛,2011,26(6):3-8.
- [5] 梁亚民,臧海明. 区域差异视角下江苏省经济增长的实证分析——基于面板数据模型的研究[J]. 统计教育,2009,114(3):38-40.
- [6] 苏良军,王芸. 中国经济增长空间相关性研究——基于“长三角”与“珠三角”的实证[J]. 数量经济技术研究,2007(12):26-37.
- [7] 张志强. 空间面板数据参数估计的小样本特性探究[J]. 数量经济技术经济研究,2012(9):122-140.
- [8] 中华人民共和国统计局. 国家数据[EB/OL]. [2013-06-10]. <http://data.stats.gov.cn/workspace/index?m=hgnd>.

(责任编辑:尹 闯)

(上接第264页)

- [3] 马燕,范植华. 基于神经网络的基因分类器[J]. 计算机工程与设计,2005,26(2):308-311.
- [4] 蔡春,苗立峰,邓乃扬. DNA 序列特征提取方法研究[J]. 北京联合大学学报:自然科学版,2008,22(4):70-72.
- [5] 孙健,申瑞民,韩鹏. 一种新颖的径向基函数(RBF)网络学习算法[J]. 计算机学报,2003,6(11):45-47.
- [6] Shu W, He B. A quantum genetic simulated annealing algorithm for task scheduling[J]. ACM Computing Sur-

veys,2006,33(1):115-127.

- [7] 龚道雄,阮晓钢. 一种基于遗传算法的DNA多序列比对方法[J]. 北京工业大学学报,2003,3(1):19-22.
- [8] Zhang Y, Waterman M S. An eulerian path approach to global multiple alignment for DNA sequences[J]. Journal of Computational Biology,2003,10(6):803-819.

(责任编辑:尹 闯)