

上海新金融研究院工作论文系列

No. SFIWP0025

中国各省经济增长是否收敛？

朱国忠 乔坤元 虞吉海

2013年12月9日

说明：上海新金融研究院是为支持上海国际金融中心建设而成立的非营利性金融类专业学术研究机构，由中国金融四十人论坛举办，并与上海市黄浦区人民政府展开战略合作。

本工作论文是上海新金融研究院研究人员在工作期间形成的、尚未公开发表的研究成果，文中观点仅代表作者本人，不代表本研究院。未经书面同意，谢绝任何形式的转载和复制。

中国各省经济增长是否收敛？^①

朱国忠 乔坤元 虞吉海

摘要：本文使用空间动态面板数据模型和 1952-2008 年省级 GDP 数据分析中国经济增长的收敛性，发现中国各省人均 GDP 总体上不存在收敛性；落后省份的增长速度并不比富裕省份的高。此外，中国各省人均 GDP 有一定的空间相关性，但总体上不是很强。我们进一步对东部、中部、西部三个地区分别进行回归，发现东部地区经济不收敛，但是空间效应明显；而中西部地区经济收敛，但是空间效应不明显。这说明中国地区发展不平衡是各省经济增长总体上缺乏收敛性的原因之一。

关键词：增长收敛性，空间动态面板数据模型，空间计量

一、引言

区域经济增长和平衡发展是宏观经济学中一个重要的议题，地区间经济水平差距长时间存在和过分的拉大不仅造成收入分配失衡和福利损失，而且很有可能会影响到经济的整体效率和社会的稳定。现存的宏观经济模型一般认为各地区的经济增长应该是呈收敛趋势的（如 Solow, 1956），并且这一预测在数据中得到了多方面的验证，例如 Barro and Sala-i-Martin (1992) 发现美国各州的经济增长明显地呈现绝对收敛的趋势。有关中国经济的收敛性，现存的大部分研究都认为中国各省 GDP 增长存在收敛性。

上述文献中的实证分析基本上都运用传统的面板数据模型和时间序列模型。这些传统模型存在一个假设：各个区域之间在经济上不存在互动关系。从理论上说，即使各区域之间完全没有互动，其经济发展水平可能还是会趋同。例如在传统的索罗模型框架下，给定技术水平，由于资本收益率递减，所有经济体都会递归到人均产出的稳态水平。然而，经济趋同还有一个很重要的原因，那就是资本、技术和劳动力的跨区域流动，这一类的区域间互动关系会加速趋同。Sachs and Warner (1997) 的实证研究表明，开放程度高的国家人均收入向稳态水平趋近的速度是开放程度低的国家的一倍以上。传统计量模型假设此类互动关系为零，可能会带来估计上的偏差。

我们运用空间动态面板数据模型来解决这个问题。这一模型不排除各个地区在经济上存在互动关系的可能性。我们称这种互动关系为空间相关性，并且估计经济的空间相关性系数和收敛系数。和传统的动态面板数据模型相比，我们使用的估计方程增加了各地区经济增长的交互项，也就是上述的空间效应。因此从本质上说，空间动态面板数据模型是对传统面板数据模型的拓展。我们根据新古典增长理论中的柯布-道格拉斯生产函数和技术的空间溢出性推导出估计方程中的交互项，因而这一拓展是建立在新古典增长理论的基础之上的。

需要特别指出的是，空间相关性只是引发经济增长收敛（或者发散）的诸多因素之一，其他因素包括各地区的自然资源、国家政策等等。因此，空间相关性的存在不能保证经济增长的收敛。反之，空间相关性的缺失也并不意味着经济增长的发散。这些情况都在我们的分区回归结果中得到了体现。因此，通过对空间相关性的估计，我们能够更好地理解引发经济增长收敛（或者发散）的原因。

^① 朱国忠，北京大学光华管理学院，电话：010-62767407，邮箱：aaronandrewzhu@gmail.com。乔坤元，上海新金融研究院、北京大学光华管理学院，电话：010-62763430，邮箱：qiaokunyu@126.com。虞吉海（通讯作者），北京大学光华管理学院，电话：010-62767407，邮箱：jihai.yu@gmail.com。地址：北京市颐和园路 5 号北京大学光华管理学院 2 号楼 360 室。本文即将发表于《经济学》（季刊）2014 年第 14 卷第 3 期，作者衷心的感谢审稿人认真细致的审稿，当然，文责自负。

我们使用的各省 1952-2008 年的数据进行参数估计，发现中国大陆各省的人均 GDP 总体上是不存在收敛性的。我们估计出的空间相关性系数在某些设定下显著，在别的设定下不显著，也即中国各省之间经济互动关系总体而言不是很强。通过对东部、中部、西部三个地区分别进行回归，我们发现东部地区各省的人均 GDP 总体上是不收敛的，但是他们之间的空间效应明显。对于中部地区和西部地区，我们发现人均 GDP 的增长是收敛的，即落后省份发展速度会快于相对发达的省份，但是它们之间的空间效应不明显。这说明中国经济增长呈现一定的俱乐部收敛的特征，但是这种收敛并不是由俱乐部成员之间的经济互动所致，因此我国经济增长收敛的源泉还需要进一步的探究，比如各个省的自然资源禀赋、国家对各个省的政策以及各个省内部的经济政策等等，一个典型的例子是我国改革开放以来对于东部沿海地区的政策偏倚，“先富带后富”的发展战略虽然在调动地方发展经济的积极性上起到了一定的作用，但实际上却在一定程度上加剧了区域发展的不平衡，一些文献认为政府的政策调整可以减缓区域发展的不平衡的现象，具体见林毅夫等（1998）、刘夏明等（2004）。

为了在最大程度上使用数据当中的信息，同时避免商业周期波动带来的影响，我们既使用每个时间段的平均值进行分析，又采取了经验文章中估计面板数据的一般办法，把整个样本时间段细分为几个较小的时间段，使用每个时间段的首位观测进行分析。在国内关于经济增长是否收敛的文献中，大部分研究采用了年度数据。Islam (1995) 指出，使用年度数据的估计结果会在很大程度上受到短期的商业周期波动的影响，如果使用每隔几年的数据作样本的话，回归方程中的干扰项就会免于商业周期的影响，并且不太可能出现干扰项之间的自相关性很强的情况。

基于这两种数据，我们首先运用传统的横截面数据模型，发现使用 4 年均值和 4 年间隔数据的大多数估计结果不支持各省的人均 GDP 增长收敛的假说。接着我们加入个体效应但忽略空间相关性，使用动态面板数据模型对 4 年均值和 4 年间隔数据进行检验，却得到了不一致的估计结果，前者支持中国各省之间人均 GDP 发散的假说，而后者则支持各省的人均 GDP 出现了收敛的说法。然后，我们在横截面数据模型当中加入空间相关性，依然无法从这一静态的空间模型中得到一致的估计结果，4 年均值数据依然支持各省之间的经济增长呈现不收敛的假说，而 4 年间隔数据则认为各省之间经济增长出现了一定程度上的收敛，这说明如果忽略各省的个体效应，如资源禀赋和文化差异等因素，也可能得到有偏的结果。最后，我们使用空间动态面板数据模型进行估计，同时考虑个体效应和空间效应。我们的估计结果一致的表明我国各省经济增长没有出现收敛，并且空间交互项的估计参数大多数是不显著的。最后，我们进一步通过对东部、中部、西部三个地区进行分别回归，发现东部地区经济增长不收敛，但空间效应明显；而中西部地区经济增长收敛，但空间效应不明显。上述层层递进的研究策略有两大好处。首先，这有助于比较我们的结果和现存的基于零空间效应假设得到的研究结果。其次，这有助于我们评估基于零空间效应和静态空间效应假设的估计结果。

本文将在下一个部分进行文献回顾，紧接着在第三部分给出理论分析的框架和相应的数据描述，而后在第四部分讨论计量模型和相应的估计方法，在第五部分呈现实证结果，最后得到结论和政策建议。

二、文献回顾

从理论角度看，索洛模型（Solow, 1956）是讨论经济增长收敛性的一个很好的起点。在索洛模型中，决定一个国家或地区经济发展水平的因素包括资本存量和劳动投入这些内生变量，以及储蓄率、人口增长率以及技术进步率这些外生变量。给定外生变量，由于资本边际产出递减，一个经济体会收敛到稳态均衡点。如果各个经济体的储蓄率、人口增长率和技术进步率这些外生因素是相同的，那么所有的经济体都会递归到同一个稳态均衡点，也就是文献中所说的绝对收敛。然而在现实当中，世界各个国家和地区的储蓄率、人口增长率和技术进步速率均存在较大的差异，因此根据索

罗模型它们会收敛到不同的稳态均衡点，这就是文献中所说的条件收敛。Barro and Sala-i-Martin (1992) 研究了条件收敛，认为各个国家经济增长的均衡水平根据各国资源禀赋或者文化差异而有所不同。由于资本的边际产出是递减的，因此给定其他条件不变，可以推出贫穷国家的资本边际报酬更高，经济增长比富裕国家更快的结论，从而最终任何经济体都会收敛到人均收入稳态。Lucas (1988) 提出的内生增长模型中，由于人力资本的积累，物质资本的边际报酬是固定的，因此贫穷的国家或者地区在人均收入上并不向富裕国家或地区收敛，但是两类国家的增长率是收敛的。诸多讨论增长收敛性的文献中都讨论到要素的流动性问题（如 Lucas, 1988），并指出如果资本、劳动力或者技术可以自由地跨区流动，那么经济增长的收敛性会得到极大地强化。

为了估计地区经济增长的条件收敛性，现存的实证分析文献主要使用三类模型：横截面数据模型、动态面板数据模型和空间动态面板数据模型。Barro and Sala-i-Martin (1992) 使用横截面数据模型探究美国 48 个毗邻的州（不包括阿拉斯加和夏威夷）的条件收敛性，而 Islam (1995) 认为各个州的个体特征会对经济增长产生一定的影响，从而提出应该使用动态面板模型来研究地区经济增长的条件收敛性。由于地区经济增长收敛性的重要原因之一是各地之间的经济互动而产生的空间溢出效应，从空间计量经济学的角度来看，忽略了空间溢出效应会使得估计出现偏差。因此一支新的文献主要使用空间动态面板数据模型的方法来探究条件收敛性的问题，如 Yu and Lee (2012)。

国内学者对于收敛性进行了大量的研究，第一大类文献没有考虑空间相关性，主要讨论了我国东、中、西部三个地区之内和之间的经济增长收敛性：一方观点认为这三个地区内部各省市的经济增长存在收敛性，如林毅夫等 (1998)、刘夏明等 (2004)，但区域之间经济增长发散；另一方的观点则认为，三个区域内部同样出现了发散的现象，并且东部地区内部增长收敛性最低，如段平忠 (2008)。第二类文献考虑中国各地经济的空间相关性，如陈安平和李国平 (2004) 检验了中国东、中、西部三个地区内和地区间人均产出的协整关系，刘生龙和张捷 (2009) 运用空间计量的方法，发现 1985-2007 年间我国三个区域经济长期存在绝对 β 收敛和条件 β 收敛，而短期内只存在条件 β 收敛。

本文对现有的文献进行了重要的补充，这主要体现在两个方面。第一，我们用的是空间动态面板数据模型的方法进行研究，而现有的文献不是忽视空间相关性，就是使用静态的空间模型。由于我们需要充分考虑个体效应（资源禀赋差异以及文化差异）和时间效应，空间动态面板数据模型就更加合适。第二，我们估计得到的空间相关性系数包含了现存研究不能提供的信息，并且发现总体上中国各省经济的互动关系不是很强。东部地区各省之间存在较强的互动关系，但是在中部和西部此类互动关系不显著。

三、收敛的新古典分析：模型和数据

（一）空间动态面板数据模型的理论推导

在最简单的新古典增长模型中，由于贫穷的国家（地区）人均资本存量低，其资本的边际产出较高，资本积累速度较快，导致人均 GDP 增幅高于富裕的国家和地区。在此基础上，如果国家（地区）之间存在由资本、技术、劳动力或者其他因素的流动而产生的空间溢出效应，那么国家（地区）之间的收敛性会更加明显。我们分析的起点就是一个存在空间溢出效应的索洛模型，参考了 Yu and Lee (2012)。

考虑一个劳动增广的索洛模型

$$Y_{i,t} = K_{i,t}^\alpha (A_{i,t} L_{i,t})^{1-\alpha}$$

其中， Y 为产量， K 为资本存量， A 是劳动效率， L 为劳动存量。那么资本在时间上的增长率为

$$\dot{K}_{i,t} = sY_{i,t} - \delta K_{i,t}$$

其中 s 为储蓄率， δ 为折旧率。

在古典增长模型中，人口和技术进步分别按照外生的 p 和 g 的速率增长，也即：

$$L_{i,t} = L_{i,0} e^{pt} \quad A_{i,t} = A_{i,0} e^{gt}$$

在本文中我们使用技术溢出来探讨空间的相关性，因此一个地区的技术进步会对别的地区产生溢出效应，我们将技术设定为：

$$A_{i,t} = A_{i,0} e^{gt} \prod_{j \neq i}^n A_{j,t}^{\lambda \omega_{i,j}}$$

区域（经济体） i 在时刻 t 的技术为 $A_{i,t}$ ，它是由初始水平 $A_{i,0}$ 和它的增长率 g 及其邻省的技术增长率 $A_{j,t}$ 共同决定的。溢出效应的大小体现在 λ 当中且 $0 < \lambda < 1$ ，而 $\omega_{i,j}$ 则度量了邻省 j 的技术溢出是否以及在多大程度上传导给了区域（经济体） i 。

对 $A_{i,t}$ 取对数，我们有

$$\ln A_{i,t} = \ln A_{i,0} + gt + \lambda \sum_{j \neq i}^n \omega_{i,j} \ln A_{j,t}$$

写成矩阵的形式，我们令 $\mathbf{A}_t = (A_{1,t}, A_{2,t}, A_{3,t}, A_{4,t}, \dots, A_{n,t})'$ ，从而有

$$\ln \mathbf{A}_t = (I - \lambda W_n)^{-1} \ln \mathbf{A}_0 + \frac{gt}{1 - \lambda} \mathbf{l}_n$$

其中 \mathbf{l}_n 为由 n 个 1 组成的列向量，矩阵 W_n 的对角元素均为 0 从而避免自身的影响，并且各个元素标准化至 0 到 1 之间。因此， $A_{j,t}$ 在时间上的增长率为

$$\frac{\dot{A}_{i,t}}{A_{i,t}} = \frac{g}{1 - \lambda}$$

令 $\hat{k}_{i,t}$ 和 $\hat{y}_{i,t}$ 为区域（经济体） i 的均衡有效人均资本存量和经济产出，即 $\hat{k}_{i,t} = \frac{k_{i,t}}{A_{i,t} L_{i,t}}$ ，

$\hat{y}_{i,t} = \frac{y_{i,t}}{A_{i,t} L_{i,t}}$ 。因此， $\hat{y}_{i,t} = \hat{k}_{i,t}^\alpha$ 。我们可以得到：

$$\dot{\hat{k}}_{i,t} = s \hat{k}_{i,t}^\alpha - \left(\frac{\dot{L}_{i,t}}{L_{i,t}} + \frac{\dot{A}_{i,t}}{A_{i,t}} + \delta_i \right) \hat{k}_{i,t} = s \hat{k}_{i,t}^\alpha - \left(p + \frac{g}{1 - \lambda} + \delta_i \right) \hat{k}_{i,t}$$

在均衡时， $\dot{\hat{k}}_{i,t} = 0$ ，意味着

$$\hat{k}_{i,t}^* = \left(\frac{s_i}{p + g / (1 - \lambda) + \delta_i} \right)^{\frac{1}{1-\alpha}} \quad \hat{y}_{i,t}^* = \left(\frac{s_i}{p + g / (1 - \lambda) + \delta_i} \right)^{\frac{\alpha}{1-\alpha}}$$

其中, s_i 和 δ_i 为区域 i 的储蓄率和折旧率。可以看到, 如果储蓄率和折旧水平一致的话, 那么经济体将收敛到相同的资本存量和产出水平。我们对有效劳均资本存量进行对数线性化, 就可以得到

$$\frac{\partial(\ln \hat{k}_{i,t} - \ln k_i^*)}{\partial t} = -\phi^i (\ln \hat{k}_{i,t} - \ln k_i^*)$$

其中

$$\phi^i = (1-\alpha) \left(p + \frac{g}{1-\lambda} + \delta_i \right)$$

是区域 i 的收敛速度。根据 Yu and Lee (2012) 的推导, 我们可以进一步假设折旧率 δ_i 全部一致, 从而各省的经济增长收敛速度都相等。

求解关于 $\ln k$ 的一阶差分方程并且使用 $\ln \hat{k}_{i,t}$ 与 $\ln \hat{y}_{i,t}$ 在时间上的路径相同的性质 (因为 $\hat{y}_{i,t} = \hat{k}_{i,t}^\alpha$), 我们可以得到

$$\ln \hat{y}_{i,t2} = -e^{-\phi t} \ln y_{i,t1} + (1 - e^{-\phi t}) \ln y_i^*$$

其中 $\tau = t_2 - t_1$, $\hat{y}_i^* = \left(\frac{s_i}{p + g / (1-\lambda) + \delta_i} \right)^{\frac{\alpha}{1-\alpha}}$ 并不取决于 t 。除了收敛速率 ϕ 更大以外, 上式与 Islam

(1995) 的完全一致。进一步, 我们令 $y_{i,t} = \frac{y_{i,t}}{L_{i,t}}$ 为人均产出, 那么 $\ln y_{i,t} = \ln \hat{y}_{i,t} + \ln A_{i,t}$, 因此我

们有

$$\ln y_{i,t2} = \ln A_{i,t2} - e^{-\phi \tau} \ln A_{i,t1} + e^{-\phi \tau} \ln y_{i,t1} + \ln y_i^* (1 - e^{-\phi \tau})$$

将上述 $\ln A_t$ 的表达式带入,

$(I - \lambda W_n) \ln y_{i,t2} = (I - \lambda W_n) e^{-\phi \tau} \ln y_{i,t1} + (1 - e^{-\phi \tau}) \ln A_0 + g(t_2 - e^{-\phi \tau} t_1) l_n + (I - \lambda W_n)(1 - e^{-\phi \tau}) \ln \bar{y}_t^*$ 因此我们要估计的方程可以写为^①:

$$\ln y_t = \lambda W_n \ln y_t + \gamma \ln y_{t-1} + \rho W_n \ln y_{t-1} + \mathbf{c}_n + \eta_t l_n + V_{n,t}$$

其中 $\gamma = e^{-\phi}$ 衡量区域是否出现经济增长收敛, 而 $\rho = -\lambda e^{-\phi}$ 则刻画区域滞后空间效应的影

响, 而个体特征的向量为 $\mathbf{c}_n = (1 - e^{-\phi \tau}) \left(\ln A_0 + \frac{\alpha}{1-\alpha} \ln X_n - \frac{\alpha \lambda}{1-\alpha} W_n \ln X_n \right)$, 其中

$\mathbf{X}_n = \left(\frac{s_1}{p + g / (1-\lambda) + \delta}, \frac{s_2}{p + g / (1-\lambda) + \delta}, \dots, \frac{s_n}{p + g / (1-\lambda) + \delta} \right)'$, $\eta_t = g(t_2 - e^{-\phi \tau} t_1)$ 是时间固

定效应, 而 $V_{n,t} = (v_{1,t}, v_{2,t}, \dots, v_{n,t})'$ 为附加的误差项。

与 Islam (1995) 相比, 我们多了 $W \ln y_t$ 和 $W \ln y_{t-1}$, 这两项描述的是由于技术外溢带来的各省之

^①我们在后面将这一估计方程写成了加总的形式, 详见后文的公式 (1)。

间经济的相互依赖性，当 $\lambda=0$ 且 $\rho=0$ 时，我们的模型与 Islam (1995) 的完全一致。另外，个体效应 c_n 不仅仅取决于不同的初始技术水平 $A_{i,0}$ ，而且和各省其他的变量，如储蓄率 s_i 有关，如果这些 X_n 随着时间变化而变化并且是外生的，那么它们应该作为额外的控制变量而非个体效应。与此同时，根据模型推导，我们对于空间效应的滞后项系数的有一个约束，即 $\rho+\gamma\lambda=0$ 。为了保证实证的结果在更加广泛的环境中适用，我们在估计过程中并不强加这个约束条件。从我们的估计结果中（见表 6）可以看到， $\rho+\gamma\lambda=0$ 的原假设不能被拒绝，说明我们的模型是广泛适用的。

（二） 数据

我们的数据来源于单一的数据源《新中国六十年统计资料汇编》，为了保证数据的完整性，我们选取 1952 年到 2008 年的数据，一共 30 个省（直辖市和自治区，不包括海南），由于重庆市于 1997 年 3 月 14 日正式成为省级行政单位，我们考虑了这方面的因素，对其原属省份四川进行了一定的调整，使得我们的数据一致^①。

我们在不含空间效应的横截面、面板数据模型以及包含空间效应的横截面数据模型的回归中，为了在最大程度上使用数据信息，主要使用从 1952 年起的 4 年均值数据进行估计。考虑到 Islam (1995) 提到的商业周期的影响，我们同时取最长的以 1952 年为起点的 4 年间隔数据进行估计。在空间动态面板数据模型的估计当中，我们同时使用了以 1952、1953、1954、1955 为起点的每 4 年的产出和从 1952 年开始每 4 年产出的平均值，从而一共有 5 个样本。我们将借助空间动态面板数据模型对这 5 个数据集进行估计^②。

我们在下表中呈现经济增长的描述性统计量，可以看出各省的人均 GDP 差距并不是那么大，我们将在接下来的部分给出计量模型和估计方法，并且探究各省经济增长是否收敛。

表 1 描述性统计量

变量	定义	观测数	均值	标准差	最小值	最大值
logpcgdp	人均 GDP 对数值	1741	6.767	1.662	3.970	11.200
logpcgdplag	人均 GDP 对数值滞后项	1710	6.708	1.616	3.970	11.103
popgr	人口增长率 (%)	1741	1.750	2.320	-30.941	50.697
inv	投资率 (%)	1741	25.742	18.716	0.303	474.479

数据来自于《新中国 60 年统计资料汇编》。

四、 计量模型和估计方法

（一） 回归模型

首先，我们使用最简单的 OLS 对横截面数据进行回归。与 Barro 和 Sala-i-Martin (1992) 相似，我们的估计方程如下：

^①我们在稳健性检验中，同时对以 1952、1953、1954、1955 为起点的四个样本的 4 年间隔数据以及 1952-1980、1980-2008 两个子样本以及 1980-2008 年包括海南在内的 31 个省级行政单位的数据（1980h-2008h）也做同样的检验，得到了类似的结论。海南于 1988 年成为正省级行政单位，我们也做了类似的处理。但限于篇幅，我们没有汇报包括各个子样本的结果。

^②我们对以 1953、1954 和 1955 年为起点的 4 年间隔和均值数据也进行了类似的估计，得到的结果与本文汇报的结果基本一致。

$$\ln y_{i,T} = \alpha + \gamma \ln y_{i,0} + v_i \quad (1)$$

在这里，因变量 $\ln y_{i,T}$ 是末期人均 GDP，自变量 $\ln y_{i,0}$ 是初期人均 GDP。

其次，我们考察动态面板数据模型，和 Islam（1995）相同，我们的估计方程如下：

$$\ln y_{i,t} = \gamma \ln y_{i,t-1} + \alpha_i + \eta_t + \varepsilon_{i,t} \quad (2)$$

其中 α_i 是个体固定效应， η_t 是时间固定效应。

再接下来，我们估计包含空间效应的横截面模型（3）和包含空间效应及其滞后项的横截面模型（4）：

$$\ln y_{i,T} = \alpha + \gamma \ln y_{i,0} + \lambda \sum_{j \neq i}^n \omega_{i,j} \ln y_{j,T} + v_i \quad (3)$$

$$\ln y_{i,T} = \alpha + \lambda \sum_{j \neq i}^n \omega_{i,j} \ln y_{j,T} + \gamma \ln y_{i,0} + \rho \sum_{j \neq i}^n \omega_{i,j} \ln y_{j,0} + v_i \quad (4)$$

以上两式中的 $\omega_{i,j}$ 是权重矩阵，在下一段我们会对其进行具体介绍。

最后，我们使用动态空间面板数据进行回归，估计方程如下：

$$\ln y_{i,t} = \alpha + \lambda \sum_{j \neq i}^n \omega_{i,j} \ln y_{j,t} + \gamma \ln y_{i,t-1} + \rho \sum_{j \neq i}^n \omega_{i,j} \ln y_{j,t-1} + \mathbf{c}_i + \eta_t + v_i \quad (5)$$

从方程中可以看出，我们的动态空间面板数据模型包含了：（1）个体效应（ \mathbf{c}_i ），这样各省市之间的资源禀赋和文化差异可以得到控制；（2）时间效应（ η_t ），这样各期的政府宏观政策或者世界经济形势波动对地方经济发展的影响也得到了控制；（3）空间相关性，包括当期空间自回归项（ λ ）和滞后空间自回归项（ ρ ）。其中的参数， λ 和 ρ 描述各省经济之间的当期和滞后期的相关程度。我们允许不同省之间的相关程度有差异性，表现为不同的权重（ $\omega_{i,j}$ ）。具体操作中我们考虑两类权重。第一类权重基于 i 、 j 两省之间是否互为邻省，如果是邻省， $\omega_{i,j}=1$ 否则为 0。第二类权重基于 i 、 j 两省之间的可行距离^①倒数的平方做元素，即 $\omega_{i,j}=1/d_{ij}^2$ 。

我们对面板数据模型用不同的方法进行估计，以确保结果的稳健性。根据对时间效应不同的处理方法，我们有直接最大似然估计法（D-MLE）和转换最大似然估计法（T-MLE），具体的估计方法将在下一节给出。

（二）估计方法——准最大似然估计

我们使用准最大似然估计法（Quasi-Maximum Likelihood Estimation, QMLE）进行估计，

$$\ln \mathbf{Y}_{nt} = \lambda_0 \mathbf{W}_n \ln \mathbf{Y}_{nt} + \gamma_0 \mathbf{Y}_{nt-1} + \rho_0 \mathbf{W}_n \ln \mathbf{Y}_{nt-1} + \mathbf{c}_n + \eta_t \mathbf{I}_n + \mathbf{V}_{nt}, t = 1, 2, 3, \dots, T \quad (6)$$

其中， $\mathbf{Y}_{nt}=(y_{1t}, y_{2t}, \dots, y_{nt})'$ ， $\mathbf{V}_{nt}=(v_{1t}, v_{2t}, \dots, v_{nt})'$ 均为 $n \times 1$ 列向量，而 \mathbf{W}_n 为 $n \times n$ 空间权重矩阵， \mathbf{c}_n 为 $n \times 1$ 列向量，刻画个体效应， η_t 是时间固定效应，我们假设 v_{it} 是在时间和个体上都是独立同分布的，方差为 σ_0^2 。

我们记 $\theta=(\delta', \lambda, \sigma^2)'$ 并且 $\zeta=(\delta', \lambda, \mathbf{c}_n', \boldsymbol{\eta}_T)'$ ，其中 $\delta=(\gamma, \rho)'$ 并且 $\boldsymbol{\eta}_T=(\eta_1, \eta_2, \eta_3, \dots, \eta_T)'$ 。在真实值之下， $\theta_0=(\delta_0', \lambda_0, \sigma_0^2)'$ 并且 $\zeta_0=(\delta_0', \lambda_0, \mathbf{c}_{n0}', \boldsymbol{\eta}_{T0})'$ ，其中 $\delta_0=(\gamma_0, \rho_0)'$ 。我们记 $\mathbf{Z}_{nt}=(\mathbf{Y}_{n,t-1}, \mathbf{W}_n \mathbf{Y}_{n,t-1})$ 。那么方程（6）

^①可行距离即交通可以到达的距离，比如调整过某些山川的阻隔等等，我们使用 Google Earth 进行查找。

的似然函数方程为：

$$\ln L_{n,T}(\theta, \mathbf{c}_{nT}) = -\frac{nT}{2} \ln 2\pi - \frac{nT}{2} \ln \sigma^2 + T \ln |S_n(\lambda)| - \frac{1}{2\sigma^2} \sum_{t=1}^T \mathbf{V}'_{nT}(\zeta) \mathbf{V}_{nT}(\zeta) \quad (7)$$

其中 $\mathbf{V}_{nT}(\zeta) = S_n(\lambda) \mathbf{Y}_{nT} - \mathbf{Z}_{nT} \delta - \mathbf{c}_n - \eta_t l_n$ 。对这个方程取一阶条件，我们得到集中的似然方程（concentrated likelihood function）：

$$\ln L_{n,T}(\theta) = -\frac{nT}{2} \ln 2\pi - \frac{nT}{2} \ln \sigma^2 + T \ln |S_n(\lambda)| - \frac{1}{2\sigma^2} \sum_{t=1}^T \tilde{\mathbf{V}}'_{nT}(\zeta) \tilde{\mathbf{V}}_{nT}(\zeta) \quad (8)$$

其中 $\tilde{\mathbf{V}}_{nT}(\zeta) = S_n(\lambda) \tilde{\mathbf{Y}}_{nT} - \tilde{\mathbf{Z}}_{nT} \delta$ ，而 $\mathbf{Y}_{nT} = \mathbf{Y}_{nT} - \frac{1}{T} \sum_{s=1}^T \mathbf{Y}_{ns}$ 是对时间均值的偏离。QMLE 得到的

$\hat{\theta}_{nT}$ 最大化集中的似然函数方程，即公式（7）。正如 Lee and Yu（2010）中提到的，由于 n 和 T 都比较大，根据渐进性质，我们有：

$$\begin{aligned} & \sqrt{nT} (\hat{\theta}_{nT} - \theta_0) + \sqrt{\frac{n}{T}} \varphi_{\theta_0, nT, 1} + \sqrt{\frac{T}{n}} \varphi_{\theta_0, nT, 2} + O_p \left(\max \left(\sqrt{\frac{n}{T^3}}, \sqrt{\frac{T}{n^3}}, \sqrt{\frac{1}{T}} \right) \right) \\ & \xrightarrow{d} N \left(\mathbf{0}, \lim_{T \rightarrow \infty} \Sigma^{-1}_{\theta_0, nT} (\Sigma_{\theta_0, nT} + \Omega_{\theta_0, nT}) \Sigma^{-1}_{\theta_0, nT} \right) \end{aligned}$$

其中 $\varphi_{\theta_0, nT, 1}$ 和 $\varphi_{\theta_0, nT, 2}$ 是个体效应和时间固定效应的 $O(1)$ 阶领先偏差项（leading bias）， $\Sigma_{\theta_0, nT}$ 为信息

矩阵， $\Omega_{\theta_0, nT}$ 获取的是扰动项的非正态的特点，因此估计的偏差是 $O \left(\max \left(\frac{1}{n}, \frac{1}{T} \right) \right)$ 阶的。

$O \left(\frac{1}{T} \right)$ 阶的误差可以通过偏差校正步骤来修正，为了避免 $O \left(\frac{1}{n} \right)$ 阶的误差，我们可以使用数据

转换机制从而保持与上述 QMLE 相同的渐进效率。我们令 $\mathbf{J}_n = \mathbf{I}_n - \frac{1}{n} l_n l_n'$ ， l_n 为各个元素均为 1 的列向量。当 $W_n l_n = l_n$ 时，也即 W_n 是一个行标准化的矩阵（也即各行之和为 1）的时候，我们有

$$\mathbf{J}_n \mathbf{W}_n = \mathbf{J}_n \mathbf{W}_n \left(\mathbf{J}_n + \frac{1}{n} l_n l_n' \right) = \mathbf{J}_n \mathbf{W}_n \mathbf{J}_n$$

因为 $\mathbf{J}_n \mathbf{W}_n l_n = \mathbf{J}_n l_n = \mathbf{0}$ 。因此，

$$\mathbf{J}_n \mathbf{Y}_{nt} = \lambda_0 (\mathbf{J}_n \mathbf{W}_n) (\mathbf{J}_n \mathbf{Y}_{nt}) + \gamma_0 (\mathbf{J}_n \mathbf{Y}_{nt-1}) + \rho_0 (\mathbf{J}_n \mathbf{W}_n) (\mathbf{J}_n \mathbf{Y}_{nt-1}) + (\mathbf{J}_n \mathbf{X}_{nt}) \beta_0 + (\mathbf{J}_n \mathbf{c}_{n0}) + (\mathbf{J}_n \mathbf{V}_{nt})$$

这样我们就可以不涉及时间固定效应，而 $\mathbf{J}_n \mathbf{c}_{n0}$ 可以被解释为转换过的个体效应， \mathbf{J}_n 作为附加的信息与特征向量的子矩阵 $\mathbf{F}_{n, n-1}$ 相对应，我们记

$$\mathbf{Y}_{nt}^* = \mathbf{F}'_{n, n-1} \mathbf{J}_n \mathbf{Y}_{nt} = \mathbf{F}'_{n, n-1} \mathbf{Y}_{nt}$$

维度为 $n-1$ 。这种等价转换是为了避免转换过后的残差线性相关^①，我们有：

$$\mathbf{Y}_{nt}^* = \lambda_0 \mathbf{W}_n^* \mathbf{Y}_{nt}^* + \gamma_0 \mathbf{Y}_{nt-1}^* + \rho_0 \mathbf{W}_n^* \mathbf{Y}_{nt-1}^* + \mathbf{X}_{nt}^* \beta_0 + \mathbf{c}_{n0}^* + \mathbf{V}_{nt}^*$$

^①由于 \mathbf{J}_n 并不是满秩的，因此 $\mathbf{J}_n \mathbf{V}_{nt}$ 是线性相关的。

其中 $W_n^* = F'_{n,n-1} W_n F_{n,n-1}$, $X_{nt}^* = F'_{n,n-1} X_{nt}$, $c_{n0}^* = F'_{n,n-1} c_{n0}$, $V_{nt}^* = F'_{n,n-1} V_{nt}$ 是一个 $n-1$ 维的扰动项向量, 均值为 0 且方差矩阵为 $\sigma_0^2 I_{n-1}$ 。上式是一个典型的空间动态面板数据模型, 因此在转换之后, 观测数量为 $T(n-1)$, 比原来的样本减少了一期的数据, 同消除时间虚拟变量损失的观测数一样。这一公式对后面的估计很有作用, 因为我们可以构建关于 Y_{nt}^* 的似然函数。假设 V_{nt} 是正态分布 $N(0, \sigma_0^2 I_n)$, 转换过后的 V_{nt}^* 的分布变为 $N(0, \sigma_0^2 I_{n-1})$, 因此似然函数的对数方程为:

$$\ln L_{n,T}(\theta, c_n) = -\frac{(n-1)T}{2} \ln 2\pi - \frac{(n-1)T}{2} \ln \sigma^2 + T \ln(1-\lambda) - T |I_n - \lambda W_n| \frac{1}{2\sigma^2} \sum_{t=1}^T V_{nT}'(\zeta) V_{nT}(\zeta) \quad (9)$$

如果 V_{nt} 不是正态分布的, 那么公式 (9) 给出了 QMLE 的估计方程, Lee and Yu (2010) 证明了, QMLE 不会产生 $O\left(\frac{1}{n}\right)$ 阶的偏差。

为了估计增长模型, 我们应该使用上述提到的这些方法, D-MLE 是直接对公式 (8) 进行估计的结果, T-MLE 是对公式 (9) 得到的估计结果。比较两者, 我们知道对 D-MLE 进行偏差修正需要 n 和 T 均比较大, 而 T-MLE 只需要 T 比较大。

五、实证结果

(一) 不含空间相关性的实证结果

表 2 不含空间效应的实证结果 (4 年均值)

方程(1)				方程(2)					
方法	简单横截面		混同横截面		方法	动态面板 D-MLE		动态面板 T-MLE	
样本周期	52-08		52-08			52-08		52-08	
常数项	5.5967***	(0.67)	0.3946***	(0.06)	时期数	T=14		T=14	
$\ln(y_{i0}) (\gamma)$	0.8174	(0.14)	0.9928	(0.01)					
\bar{R}^2	0.5492		0.9683		$\ln(y_{i,t-1}) (\gamma)$	1.0661**	(0.03)	1.0661**	(0.03)
收敛速度	0.0036		0.0018		收敛速度(Φ)	-0.016		-0.016	
(Φ)									
($\gamma = e^{-\Phi}$)	(s=60)								

其中, ***代表 1% 显著性水平, **代表 5% 显著性水平, *代表 10% 显著性水平。括号内为标准误。

系数中, ρ 和 λ 是和 0 作比较, 而 γ 是和 1 作比较。

数据来自于《新中国 60 年统计资料汇编》。

之前的学者在没有考虑空间效应的情况下得到中国各省之间经济增长收敛的结论, 如林毅夫等 (1998)、刘夏明 (2004) 等等。我们首先忽略空间效应进行考察, 并且进一步加入个体效应进行讨论。

我们首先使用横截面模型对中国的区域增长进行分析。对于 1952-2008 年的样本, 我们首先使用 $\ln(y_{2008})$ 对 $\ln(y_{1952})$ 进行回归, 接着我们取 4 年均值的数据, 将 $\ln(y_{2008})$ 对 $\ln(y_{2004})$, \dots , $\ln(y_{1956})$

对 $\ln(y_{1952})$ 进行回归，通过对这些估计值求平均数来实现混同横截面回归。估计的结果在表 2 中的左栏展示。

可以看到，简单横截面回归得到的收敛速度为 0.0036。混同横截面得到的结果与简单横截面的一致，整体的收敛速度依然为正。我们发现我国各省经济增长总体而言并不是收敛的，估计的系数 γ 并不显著地小于 1。

使用固定效应模型进行估计，我们发现了更大的不同：整体样本的收敛速率变为负数，这与之前的大多数文献得到的结果不一致，说明中国区域经济增长并没有出现收敛的现象。

我们同时也对 1952+4t, 1953+4t, 1954+4t 和 1955+4t 的 4 年间隔数据进行了回归，表 3 是以 1952 年为起点的 4 年间隔数据的结果。我们发现间隔数据更加倾向于支持收敛的假说，而使用面板数据模型考察间隔数据则得到了与均值数据完全不同的结果。

表 3 不含空间效应的实证结果（4 年间隔）

方程(1)				方程(2)				
方法	简单横截面	混同横截面		方法	动态面板 D-MLE	动态面板 T-MLE		
样本周期	52-08	52-08			52-08	52-08		
常数项	6.4527***	(0.72)	0.4861***	(0.07)	时期数	T=14	T=14	
$\ln(y_{i0})(\gamma)$	0.7510*	(0.15)	0.9840	(0.01)	D-MLE			
\bar{R}^2	0.4544	0.9444		$\ln(y_{i,t-1})(\gamma)$	0.8763***	(0.03)	0.8763***	(0.03)
收敛速度	0.0048	0.0040		收敛速度	0.033	0.033		
(Φ)				(Φ)				
($\gamma=e^{-\Phi s}$)	(s=60)							

其中，***代表 1% 显著性水平，**代表 5% 显著性水平，*代表 10% 显著性水平。括号内为标准误。

系数中， ρ 和 λ 是和 0 作比较，而 γ 是和 1 做比较。

数据来自于《新中国 60 年统计资料汇编》。

这一部分的结果说明如果模型忽略了空间效应可能会得到不正确的结果，我们发现使用均值数据和间隔数据得到了截然相反的结论。这说明如果忽略各省市的空间效应，如劳动力和资本的跨区域流动、市场的整合、以及技术的转移等，估计的结果可能有偏。我们将在下一部分进一步放松假设，在横截面数据中加入空间效应来探究我国各省的经济增长是否收敛。

（二）包含空间效应的横截面回归结果

刘生龙和张捷（2009）等学者使用静态的空间计量模型考察我国各省经济增长的收敛性，他们使用包含空间效应的横截面数据进行考察，得到了我国各省之间经济增长存在一定的收敛性的结论。因此我们在这一节忽略个体效应，进一步放松假设，在横截面模型中加入空间因素来进一步验证我们的结果，看看能否得出与之前学者一致的结论。

表 4 报告了估计结果。首先，如果不考虑空间效应的滞后项，简单的横截面回归结果告诉我们 1952-2008 年间经济增长是收敛的，收敛速度为 0.0069，而用混同横截面得到的结果是不收敛的，系数 (γ) 并不是显著地小于 1。接着我们加入空间效应的滞后项，结果基本一致。我们使用简单横截面回归得到收敛的结论，而用混同横截面回归却得到不收敛的结果。就收敛的速度而言，是否加入空间滞后项的影响很小，对于简单横截面得到的收敛速度是 0.0065，而混同横截面则得到 0.0031。

表 4 包含空间效应的横截面回归（4 年均值）

不含滞后项，即方程(3)	简单横截面	1952-2008	混同横截面	1952-2008
常数项	2.3366	(1.46)	0.2618**	(0.11)

$\ln(y_{i0}) (\gamma)$	0.6811***	(0.13)	0.9875	(0.01)
$W_n \ln(y_{it})(\lambda)$	0.41**	(0.17)	0.0174	(0.02)
\bar{R}^2	0.5225		0.9698	
收敛速度(Φ)	0.0069		0.0032	
$(\gamma=e^{-\Phi_s})$	(s=60)			
包含滞后项, 即方程(4)	简单横截面	1952-2008	混同横截面	1952-2008
常数项	2.5249*	(1.47)	0.3047***	(0.11)
$\ln(y_{i0}) (\gamma)$	0.735*	(0.14)	0.9878	(0.01)
$W_n \ln(y_{i0})(\rho)$	-0.2785	(0.30)	0.1378*	(0.08)
$W_n \ln(y_{it})(\lambda)$	0.504**	(0.20)	-0.1196	(0.07)
\bar{R}^2	0.5276		0.9685	
收敛速度(Φ)	0.0065		0.0031	
$(\gamma=e^{-\Phi_s})$	(s=60)			

其中, ***代表 1% 显著性水平, **代表 5% 显著性水平, *代表 10% 显著性水平。括号内为标准误。

系数中, ρ 和 λ 是和 0 作比较, 而 γ 是和 1 做比较。

数据来自于《新中国 60 年统计资料汇编》。

在简单横截面回归中, 空间效应的当期项是显著的。在混同横截面模型中, 空间效应的当期项不显著, 但是滞后项是显著的。总的说来, 空间效应并不总是显著, 但是也不能排除它的影响。

我们同时使用以 1952、1953、1954 和 1955 为起点的间隔数据考察收敛性。表 5 汇报了以 1952 年为起点的实证结果, 我们发现使用简单横截面数据得到的结果和均值数据的很接近, 并且收敛速度 0.009 比使用均值数据的 0.007 更快。而混同横截面数据模型得到的 γ 在 1% 水平下显著地小于 1, 并且收敛速度 0.006 也大于使用均值数据得到的 0.003。总体而言, 间隔数据更加倾向于支持收敛的假说, 但是忽略了个体效应得到的结果依然不是很稳健, 我们依旧无法得到一致的结论。这说明忽略各省的资源禀赋和文化差异等因素, 也会得到有偏的结果, 因此我们应该同时考虑个体效应和空间效应, 使用空间动态面板数据模型进行估计。

表 5 包含空间效应的横截面回归 (4 年间隔)

不含滞后项, 即方程(3)	简单横截面	1952-2008	混同横截面	1952-2008
常数项	3.0605*	(1.70)	0.2950**	(0.14)
$\ln(y_{i0}) (\gamma)$	0.5887***	(0.15)	0.9766*	(0.01)
$W_n \ln(y_{it})(\lambda)$	0.4160**	(0.19)	0.0286	(0.02)
\bar{R}^2	0.4611		0.9465	
收敛速度(Φ)	0.0088		0.0059	
$(\gamma=e^{-\Phi_s})$	(s=60)			
包含滞后项, 即方程(4)	简单横截面	1952-2008	混同横截面	1952-2008
常数项	3.14*	(1.72)	0.2634*	(0.14)
$\ln(y_{i0}) (\gamma)$	0.6001**	(0.17)	0.9775*	(0.01)
$W_n \ln(y_{i0})(\rho)$	-0.0369	(0.35)	0.0685	(0.08)
$W_n \ln(y_{it})(\lambda)$	0.4200*	(0.22)	-0.0292	(0.07)
\bar{R}^2	0.4420		0.9446	
收敛速度(Φ)	0.0085		0.0057	
$(\gamma=e^{-\Phi_s})$	(s=60)			

其中, ***代表 1% 显著性水平, **代表 5% 显著性水平, *代表 10% 显著性水平。括号内为标准误。

系数中， ρ 和 λ 是和 0 作比较，而 γ 是和 1 做比较。

数据来自于《新中国 60 年统计资料汇编》。

（三）空间动态面板数据模型的实证结果

在表 2-3 中，我们先忽略了空间效应。表 3 汇报的估计结果与文献是一致的，说明忽略了空间效应，如劳动力和资本的跨区流动、市场的整合以及技术的转移等可能会使得估计有偏。在表 3-6 中我们加入了空间效应但是忽略了个体效应，但是由于各省市之间的资源禀赋和文化差异，忽略个体效应也会带来估计结果的偏差。另外，如果忽略了个体效应，空间的互动性可能扮演着重要的角色，因此所估计的空间效应的系数也可能是有偏的，更容易出现在统计上显著为正的现象。

在这一节里，我们进一步放松假设，同时考虑个体效应和空间效应，以避免错误的模型设定（misspecification）带来的问题。我们整合原有的数据，使用空间动态面板数据模型来考察各省经济增长是否收敛的问题。基于方程（5）估计的参数值和标准误汇报于表 6。

我们首先关注 γ （即 $\ln(y_{it-1})$ 项的系数）的估计值。从表 6 中可以看出，对于 1952-2008 年的 4 年均值样本而言，各种方法估计出来的 γ 和 1 都非常接近，并且都不显著地大于 1 或者小于 1。在增长理论中，如果 γ 小于 1，那么我们可以得到增长收敛的结论，即落后地区的经济增长速度相对富裕地区更快。我们的估计结果表明，中国各省市的经济并不具有收敛性，即落后地区的经济增长并不比富裕地区快。对于以 1952, 1953、1954、1955 为起始点的每 4 年间值的估计， γ 值也大多数很接近于 1，并且在统计上不显著，这和 4 年均值样本的估计结果是一致的。唯一的例外是，以 1952 年为起始点的每 4 年间隔数据的估计表明中国各省经济是显著收敛的。因此我们的结论是，中国各省经济的发展不存在收敛性。当然，由于 γ 并不显著性的大于 1，我们的结果也不支持发散的假说，即越富裕的地区经济增长越快。

其次我们看 λ 的估计值：使用 4 年间隔数据得到的结果大多支持空间相关性，但是用均值样本得到的结果不支持空间相关性。因此，我们认为存在一定的空间相关性，但不是很强，对空间滞后项的系数 ρ 的估计也是如此。在 Wald 检验中，我们发现多数情况下都无法拒绝 $\rho+\gamma\lambda=0$ 的原假设。

我们得到的结果与之前学者得到的不一致，但是我们在使用与之前学者一致的方法也同样可以得到中国各省经济增长收敛的结论，这说明我们的数据与之前学者是基本一致的。但是我们的模型在同时考虑了空间效应和个体效应之后，避免了错误模型设定的问题，是比较稳健的模型，因此估计结果也更加可信，与此同时，我们的结果说明了忽略了空间效应或者个体效应，都可能会得到错误的结果。

那么为什么我国各省经济增长没有出现收敛呢？从表 6 的结果来看，我们认为一个可能的解释是地区之间互动性比较弱，另一种可能的解释则是林毅夫等（1998）、刘夏明等（2004）提出的我国区域发展不平衡导致。我们将通过对东、中、西部的子样本进行回归来一探究竟。

表 6 空间动态面板

使用相邻矩阵	52-08 (均值)		52-08 (间隔)		53-05 (间隔)		54-06 (间隔)		55-07 (间隔)	
时期数	T=14		T=15		T=14		T=14		T=14	
D-MLE										
$\ln(y_{it-1}) (\gamma)$	1.0529	(0.04)	0.8637***	(0.04)	0.9791	(0.03)	0.9513	(0.03)	0.9708	(0.04)
$W\ln(y_{it-1})(\rho)$	-0.0238	(0.33)	-0.0754	(0.10)	0.0086	(0.13)	-0.149	(0.11)	-0.261**	(0.11)
$W\ln(y_{it})(\lambda)$	0.0737	(0.30)	0.1526*	(0.08)	0.0736	(0.12)	0.1787*	(0.10)	0.269***	(0.09)
收敛速度(ϕ)	-0.0129		0.0366		0.0053		0.0125		0.0074	
Wald $\rho+\gamma\lambda=0$	1.3377		0.9815		2.1165		0.1496		0	

T-MLE										
$\ln(y_{i,t-1}) (\gamma)$	1.0317	(0.03)	0.8635	*** (0.04)	0.9761	(0.03)	0.9524	(0.03)	0.9716	(0.04)
$Wnln(y_{i,t-1})(\rho)$	-0.0258	(0.10)	-0.091	(0.09)	0.0074	(0.10)	-0.1717*	(0.10)	-0.279***	(0.09)
$Wnln(y_{it})(\lambda)$	0.0669	(0.08)	0.1516*	(0.08)	0.0492	(0.08)	0.188**	(0.08)	0.272***	(0.08)
收敛速度(Φ)	-0.0078		0.0367		0.0061		0.0122		0.0072	
Wald $\rho+\gamma\lambda=0$	0.8990		0.5008		1.0195		0.0192		0.0712	

其中，***代表 1% 显著性水平，**代表 5% 显著性水平，*代表 10% 显著性水平。括号内为标准误。

系数中， ρ 和 λ 是和 0 作比较，而 γ 是和 1 做比较。

数据来自于《新中国 60 年统计资料汇编》。

（四） 各个区域内经济增长收敛吗？

表 6 说明中国各省的人均 GDP 总体上并不是收敛的。并且各省经济有一定的空间相关性，但不是非常强。由于中国的东、中和西部的经济发展不平衡，因此很有必要对这三个地区的经济增长和空间相关性分别作研究。

表 7 给出了对这三个地区分别做回归的结果。我们按照陈安平和李国平（2004）的划分方法，对于原样本进行区域划分，其中，东部区域包括北京、天津、河北、广东、山东、福建、浙江、江苏、上海、辽宁和广西，中部地区包括山西、内蒙古、吉林、黑龙江、安徽、江西、河南、湖北和湖南，西部地区包括四川、贵州、云南、西藏、重庆、陕西、甘肃、青海、宁夏和新疆^①。

从表 7 中可知，对东部地区进行回归的结果不支持收敛的假设，而且一些均值样本的回归甚至支持发散。从空间效应上，我们可以看到东部地区各省互动性较强，空间相关性系数显著。相反的是，中部地区和西部地区的经济存在收敛性，但是总体上讲空间效应不显著。

表 7 中的回归结果包含了丰富的信息。首先，和现存的部分文献（见段平忠，2008）一致，中国各省经济呈现出一定的“俱乐部收敛”的特征。更重要的是，中西部地区的俱乐部成员之间空间相关性不显著，表明中国经济的“俱乐部收敛”并非由于俱乐部成员之间的较强的经济互动关系所引起。因此在未来的研究中，有必要考察自然条件、地理特征和国家扶持的因素。另外，东部地区各省尽管存在较强的经济互动关系，但是人均 GDP 并没有收敛的趋势，说明东部各省之间存在着很强的引发经济发散的因素，例如虹吸效应或者工业聚集效应。

表 7 空间动态面板：子样本数据

	52-08 (均值)	52-08 (间隔)	53-05 (间隔)	54-06 (间隔)	55-07 (间隔)					
时期数	$T=14$	$T=15$	$T=14$	$T=14$	$T=14$					
区域	东部地区									
D-MLE										
$\ln(y_{i,t-1}) (\gamma)$	1.7561***	(0.11)	0.8795**	(0.06)	1.2877***	(0.05)	1.6972***	(0.08)	1.0539	(0.06)
$Wnln(y_{i,t-1})(\rho)$	-1.395***	(0.44)	-0.207	(0.13)	-0.488***	(0.18)	-1.539***	(0.19)	-0.413***	(0.13)
$Wnln(y_{it})(\lambda)$	0.138	(0.35)	0.2047**	(0.09)	0.1432	(0.13)	0.1619	(0.15)	0.264***	(0.09)
收敛速度(Φ)	-0.1408		0.0321		-0.0632		-0.1323		-0.0131	
Wald $\rho+\gamma\lambda=0$	-43.3498		0.0669		13.3257		95.6895		1.91	

T-MLE

^①我们的划分方法与陈安平和李国平（2004）的划分并不完全一致，上述的文章由于重庆和西藏数据的缺失问题而放弃了这两个省级行政单位，而本文使用的《新中国 60 年统计资料汇编》则没有这个问题，因此我们将重庆和西藏包含在分区域的样本当中，并且根据它们的地理位置将它们划分到了西部。

$\ln(y_{i,t-1}) (\gamma)$	1.8522*** (0.05)	0.9154 (0.06)	1.4482*** (0.05)	1.7286*** (0.05)	1.0855 (0.07)
$Wnln(y_{i,t-1}) (\rho)$	-1.011*** (0.23)	-0.0364 (0.15)	-0.119 (0.20)	-0.792*** (0.22)	-0.3092 (0.19)
$Wnln(y_{it}) (\lambda)$	0.2687** (0.14)	0.3128 *** (0.09)	0.2843*** (0.11)	0.376*** (0.12)	0.4714*** (0.09)
收敛速度(Φ)	-0.1541	0.0221	-0.0926	-0.1368	-0.0205
Wald $\rho+\gamma\lambda=0$	37.1901	5.4171	11.6811	2.5909	51.3259
区域	中部地区				
D-MLE					
$\ln(y_{i,t-1}) (\gamma)$	0.9029* (0.06)	0.826*** (0.07)	0.8261*** (0.07)	0.7176*** (0.06)	0.9154 (0.07)
$Wnln(y_{i,t-1}) (\rho)$	-0.1297 (0.13)	0.006 (0.14)	-0.172 (0.13)	-0.144 (0.11)	-0.1463 (0.15)
$Wnln(y_{it}) (\lambda)$	0.1437 (0.12)	0.1429 (0.11)	0.1441 (0.11)	0.1443 (0.11)	0.144 (0.11)
收敛速度(Φ)	0.0255	0.0478	0.0478	0.0829	0.0221
Wald $\rho+\gamma\lambda=0$	0	1.4926	0.3039	0.2317	0.0186
T-MLE					
$\ln(y_{i,t-1}) (\gamma)$	1.0117 (0.06)	0.8495** (0.08)	0.8298** (0.07)	0.7213*** (0.07)	1.125 (0.08)
$Wnln(y_{i,t-1}) (\rho)$	0.0785 (0.18)	0.1137 (0.15)	-0.1047 (0.14)	-0.0743 (0.13)	0.234 (0.23)
$Wnln(y_{it}) (\lambda)$	0.1728 (0.14)	0.0476 (0.12)	0.0036 (0.13)	0.0015 (0.14)	0.2783** (0.14)
收敛速度(Φ)	-0.0029	0.0408	0.0466	0.0817	-0.0295
Wald $\rho+\gamma\lambda=0$	10.3958	2.4163	1.1864	0.7859	24.0551
区域	西部地区				
D-MLE					
$\ln(y_{i,t-1}) (\gamma)$	0.8404** (0.07)	0.7737*** (0.07)	0.7423*** (0.08)	0.6929*** (0.08)	0.7731*** (0.07)
$Wnln(y_{i,t-1}) (\rho)$	-0.2439 (0.15)	-0.1902 (0.15)	-0.1176 (0.17)	-0.3453** (0.16)	-0.3637** (0.15)
$Wnln(y_{it}) (\lambda)$	0.5516*** (0.08)	0.5531*** (0.08)	0.5524*** (0.08)	0.5485*** (0.08)	0.5489*** (0.08)
收敛速度(Φ)	0.0435	0.0641	0.0745	0.0917	0.0643
Wald $\rho+\gamma\lambda=0$	2.1997	2.5751	2.6063	0.0462	0.171
T-MLE					
$\ln(y_{i,t-1}) (\gamma)$	0.8127*** (0.07)	0.7571*** (0.07)	0.7118*** (0.08)	0.6706*** (0.08)	0.752*** (0.07)
$Wnln(y_{i,t-1}) (\rho)$	0.0378 (0.19)	0.0772 (0.18)	0.1133 (0.20)	-0.1641 (0.19)	-0.1643 (0.19)
$Wnln(y_{it}) (\lambda)$	0.0028 (0.17)	0.0027 (0.16)	0.003 (0.16)	0.1005 (0.18)	0.1189 (0.17)
收敛速度(Φ)	0.0518	0.0696	0.085	0.0999	0.0713
Wald $\rho+\gamma\lambda=0$	0.0979	0.3599	0.5648	0.4369	0.3166

其中，***代表 1% 显著性水平，**代表 5% 显著性水平，*代表 10% 显著性水平。括号内为标准误。

系数中， ρ 和 λ 是和 0 作比较，而 γ 是和 1 作比较。

数据来自于《新中国 60 年统计资料汇编》。

六、稳健性检验

人们可能会担心之前没有控制变量的估计结果会有偏差，我们按照新古典理论，选取人口增长率 ($popgr$) 和储蓄/投资率 (inv)^① 做控制变量进行研究。结果汇报于表 8 和表 9 中。

^① 在实际操作中，由于广东省 1952-1977 年大部分、重庆市 1952-1955 年、宁夏回族自治区 1952-1977 年大部分以及西藏自治区 1952-1958 年的储蓄数据缺失，我们使用固定资产投资额 (inv) 作为替代。由于新古典理论解释经济增长的一个重要的原因在于储蓄，而储蓄影响经济增长的方式是投资，因此我们使用投资率来替代储蓄率。与此同时，

我们在表 8 中汇报了使用空间动态面板模型进行回归的结果，得到了和表 6 基本一致的估计：1952-2008 年的四年均值样本和以 1952、1953、1954 以及 1955 为起始点的四年间隔样本得到的对 γ 的估计值均和 1 比较接近，并且大部分都不显著，这进一步佐证了我国各省市的经济并不具有收敛性的结论。而我们得到的大部分结果也说明 λ 是显著的，也即存在空间相关性，但这一效应不是很强，而对 ρ 的估计大部分都不显著。在 Wald 检验中，我们发现在多数情况下都无法拒绝 $\rho+\gamma\lambda=0$ 的原假设。另外，控制变量中，大部分结果表明人口增长率负面影响经济增长率而储蓄/投资率正面影响经济增长率，与传统的新古典模型的预测一致。表 8 说明我们对省际经济增长的收敛性和空间相关性的实证结果是稳健的。

表 8 空间动态面板（包含控制变量）

使用相邻矩阵	52-08（均值）		52-08（间隔）		53-05（间隔）		54-06（间隔）		55-07（间隔）	
时期数	T=14		T=15		T=14		T=14		T=14	
D-MLE										
$\ln(y_{i,t-1}) (\gamma)$	1.0574*	(0.03)	0.8661***	(0.04)	0.9888	(0.03)	0.9609	(0.03)	0.9363*	(0.03)
$W\ln(y_{i,t-1})(\rho)$	-0.0712	(0.23)	-0.0817	(0.10)	-0.0052	(0.13)	-0.1352	(0.11)	-0.3008***	(0.10)
$W\ln(y_{it})(\lambda)$	0.0659	(0.21)	0.1348*	(0.09)	0.0715	(0.12)	0.1745*	(0.10)	0.3097***	(0.08)
<i>popgr</i>	0.0117	(0.37)	0.2493	(0.23)	-0.9203**	(0.45)	-1.3083**	(0.55)	0.8263***	(0.22)
<i>inv</i>	0.1751***	(0.06)	0.1036	(0.08)	0.1315	(0.09)	0.0691	(0.08)	0.0359	(0.03)
收敛速度(Φ)	-0.014		0.0359		0.0028		0.01		0.0165	
Wald $\rho+\gamma\lambda=0$	0.001		0.373		1.392		0.3533		0.0417	
T-MLE										
$\ln(y_{i,t-1}) (\gamma)$	1.0455	(0.03)	0.8662***	(0.04)	0.9887	(0.03)	0.9616	(0.03)	0.936*	(0.03)
$W\ln(y_{i,t-1})(\rho)$	-0.0782	(0.10)	-0.0982	(0.09)	-0.0035	(0.10)	-0.1523*	(0.10)	-0.3172***	(0.09)
$W\ln(y_{it})(\lambda)$	0.073	(0.08)	0.1376*	(0.08)	0.0444	(0.08)	0.1755**	(0.08)	0.3137***	(0.08)
<i>Popgr</i>	0.0418	(0.37)	0.2275	(0.22)	-0.9268**	(0.44)	-1.2518**	(0.55)	0.8639***	(0.22)
<i>inv</i>	0.1771***	(0.06)	0.1033	(0.08)	0.1362	(0.09)	0.0683	(0.09)	0.0374	(0.03)
收敛速度(Φ)	-0.0111		0.0359		0.0028		0.0098		0.0165	
Wald $\rho+\gamma\lambda=0$	0.0016		0.1358		0.5367		0.0925		0.2026	

其中，***代表 1% 显著性水平，**代表 5% 显著性水平，*代表 10% 显著性水平。括号内为标准误。

系数中， ρ 和 λ 是和 0 作比较，而 γ 是和 1 做比较。

数据来自于《新中国 60 年统计资料汇编》。

接着我们按照东、中、西部划分子样本，同样加入控制变量进行估计，结果汇报于表 9 当中。我们得到的估计值无论在大小还是在显著性上，都与表 7 的结果基本一致：东部地区并没有出现收敛，甚至大多数的样本结果支持发散的说话。空间效应方面，我们可以看到东部地区各省互动性较强，空间相关性系数显著。与之相反，中西地区的经济均存在收敛性，但总体而言各自的空间效应不显著。这说明我们子样本的估计结果也依然是稳健的。

表 9 空间动态面板：子样本数据（包含控制变量）

	52-08（均值）		52-08（间隔）		53-05（间隔）		54-06（间隔）		55-07（间隔）	
时期数	T=14		T=15		T=14		T=14		T=14	
区域	东部地区									

D-MLE

我们对储蓄和投资进行了 Pearson 相关系数检验，发现两者之间的相关系数为 0.7912 ($p<0.0001$)，而使用两者生成的储蓄率和投资率之间的相关系数也有 0.5280 ($p<0.0001$)，因此我们使用投资率代替储蓄率的做法是合理的。

$\ln(y_{i,t-1}) (\gamma)$	1.6756** (0.31)	0.8915 * (0.06)	1.3414*** (0.05)	1.4625*** (0.06)	0.9932 (0.06)
$Wnln(y_{i,t-1})(\rho)$	-1.1941 (1.45)	-0.2105 (0.13)	-0.612*** (0.18)	-1.049*** (0.15)	-0.379*** (0.13)
$Wnln(y_{it})(\lambda)$	0.1654 (1.14)	0.2131** (0.09)	0.1279 (0.14)	0.2049* (0.11)	0.3436*** (0.08)
<i>popgr</i>	0.7062 (0.46)	0.1574 (0.30)	-1.4359* (0.80)	-0.0605 (0.86)	0.9921 (0.26)
<i>inv</i>	0.3064 (0.34)	0.1506 (0.19)	0.0878 (0.17)	0.1775 (0.18)	-0.1096 (0.18)
收敛速度(Φ)	-0.129	0.0287	-0.0734	-0.095	0.0017
Wald $\rho+\gamma\lambda=0$	3.1235	0.0389	26.3595	58.5604	0.1753

T-MLE

$\ln(y_{i,t-1}) (\gamma)$	1.7931*** (0.05)	0.9291 (0.06)	1.4975*** (0.05)	1.598*** (0.05)	1.0728 (0.07)
$Wnln(y_{i,t-1})(\rho)$	-0.833*** (0.24)	-0.031 (0.15)	-0.2573 (0.20)	-0.500** (0.22)	0.3147 (0.20)
$Wnln(y_{it})(\lambda)$	0.3014** (0.13)	0.3122 *** (0.09)	0.2725** (0.12)	0.410*** (0.12)	0.545*** (0.09)
<i>popgr</i>	0.8845* (0.46)	0.1323 (0.30)	-0.9956 (0.81)	0.4471 (0.86)	1.154*** (0.26)
<i>inv</i>	0.3344** (0.17)	0.2031 (0.19)	0.0645 (0.18)	0.1752 (0.18)	-0.097*** (0.18)
收敛速度(Φ)	-0.146	0.0184	-0.101	-0.1172	-0.0176
Wald $\rho+\gamma\lambda=0$	13.1167	5.7977	3.1664	3.0354	63.2625

区域

中部地区

D-MLE

$\ln(y_{i,t-1}) (\gamma)$	0.9019* (0.06)	0.8607* (0.08)	0.8077*** (0.07)	0.7294*** (0.07)	0.82** (0.08)
$Wnln(y_{i,t-1})(\rho)$	-0.1198 (0.14)	0.049 (0.15)	-0.1959 (0.13)	-0.1303 (0.11)	-0.2015 (0.14)
$Wnln(y_{it})(\lambda)$	0.1386 (0.11)	0.1433 (0.11)	0.1375 (0.11)	0.1448 (0.11)	0.1389 (0.11)
<i>popgr</i>	-0.634 (1.17)	-0.6359 (0.57)	0.2134 (1.05)	-0.9455 (1.03)	1.6929 (1.15)
<i>inv</i>	0.1865 (0.16)	0.0635 (0.16)	0.3692** (0.16)	-0.0011 (0.17)	0.0155 (0.16)
收敛速度(Φ)	0.0258	0.0375	0.0534	0.0789	0.0496
Wald $\rho+\gamma\lambda=0$	0.0032	2.3937	0.7866	0.0834	0.6307

T-MLE

$\ln(y_{i,t-1}) (\gamma)$	1.0335 (0.07)	0.8981 (0.08)	0.812*** (0.07)	0.7324*** (0.07)	0.8234** (0.08)
$Wnln(y_{i,t-1})(\rho)$	0.2739 (0.17)	0.1745 (0.16)	-0.1279 (0.14)	-0.0617 (0.13)	-0.1364 (0.16)
$Wnln(y_{it})(\lambda)$	0.1392*** (0.14)	0.0487 (0.12)	0.0011 (0.13)	0.002 (0.14)	0.0079 (0.13)
<i>popgr</i>	-3.987*** (1.20)	-0.8472 (0.58)	0.234 (1.06)	-0.8937 (1.04)	1.8407 (1.16)
<i>inv</i>	0.0744 (0.16)	0.0584 (0.16)	0.3647** (0.16)	-0.0003 (0.17)	0.0145 (0.16)
收敛速度(Φ)	-0.0082	0.0269	0.0522	0.0779	0.0486
Wald $\rho+\gamma\lambda=0$	21.5116	4.1428	1.881	0.5175	1.5067

区域

西部地区

D-MLE

$\ln(y_{i,t-1}) (\gamma)$	0.8585** (0.07)	0.7666*** (0.07)	0.7474*** (0.08)	0.6981*** (0.08)	0.7379*** (0.07)
$Wnln(y_{i,t-1})(\rho)$	-0.2923 (0.15)	-0.2069 (0.15)	-0.1267 (0.18)	-0.3153** (0.16)	-0.441*** (0.15)
$Wnln(y_{it})(\lambda)$	0.5472*** (0.08)	0.5512*** (0.08)	0.5512*** (0.08)	0.549*** (0.08)	0.5428*** (0.08)
<i>popgr</i>	0.3577 (0.95)	0.7199 (0.82)	-0.3953 (0.75)	-1.5553 (0.98)	1.2957* (0.68)
<i>inv</i>	0.1772** (0.08)	0.0256 (0.13)	0.0868 (0.18)	0.0911 (0.15)	0.0231 (0.03)
收敛速度(Φ)	0.0382	0.0664	0.0728	0.0899	0.076
Wald $\rho+\gamma\lambda=0$	1.4799	2.1296	2.4477	0.1784	0.0768

T-MLE

$\ln(y_{i,t-1}) (\gamma)$	0.8294** (0.07)	0.7527*** (0.07)	0.7176*** (0.08)	0.6736*** (0.08)	0.7256*** (0.07)
$W\ln(y_{i,t-1}) (\rho)$	0.0041 (0.19)	0.0655 (0.18)	0.1101 (0.20)	-0.1405 (0.19)	-0.2563 (0.19)
$W\ln(y_{it}) (\lambda)$	-0.0003 (0.17)	0.0014*** (0.15)	0.002 (0.16)	0.1056 (0.17)	0.1701 (0.17)
$popgr$	0.5165 (0.91)	0.7749 (0.79)	-0.2836 (0.71)	-1.4024 (0.97)	1.252* (0.69)
inv	0.1327** (0.08)	0.0365 (0.13)	0.0712 (0.17)	0.0737 (0.15)	0.0154 (0.03)
收敛速度(Φ)	0.0468	0.071	0.083	0.0988	0.0802
Wald $\rho+\gamma\lambda=0$	0.0009	0.2555	0.5205	0.2258	0.9701

其中，***代表 1% 显著性水平，**代表 5% 显著性水平，*代表 10% 显著性水平。括号内为标准误。

系数中， ρ 和 λ 是和 0 作比较，而 γ 是和 1 做比较。

数据来自于《新中国 60 年统计资料汇编》。

七、 结论、讨论和政策建议

本文分析了中国各省人均 GDP 的收敛性问题，这是一个关系到一个国家或区域经济发展是否平衡、收入分配是否公平的重大问题。本文的创新之处在于，我们运用了空间动态面板数据模型，着重强调了各省经济在空间上的相关性，同时也考虑了各省资源禀赋和文化差异等因素，这样的估计结果更加可靠。

我们发现中国各省经济在总体上并不是收敛的，各省之间存在着一定的经济互动，但不是很强。较弱的空间相关性可以部分解释为什么总体上各省经济不收敛，而不收敛的另一个原因是东、中、西部经济发展不平衡。我们对东、中、西部三个地区分别进行回归，发现东部地区各省经济没有收敛性，但是中部和西部各省经济呈收敛性。换言之，中国经济发展部分呈现出“俱乐部收敛”的特征。

在分地区回归分析中，我们发现中部和西部各省的空间相关性不显著。因此，中西部的“俱乐部收敛”并不是因为中西部各省经济上互动关系较强导致的，需要从地理特征、自然资源、国家扶持政策等其他因素去找原因。而东部地区各省尽管在发展上没有呈现出收敛性，但是空间相关性却很显著，说明在东部地区存在导致经济增长发散的因素，例如虹吸效应和工业聚集效应等等。

本文的研究结果对于经济政策有一定的启示作用，面对总体上经济缺乏收敛性的事实，政府可以采取降低省际技术、贸易壁垒，推进人才、技术和资本的流动，从而增强各省经济的有效互动，达到缩小省际经济水平差距的目的。考虑到东部地区各省经济互动强，但是缺乏收敛性，政策制定者需要分析原因并提出相应的解决方案。中西部地区俱乐部收敛特征明显，有可能是国家对中西部落后地区增加人力资本投入、扩大对外开放、提升工业化水平等扶持政策所致。进一步分析这一类因素将有助于政府改进扶持政策，以期收到更好的效果。

参考文献

- [1] Barro, R. and X. Sala-i-Martin, “Convergence”, *Journal of Political Economy*, 1992, 100(2), 223-251.
- [2] 陈安平、李国平, “中国地区经济增长的收敛性: 时间序列的经验研究”, 《数量经济技术经济研究》, 2004 年第 11 期, 第 31-35 页。
- [3] 段平忠, “我国人口流动对区域经济增长收敛效应的影响”, 《人口与经济》, 2008 年第 4 期, 第 1-5 页。
- [4] Islam, N., “Growth Empirics: A Panel Data Approach”, *The Quarterly Journal of Economics*, 1995,

- 110(4), 1127-1170.
- [5] Lee, L., and J. Yu, “A Spatial Dynamic Panel Data Model With Both Time and Individual Fixed Effects”, *Econometric Theory*, 2010, 26(2), 564-597.
- [6] 林毅夫、蔡昉、李周, “中国经济转型时期的地区差距分析”, 《经济研究》, 1998 年第 6 期, 第 3-10 页。
- [7] 刘生龙、张捷, “空间经济视角下中国区域经济收敛性再检验——基于 1985-2007 年省级数据的实证研究”, 《财经研究》, 2009 年第 12 期, 第 16-26 页。
- [8] 刘夏明、魏英琪、李国平, “收敛还是发散? ——中国区域经济发展争论的文献综述”, 《经济研究》, 2004 年第 7 期, 第 70-81 页。
- [9] Lucas Jr, R., “On the Mechanics of Economic Development”, *Journal of Monetary Economics*, 1988, 22(1), 3-42.
- [10] Sachs, J. and A. Warner, “Fundamental Sources of Long-run Growth”, *The American Economic Review*, 1997, 87(2), 184-188.
- [11] Solow, R., “A Contribution to the Theory of Economic Growth”, *The Quarterly Journal of Economics*, 1956, 70(1), 65-94.
- [12] Yu, J. and L. Lee, “Convergence: A Spatial Dynamic Panel Data Approach”, *Global Journal of Economics*, 2012, 1(1), 1-37.

Is Provincial Economic Growth Convergent in China?

Abstract This paper studies the convergence of provincial economies in China by employing Spatial Dynamic Panel Data model and GDP data from 1952 to 2008. We find that there is no overall growth convergence among provinces in China and the spatial correlation is weak. We further study “club convergence” by grouping provinces into eastern, middle, and western regions respectively and find that provinces in eastern region exhibit divergence with overall significant spatial correlation while those in other two regions show convergence with little spatial correlation. This suggests that the lack of convergence is partially caused by the unbalanced development across regions.

Key Words Growth Convergence; Spatial Dynamic Panel Data Model; Spatial Econometrics

JEL: C33; E01; E13;