

数据空间关系的构建与表达

2019 Annual Stata Conference

2019.06-16

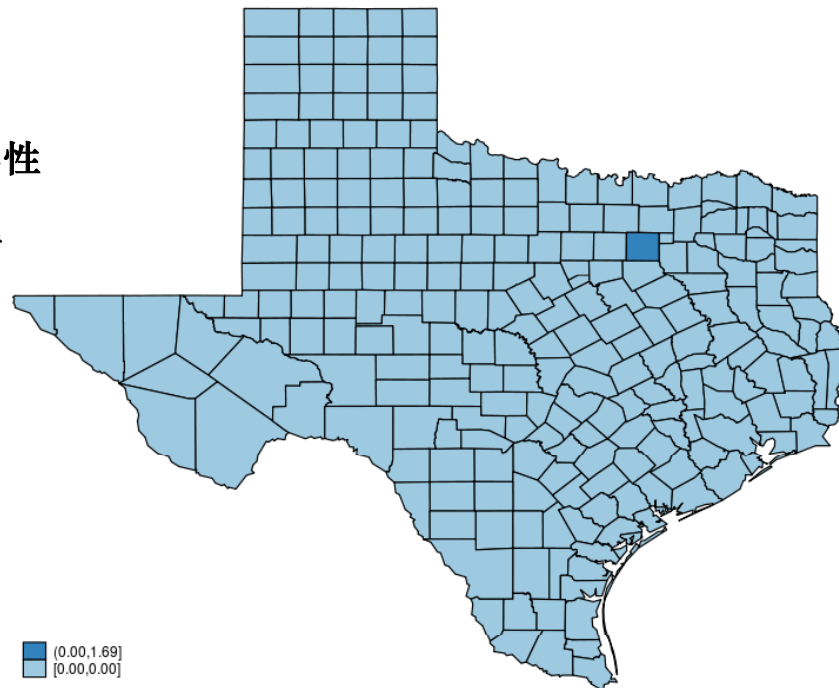
肖光恩

武汉大学经济与管理学院

Email: xge2000@163.com

一、数据空间关系的探索是当前计量经济学发展的一个重要方向

- 地理区位的定义
- 地理数据与数据的地理属性
- 数据地理属性规律的探索

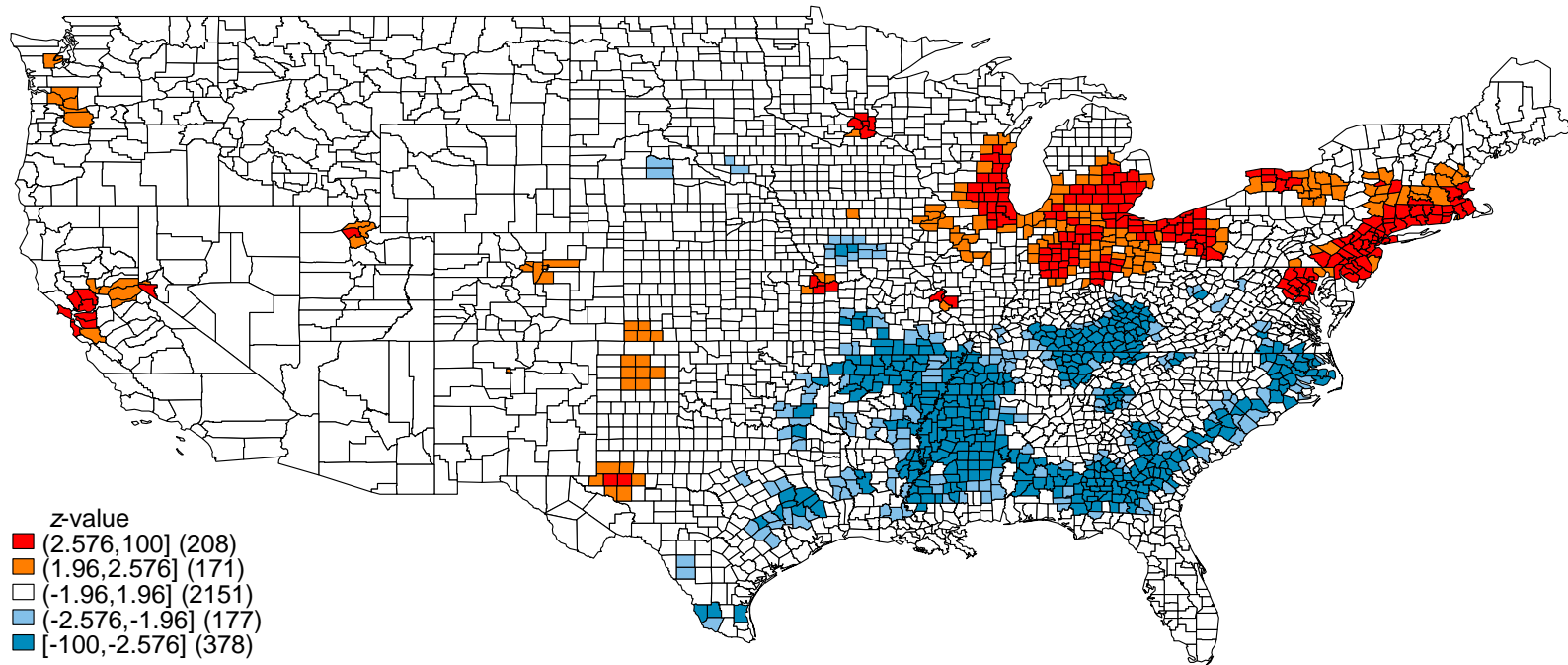


一、数据空间关系的探索是当前计量经济学发展的一个重要方向

—— (1) 理论探索

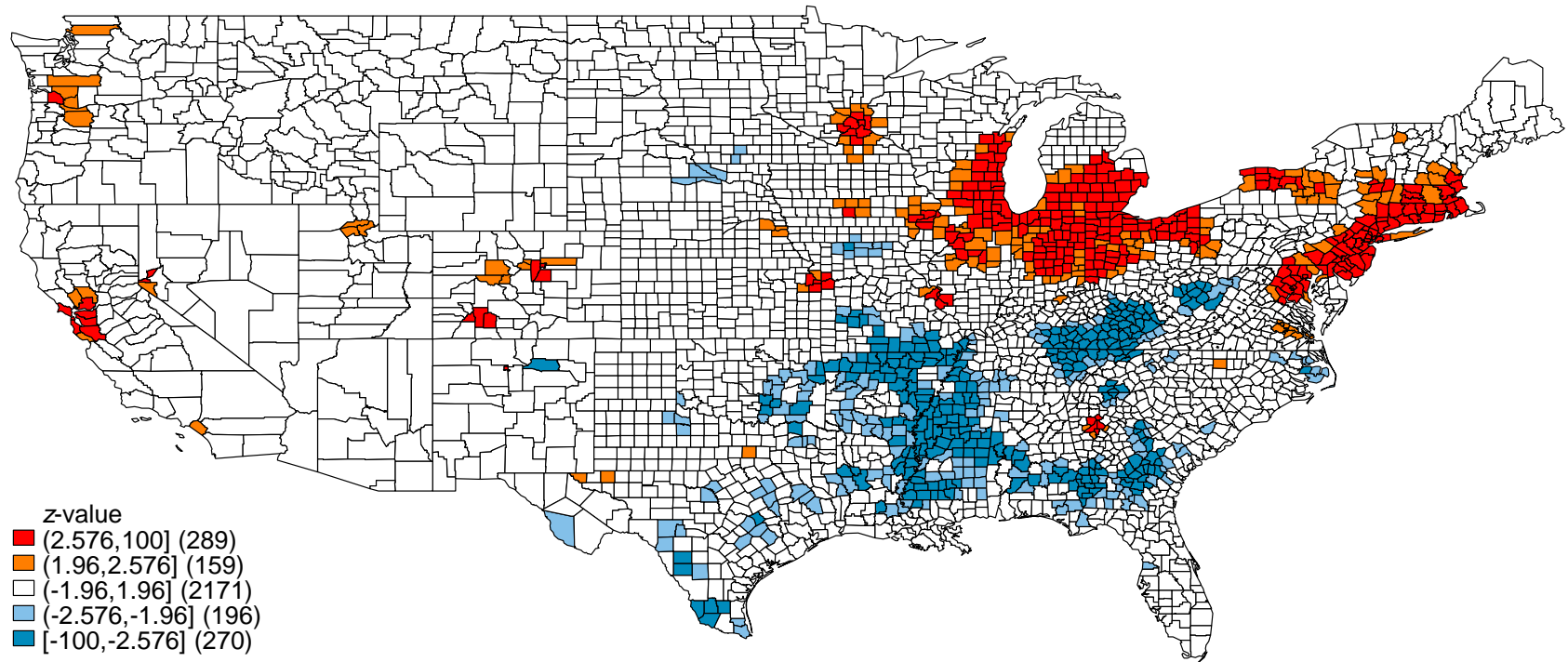
- 区位理论的古老起源：location theory Richard Cantillon (1680-1734)
- 新经济地理理论：Krugman：
 - 修正了经济学第一假设：资源是有限的且是不均分分布的（即利用资源的经济活动的空间分布是非均衡的——Agglomeration; Clustering; Concentration; Unbalance; Block; Lump; Spatial ）；
 - 地理基尼指数 (GGI) : Geographical Gini Index; Spatial Gini Coefficient; Krugman Specialization Index
 - 地理集中度指数 (EGI) : Elilison & Glaeser(1997)提出: Geographical Concentration Index:
 - 产业集中度 (HHI) :Herfindahl-Hirschman Index, 是地理集中度指数的组成部分
- 地理第一定理和第二定理： 社会科学研究与自然科学研究的分野
 - 1. Tobler's First Law of Geography (TFL) : All things are related, but nearby things are more related than distant things. Moran's I; Geary's C; General G. (社会科学研究重点关注的领域：变量在不同地区的相关性)
 - 2. Goodhild's Second Law of Geography (GSL) : Law of Spatial Heterogeneity; Spatial local Heterogeneity; Spatial stratified Heterogeneity: LISA; Getisord (GI); SatScan; Q-statistics (自然科学研究重点关注的领域：变量在不同地区的差异性)

一、数据空间关系的探索是当前计量经济学发展的一个重要方向 —— (2) 可视化探索



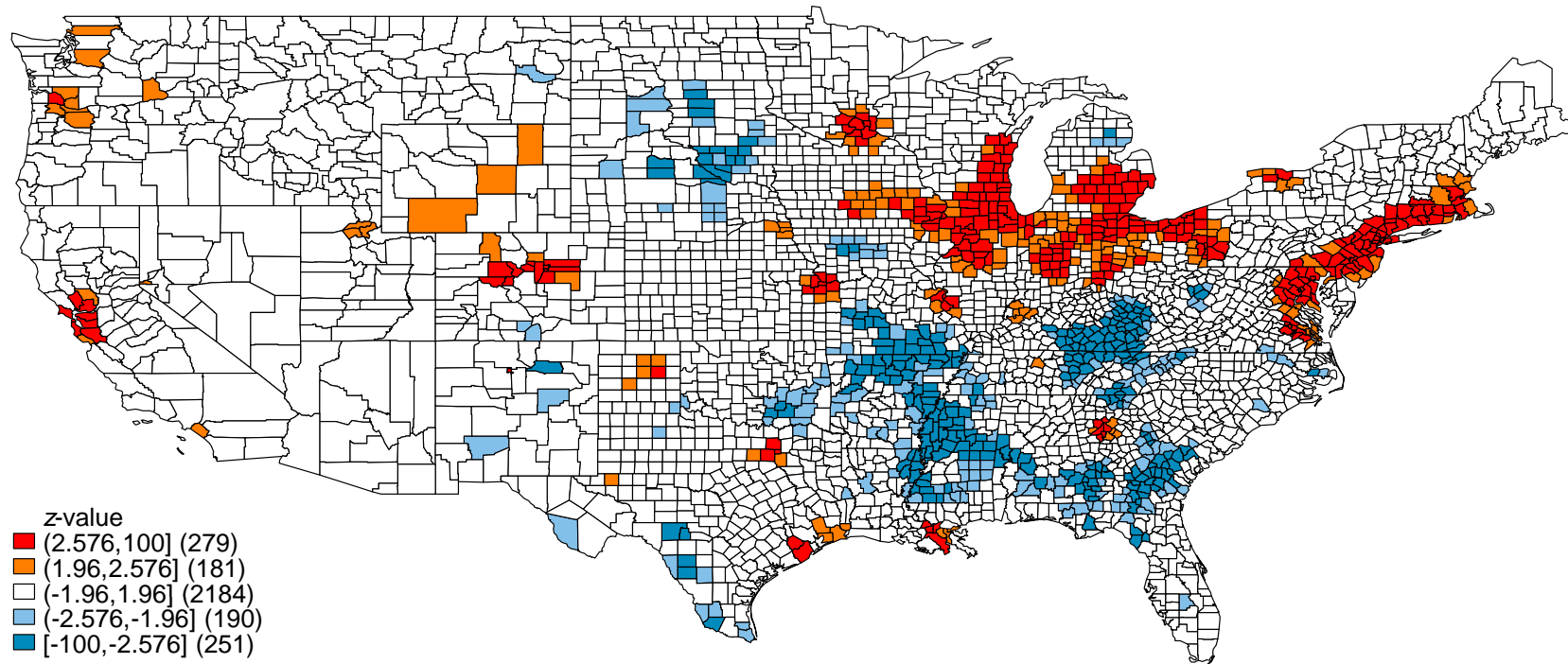
Getisord: Median Family Income in 1959

一、数据空间关系的探索是当前计量经济学发展的一个重要方向 —— (2) 可视化探索



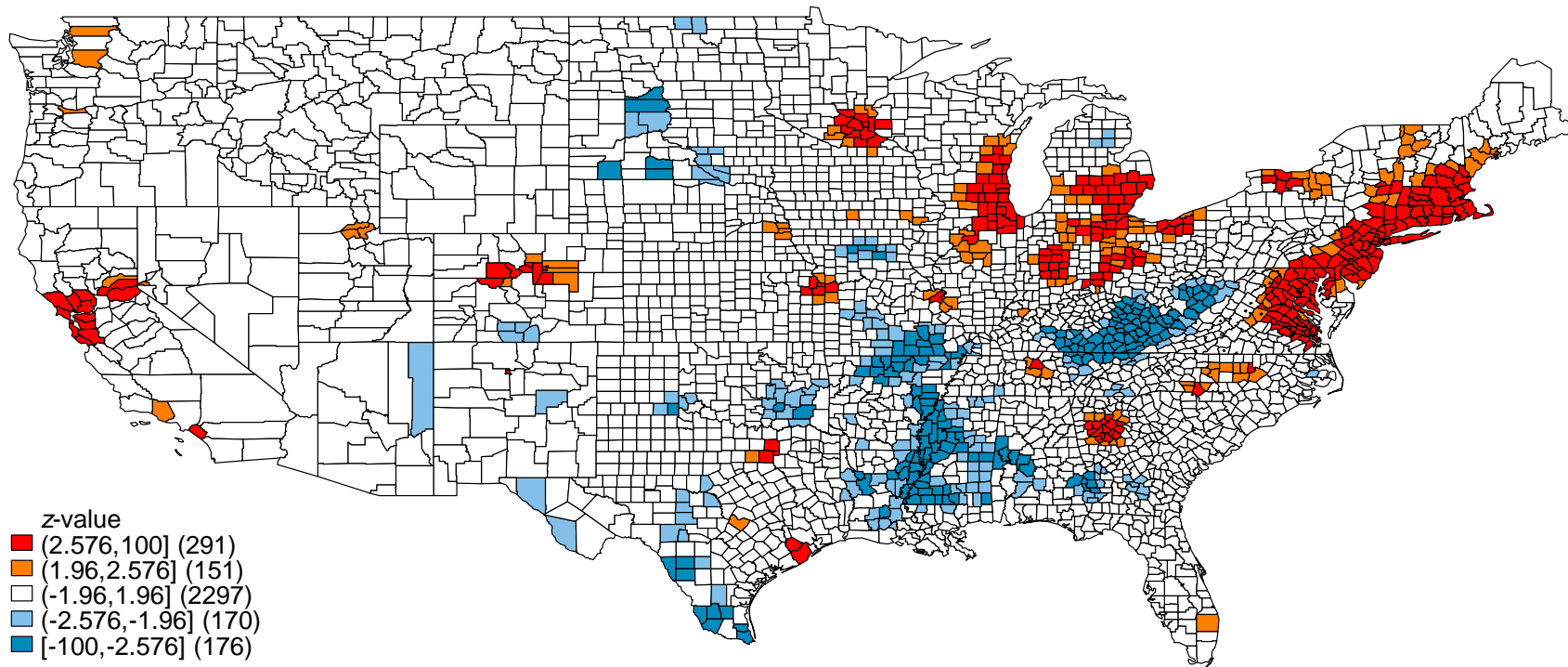
Getisord: Median Family Income in 1969

一、数据空间关系的探索是当前计量经济学发展的一个重要方向 —— (2) 可视化探索



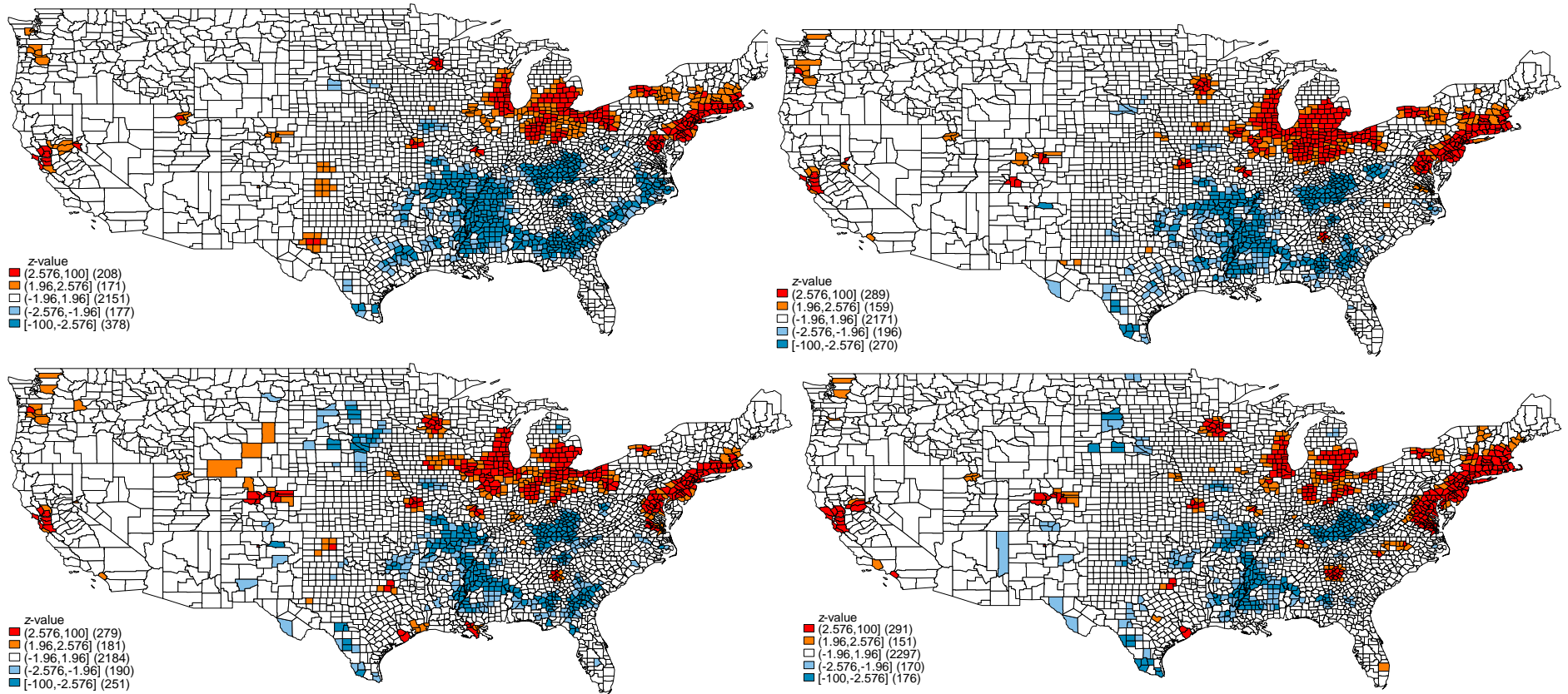
Getisord: Median Family Income in 1979

一、数据空间关系的探索是当前计量经济学发展的一个重要方向 —— (2) 可视化探索



Getisord: Median Family Income in 1989

一、数据空间关系的探索是当前计量经济学发展的一个重要方向 —— (2) 可视化探索



Getisord: Median Family Income in 1959、1969、1979、1989

一、数据空间关系的探索是当前计量经济学发展的一个重要方向

—— (3) 统计学的探索：空间相关性的识别

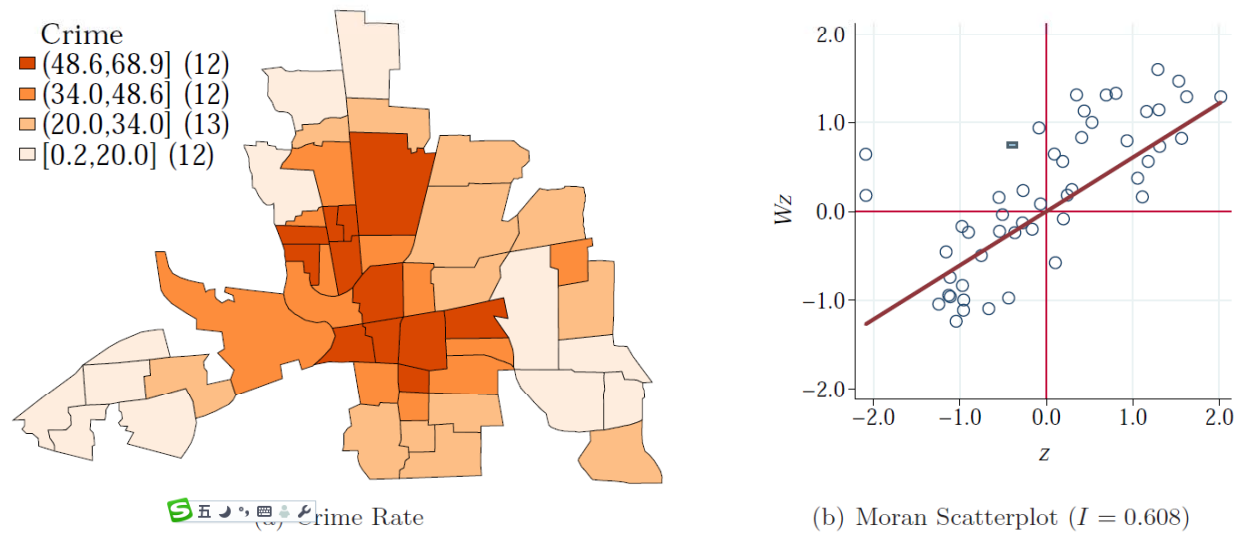


Figure 1: Spatial Variation in Crime Rate

Morans' I: Spatial Autoregression

一、数据空间关系的探索是当前计量经济学发展的一个重要方向 ——（4）分析模型的探索

➤当前（社会科学）对空间地理数据的开发利用的主要方法：

➤空间地理统计学：

- 基础地理数据的处理与估值、地理矩阵、地理多元统计、地理空间统计、地理随机过程、神经网络分析、地理模糊数问题

➤空间计量经济学分析：

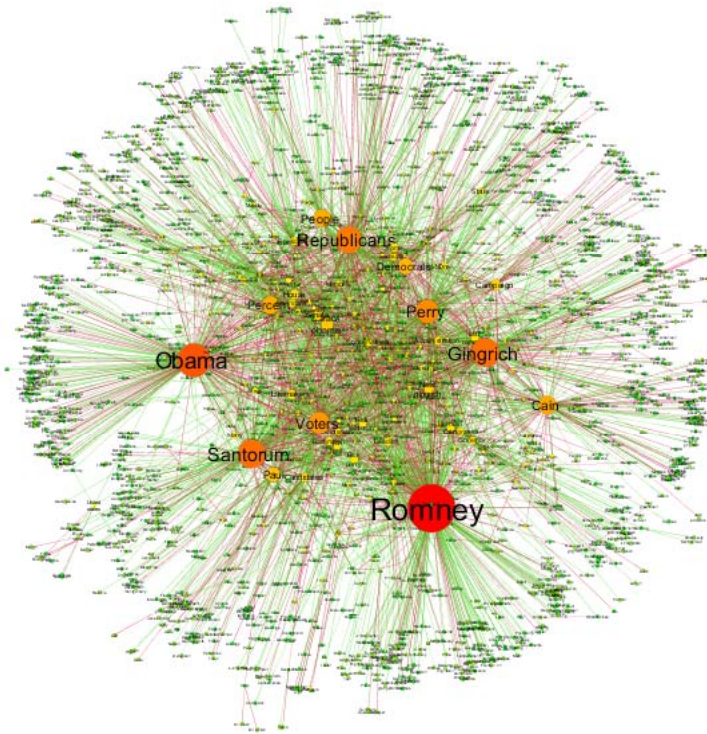
- 空间溢出效应识别、横截面与面板模型、时间空间模型、空间离散数据模型……

➤空间地理或背景变量的测量模型：

- 分层计量分析模型（Multilevel Model Analysis）：适用于抽样调查与分层调查数据
- 结构方程分析模型（SEM: Structural Equation Model）：（包括相关分析、子分析、主成分分析、聚类分析）：重点是 ideas building
- 网络分析法（Networking Analysis）：空间网络节点、线、关系与网络集中度等分析

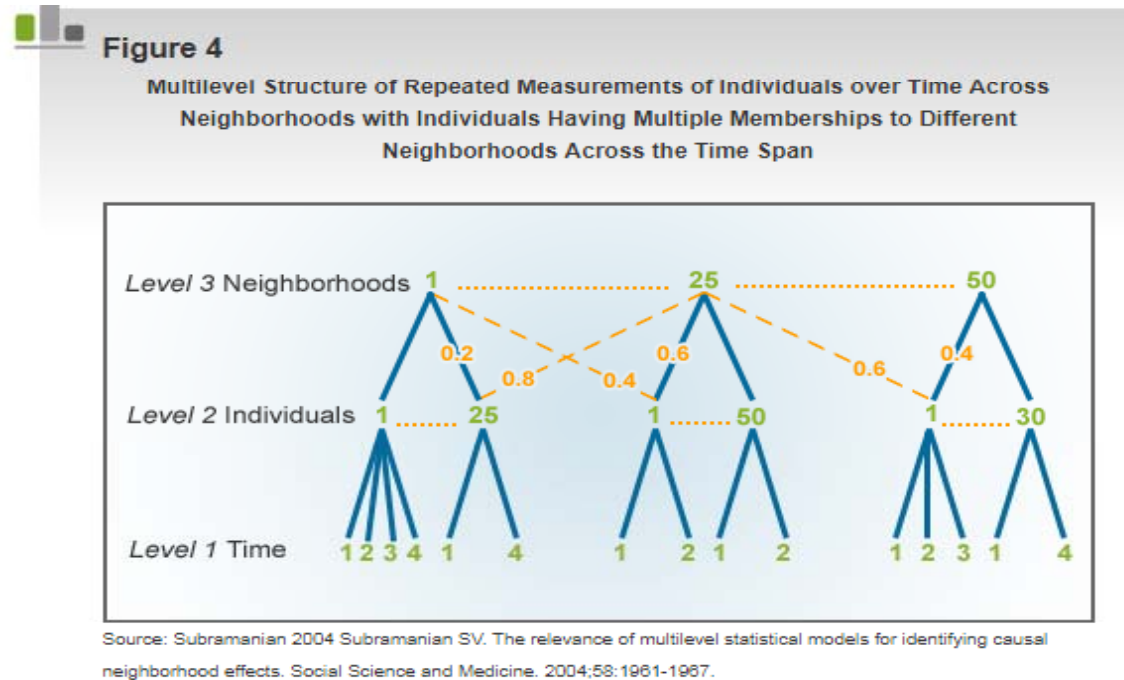
一、数据空间关系的探索是当前计量经济学发展的一个重要方向 ——（4）分析模型的探索

网络数据分析模型 (NWM)



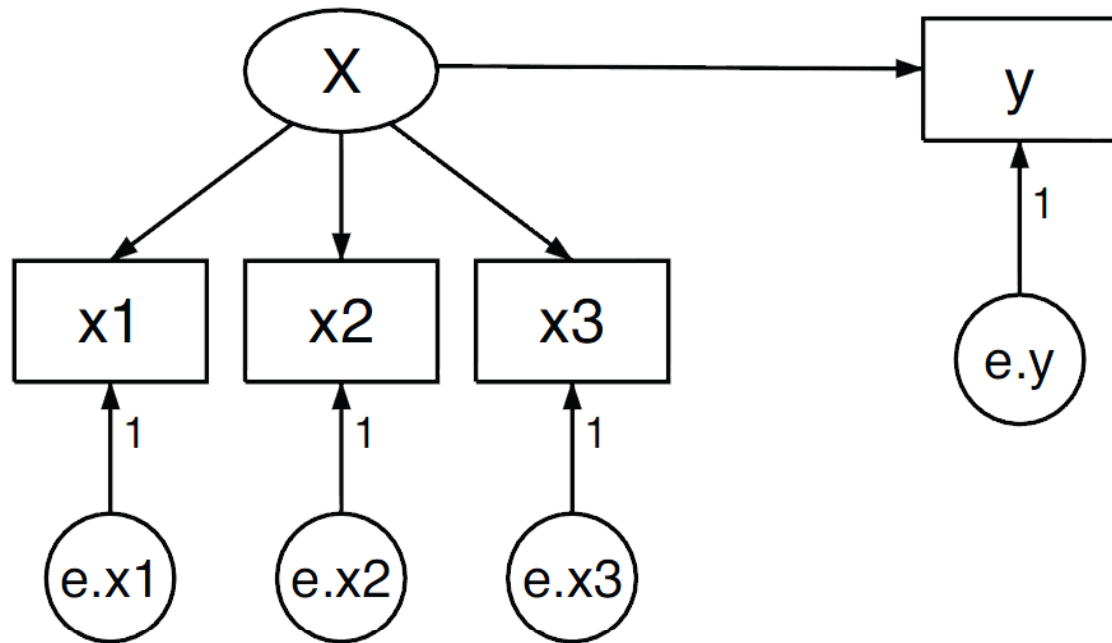
一、数据空间关系的探索是当前计量经济学发展的一个重要方向 ——（4）分析模型的探索

分层数据模型(MLM)



一、数据空间关系的探索是当前计量经济学发展的一个重要方向 ——（4）分析模型的探索

➤ 结构方程模型 (SEM)



一、数据空间关系的探索是当前计量经济学发展的一个重要方向 —— (5) 计量经济学的探索：源于OLS经典假设的放宽

□传统计量经济学模型及其假设

□ $E(y|x_i) = \beta_1 + \beta_i x_i, E(\epsilon) = 0, \text{var}(\epsilon) = \text{var}(y) = \sigma^2; i = 1, 2, \dots, N$

□ $\text{var}(y|x_i) = \sigma^2$

□ x_i is given and more than 2 observations, $y \sim N[\beta_1 + \beta_i x_i, \sigma^2]$ and random

□ $\text{cov}(y_i, y_j) = 0; (i \neq j)$ or $\text{cov}(y_t, y_{t-1}) = 0; (t = 1, 2, 3 \dots n)$: 不完全相关

□ $\text{cov}(\epsilon_i, \epsilon_j) = \text{cov}(y_i, y_j) = 0; e \sim N[0, \sigma^2]; (i \neq j)$: 与观测单位不相关

□ The values of each x_{ik} are not random and are not exact linear functions of the other explanatory variables.

□空间计量经济学是传统计量经济学的拓展（计量经济学变成了空间计量经济学中一个分支）

□ Tobler's First Law of Geography (TFL) : All things are related, but nearby things are more related than distant things.

□ 生在淮南则桔，生在淮北则为枳；近朱则赤，近墨者则黑！

一、数据空间关系的探索是当前计量经济学发展的一个重要方向
—— (5) 计量经济学的探索：源于OLS经典假设的放宽

➤ 空间计量模型的一般形式 (Manski, 1993):

$$Y = \rho WY + \alpha I_N + \beta X + \theta WX + \mu,$$

$$\mu = \gamma W\mu + \epsilon \quad (\text{corrected latent effects})$$

➤ WY : $y_i = f(y_j)(i, j = 1, 2, \dots, n; i \neq j)$: 内生性交互项 (被解释变量之间的相互影响)

➤ WX : $y_i = f(\beta_i x_i + \theta w x_j + \epsilon_i)$: 外生性交互项 (解释变量间的相互影响)

➤ $W\mu$: 误差相关 (外生性交互) 项 (干扰项间的相互影响)

➤ ρ : (被解释变量) 空间自回归系数, spatial autoregressive coefficient

➤ θ : (解释变量) 空间自相关系数: spatial autocorrelation coefficient

➤ γ : (干扰项) 空间自相关系数: spatial autocorrelation coefficient

一、数据空间关系的探索是当前计量经济学发展的一个重要方向
—— (5) 计量经济学的探索：源于OLS经典假设的放宽

□ 空间计量经济学分析应用是计量经济学的新发展：

□ 内生的溢出效应 (spatial externality) : WY

□ 外生的溢出效应(spatial spillover): WX

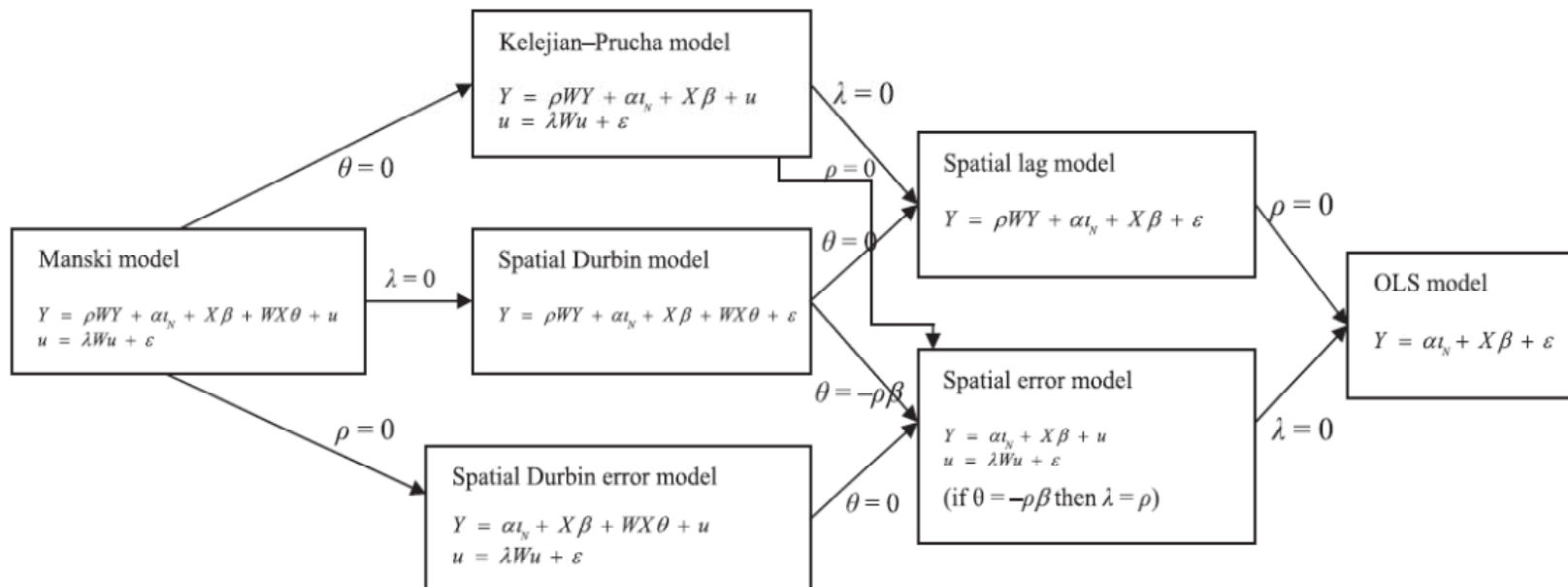
□ 遗漏变量的溢出效应(omission of variables or correlated latent effects):
 $W\mu$

□ 混合效应(spatial mixed effects): $GSAR=SAR+SEM+SDM+SLX$

□ 空间异质性(spatial heterogeneity): GWR

空间计量经济学分析的基本模型及其变形

□ The relationships between different spatial dependence models for cross-section data



二、空间计量经济学中空间权重矩阵的最新发展

□ 度量权重矩阵是空间计量经济学模型估计的重要工具

□ 地理第一定理：Waldo Tobler (University of California, Santa Barbara, UCSB): Everything is related to everything else, but near things are more related to each other.

□ 空间权重矩阵：即“空间相关”或“地理相邻”的数学化描述或度量。

□ 相邻空间权重：

□ 线性相邻 (linear contiguity) : 左或右共一条边

□ 车相邻 (rock contiguity) : 共一条边即可

□ 象相邻 (bishop contiguity) : 共一个点

□ 后相邻 (queen contiguity) : 共一个点或一条边

□ 距离空间权重：即把空间相关的影响限定在一定的地理“带宽”中（距离或距离倒数）：Euclidean Distance (threshold)

二、空间计量经济学中空间权重矩阵的最新发展

□空间权重（度量空间相关的重要工具）构建的新发展：

□距离空间权重矩阵的数学表达：摩擦参数

$$W_{ij} = \begin{cases} 0 & , \text{ if } dij \leq lb \text{ or } dij \geq ub \\ 1/d_{ij}^f & , \text{ if } lb < dij < ub \end{cases}$$

或者下式：距离衰减参数的选择要基于研究对象的性质、理论或社会共识。

$$W_{ij} = f(\theta, d_{ij})$$

二、空间计量经济学中空间权重矩阵的最新发展

□ 空间权重（度量空间相关的重要工具）构建的新发展：

□ 空间权重的最新发展：

- 网络空间权重：（网络分析化）
- 动态化空间权重：用社会化时变的变量构建经济空间权重矩阵
- 指数化空间权重：即出现矩阵指数空间模型（LeSage and Kelley Pace,2009）
- 异质化空间权重： W_1 、 W_2 …

□ 构建空间权重要注意的问题：

- 是一种定性分析：人为设定，即“相信即科学”，是一种价值判断，是一种定（质）性分析
- 要有基本的理论根据或社会共识

二、空间计量经济学中空间权重矩阵的最新发展

□空间计量模型及其估计的新发展:

□横截面空间计量模型

□(1) 线性模型: 连续被解释变量和离散被解释变量

□(2) 非线性模型:

□空间面板计量模型:

□空间-时间计量模型

□最新发展空间计量模型

□空间VAR模型

□空间联立方程组模型

□空间因果关系 (Spatial Granger causal analysis) 分析

□建模思路:

□ (1) 从特殊到一般; (2) 从一般到特殊

□空间计量模型的估计:

□常用的空间权重: (1) 二值相邻空间权重 (q个最近近邻) (2) (逆) 距离空间权重;

□空间权重的选择方法: (1) 根据理论人为设定; (2) 根据拟合优度; (3) 根据贝叶斯后验概率确定

□估计方法: (1)最大似然估计(ML);(2) 矩估计(IV/GMM:2SLS); (3)贝叶斯估计(MCMC)

三、对空间计量经济学发展的评价

□我对空间计量经济学的评价（见《空间计量经济学导论》译者序）：

- 空间计量经济学是**计量经济学空间化或是地理统计学计量化**的结果，它融合了地理统计、计量经济学、空间经济学、计算机编程和计量软件开发等多学科最新发展成果。
- 空间计量经济学的出现，彻底改变了传统经济学忽视经济资源和经济活动对地理区位或空间的依赖的现状，而且**全面颠覆了传统计量经济学对观测个体相互独立的经典假设**。
- 空间计量经济学快速发展的重要意义，不仅**重塑了传统计量经济学的分析框架和基本思路，而且把观测个体（或单位）在空间和时间上溢出效应作为统计度量和计量识别分析的重要领域，重点对观测个体在一些地理区位上的空间依赖或空间溢出关系进行量化研究**，从而推动了计量经济学的全新发展。

四、空间计量经济学的主要参考书

➤请参阅作者已经出版的译著或编著

- 肖光恩等译，《空间计量经济学导论》，北京大学出版社，2014年1月出版。
- 肖光恩独译，《空间计量经济学：从横截面数据到空间面板》，中国人民大学出版社，2015年4月出版。
- 肖光恩等编著，《空间计量经济学：基于Matlab的应用分析》，北京大学出版社，2018年8月出版。
- 肖光恩等译，《空间计量经济学入门：在R中的应用》，中国人民大学出版社，2018年5月出版。
- 肖光恩等译，《用Stata学微观计量经济学》，重庆大学出版社，2015年4月出版。
- 闫吉丽、肖光恩译，《空间计量经济学：空间数据分位数的回归》，中国人民大学出版社，2018年3月出版。
- 张璐、肖光恩等译，《空间数据分析：模型、方法与技术》，中国人民大学出版社，2018年8月。