# 新冠疫情数据分析系列公益讲座 COVID-19 Data Analysis Webinars

## 第九讲、基于面板数据的地区疫情经济影响分析

主讲人：吴玉鸣，华东理工大学

主持：沈体雁，北京大学

合办单位：华东理工大学商学院

讲座时间：北京时间2020年6月26日上午9-10点

9:00-10:00 PM, June 25, 2020, EDT (New York Time)

# 议题 Topics

- 研究目标与背景Objectives & Background
- 以往研究及问题Previous Study & Problems
- 数据来源与变量Data Sources & Variables
- 方法与云计算平台Methodology & Platform
- 模型分析与结果Model Analysis & Results
- 总结与讨论Summary and Discussions
- 参考文献和致谢References & Acknowledgements

# 研究目标Objectives

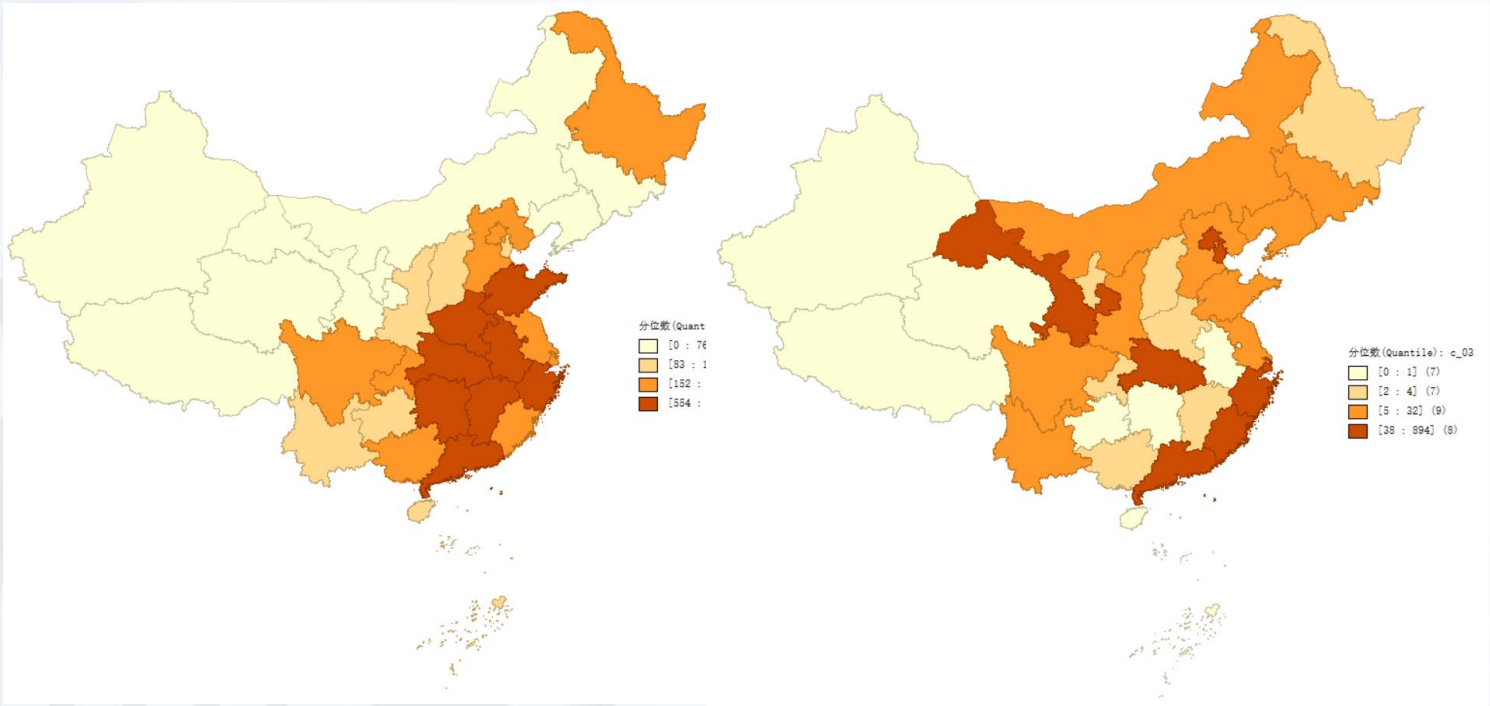
- 新型冠状病毒肺炎是近百年来人类遭遇的影响范围最广的全球性大流行病，是新中国成立以来发生的传播速度最快、感染范围最广、防控难度最大的一次重大突发公共卫生事件（钟瑛与陈盼，2020），微观层面的个体行为通过各行业的生产要素供给、商品价格、投资的变动，反映到宏观层面进而辐射影响到产业经济上。（何诚颖等，2020 ）
- 以2020年2-4月期间中国工业经济为例，采用探索性空间数据分析方法，分析中国省域新冠肺炎病毒肺炎疫情（以下简称为新冠疫情）的空间分布特征；基于空间杜宾面板计量经济模型，分析新冠疫情对中国区域工业经济发展的影响。
- 探究省域新冠疫情的空间分布模式，检验其对工业经济增长的本地-邻地效应、直接-间接效应，初步探究新冠肺炎疫情对中国省域工业经济增长的影响程度及具体作用途径（定资产投资、居民消费、商品零售），为通过投资、消费、贸易等路径抗击新冠疫情对工业经济增长的影响提供政策参考。

# 研究背景Background

- 随着经济社会全球化和区域经济一体化程度的不断提高，以及劳动、资本、技术和信息等要素在地区之间流动性的日益增强，突发重大疫情对经济社会发展的影响及传到途径错综复杂，研究重大传染病疫情对区域经济发展的影响具有重要的意义。
- 自从2019年12月新冠肺炎肺炎疫情在武汉爆发以来，已经扩散并波及到中国大陆31个省级区域，生产下降、消费减少、投资下降、贸易疲软，产业经济增长遭受严重损失。
- 新冠疫情爆发及传播具有深度不确定性，其对中国各个地区尤其是省级区域的影响深远，我国各个地区正面临着新冠疫情的严峻考验。

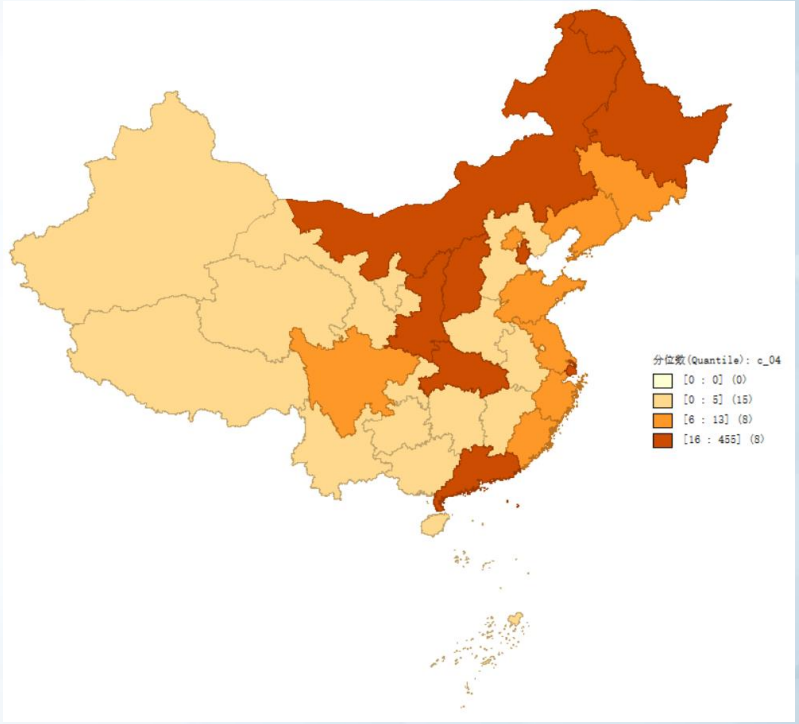
# 研究背景Background

2020年2-4月31个省域新冠疫情空间分布四分位图



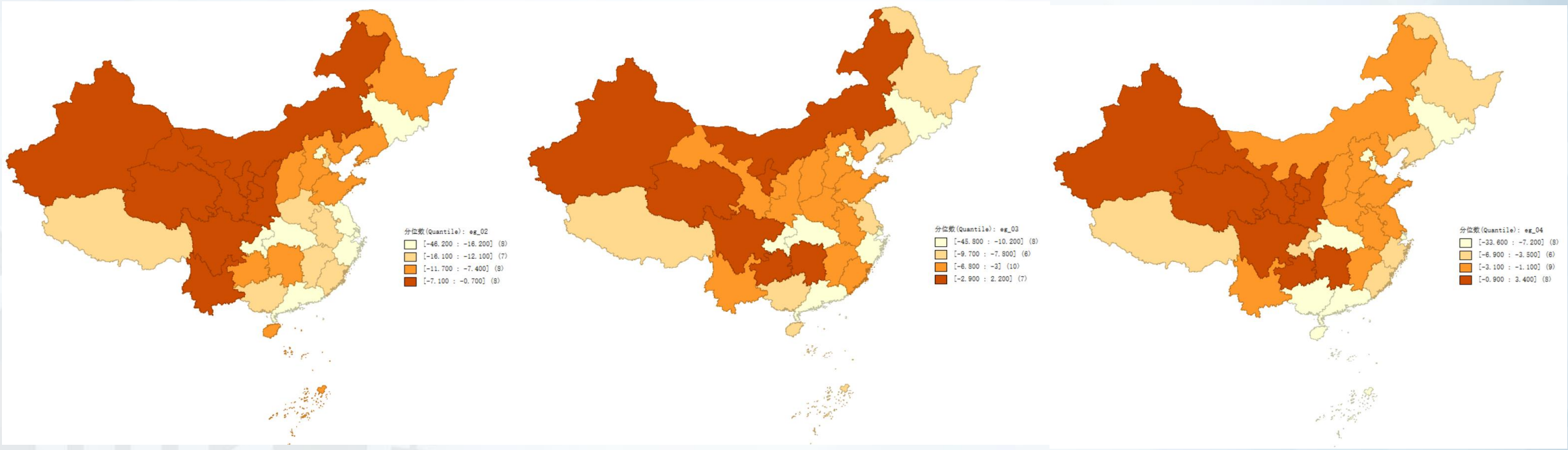






# 研究背景Background

2020年2-4月31个省域工业经济增长空间分布四分位图



# 以往研究及问题Previous Study & Problems

- 武汉新冠肺炎疫情最大受害对象是老年群体，新冠肺炎老年群体重症率要远高于其他年龄群体（Huang et al.，2020）。
- 疫情对生产性服务业的负面影响要大于生活性服务业，疫情对中国制造业的影响可能要大于中国服务业（刘志彪，2020）。
- 刘世锦等（2020）将网络分析法应用于投入产出体系，深入分析了新冠肺炎疫情对经济的冲击路径的研究发现，湖北省经济具有“内向型”特征，新冠肺炎疫情对其省内经济的冲击大于省外；与湖北省经济依存度较高的江苏、广东、浙江等省级行政单位，面临的直接经济冲击风险较大。

# 以往研究及问题Previous Study & Problems

- 尹彦辉等（2020）基于新凯恩斯DSGE模型，分析了新冠肺炎疫情对中国宏观经济的影响发现，新冠肺炎疫情对中国宏观经济的影响具有阶段性特征，以短期冲击为主，长期冲击不显著；一定程度上，政府增加**投资性支出**和转移性支出都可以改善消费需求萎靡的状况。
- 朱军等（2020）基于动态随机一般均衡模型分析新冠肺炎疫情影响社会经济的机制路径和影响程度，并对财政政策干预的效果进行了量化评估的研究发现，新冠肺炎疫情主要通过缩小有效劳动力规模来破坏**市场正常交易**机制，恢复公众信心是有效控制新冠肺炎疫情的重要手段。
- 田盛丹（2020）基于可计算的一般均衡模型，对新冠肺炎疫情对经济增长可能产生的冲击进行了模拟研究发现，新冠肺炎疫情对各经济主体、行业部门的冲击严重，特别是对于**居民消费**，财政政策可以有效缓解冲击。

# 以往研究及问题Previous Study & Problems

- 综上文献，现有研究在新冠疫情对经济社会的负面影响负面取得了一些进展。
- 由于数据可得性问题及一些技术性问题，目前的研究大多数停留在定性分析或间接的定量分析阶段，主要通过时间序列数据分析或CGE/DSGE模型等进行模拟分析，并未直接采用新冠肺炎确诊病例、死亡病例及治愈出院病例数据进行空间计量模型研究。
- **新冠疫情具有明显的空间扩散特征，而现有文献对新冠肺炎疫情的研究缺乏对空间因素的足够关注，对疫情影响经济社会的机制路径和影响程度关注不够。**

# 数据来源与变量Data & Variables

- 2020年2-4月中国大陆31个省域新冠肺炎疫情面板数据。
- 新冠肺炎病例数据源于哈佛大学中国数据实验室、Wind数据库，宏观经济数据源于国家统计局网站。

**表1 变量及其相关的信息**

| 变量类别 | 变量名称 | 对应的英文简称 | 备注 |
|---|---|---|---|
| 被解释变量 | 工业经济增长 | EG | 工业增加值累计增加（%） |
| 核心解释变量 | 新冠肺炎确诊病例 | CONFIRMED | 增量（例） |
| | 新冠肺炎死亡病例 | DEATHS | 增量（例） |
| | 新冠肺炎治愈出院病例 | RECOVERED | 增量（例） |
| 控制变量 | 居民消费价格指数 | CPI | % |
| | 商品零售价格指数 | RPI | % |

# 数据来源与变量Data & Variables

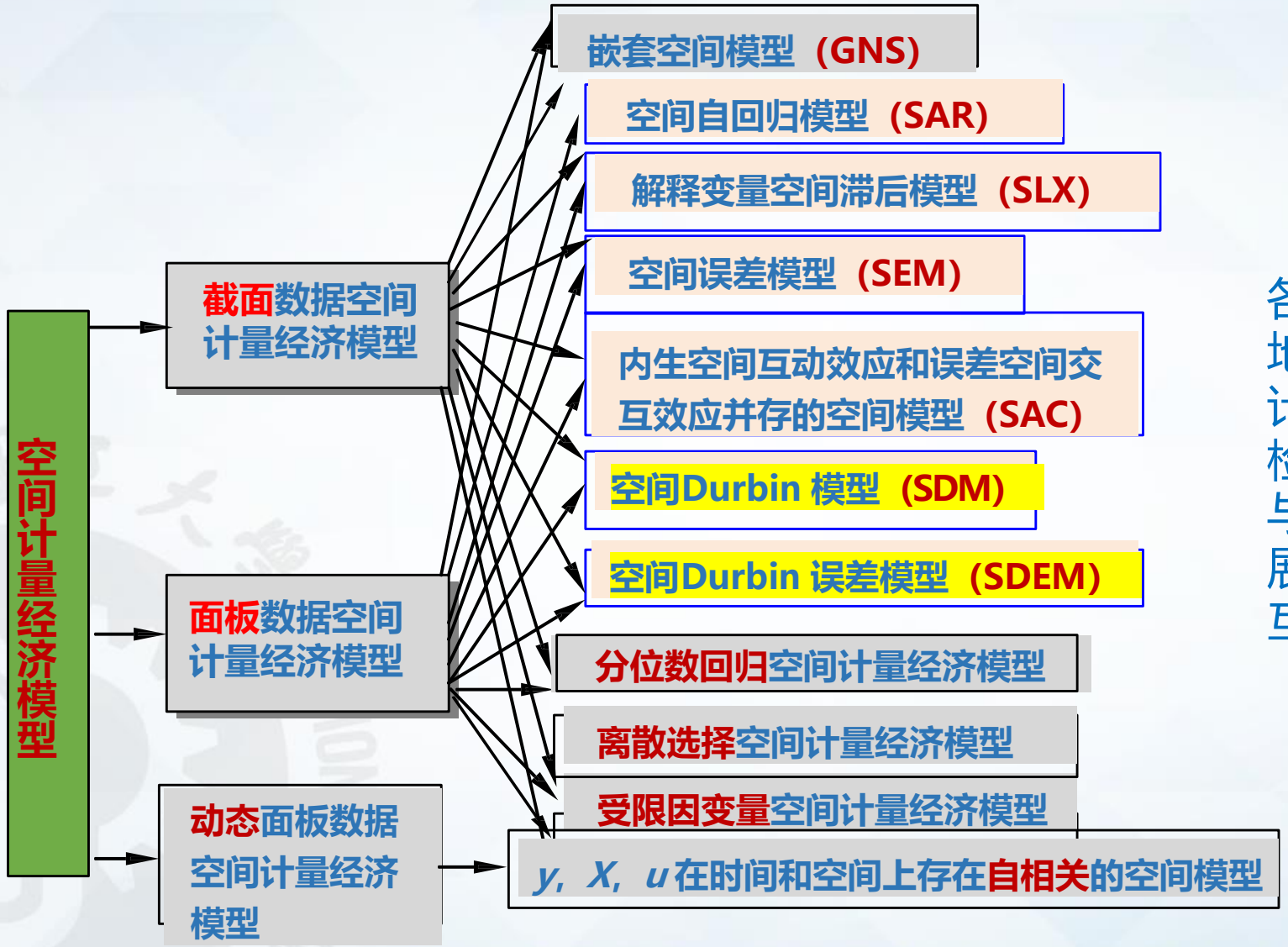
表2 描述性统计

| 变量名称 | 指标 | 单位 | 观测值 | 均值 | 标准差 | 最小值 | 最大值 |
|---|---|---|---|---|---|---|---|
| 工业经济增长 | EG | % | 93 | -8.586 | 8.629 | -46.200 | 3.400 |
| 新冠肺炎确诊病例 | CONFIRMED | 人 | 93 | 764.215 | 6187.040 | 0.000 | 59754.000 |
| 新冠肺炎死亡病例 | DEATHS | 人 | 93 | 47.032 | 295.664 | 0.000 | 2512.000 |
| 新冠肺炎治愈出院病例 | RECOVERED | 人 | 93 | 832.032 | 4577.236 | 0.000 | 31966.000 |
| 固定资产投资 | FAI | % | 93 | -12.628 | 16.284 | -82.800 | 19.100 |
| 居民消费价格指数 | CPI | % | 93 | 104.022 | 1.177 | 101.700 | 106.900 |
| 商品零售价格指数 | RPI | % | 93 | 102.793 | 1.444 | 100.000 | 106.800 |

# 方法与云计算平台Methodology & Platform

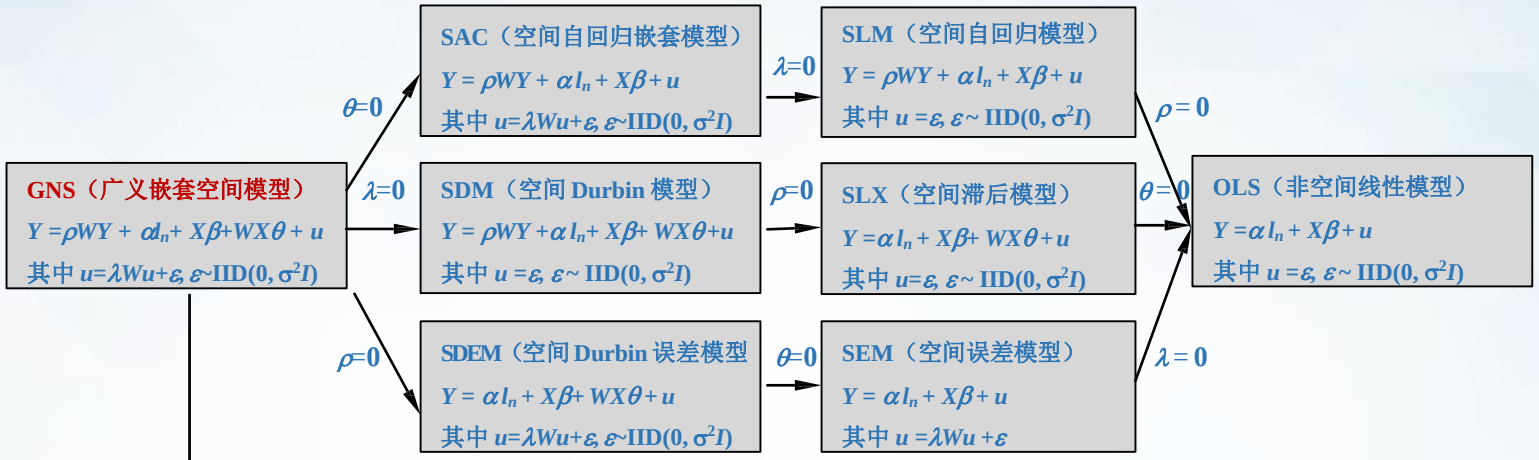
- 在经济社会全球化和区域经济一体化的趋势下，技术经济、投资贸易和人文交流等活动越来越密切，**区域经济之间的空间互动性（依赖性和差异性）越来越强**，经济发达和欠发达发达国家和地区对诸如新冠疫情等重大公共卫生事情的差异性应对举措（同舟共济还是以邻为壑？），将产生迥异和难以预料的结果。
- 空间计量经济学方法：测度空间互动的直接效应和间接效应？
  - ✓探索性空间数据分析（ESDA）
  - ✓空间面板计量数据经济分析（Spatial Panel Data Econometrics Analysis）

# 方法与云计算平台Methodology & Platform




各种空间经济学或新经济地理学理论需要运用空间计量经济学理论与方法去检验，而这些大量的理论与实证相结合的研究的开展极大地推动了包括空间互动行为研究的快速发展。

## 截面数据、面板数据空间计量经济模型




## 空间和时间上的广义动态空间依赖模型

$$Y_t=\tau Y_{t-1}+\rho W Y_t+\eta W Y_{t-1}+X_t\beta_1+W X_t\beta_2+X_{t-1}\beta_3+W X_{t-1}\beta_4+Z\pi+u$$

其中 $u_t=\gamma u_{t-1}+\lambda W u_{t-1}+\mu+\xi I+\varepsilon, \varepsilon\sim IID(0,\sigma^2 I)$

$\mu=\kappa W\mu+\xi$

$Y$ 是 $N\times1$ 阶被解释变量列向量，$I$ 是 $N\times1$ 阶单位列向量，$\alpha$是常数项。$X$ 是 $N\times K$ 阶外生解释变量矩阵，$\beta$是 $K\times1$ 阶待估计回归 系数列向量。$\varepsilon$是 $N\times1$ 阶服从独立同分布的随机误差项列向量。

空间互动效应有三类：（1）$Y_i$与 $Y_j$ 间的内生变量互动效应。（2）$X_i$ 与 $X_j$ 间的外生变量互动效应。（3）$u_i$ 与 $u_j$ 间的随机误差项互动效应。

GNS（广义嵌套空间模型）是具有所有类型互动效应的模型。其中 $WY$ 表示被解释变量之间的内生变量互动效应。$WX$ 表示解 释变量之间的外生变量互动效应。$Wu$ 表示随机误差项之间的互动效应。

$\rho$表示空间自回归系数。$\lambda$表示空间自相关系数。$\theta$表示 $K\times1$ 阶待估计回归系数列向量。$W$ 表示 $N\times N$ 阶非负矩阵，其主对角线上 元素为零（自己不能与自己为邻），只设定不同个体的空间组合和安排。

上述模型可以用极大似然法（ML），也可以用广义距方法（GMM）估计。

$\tau$、$\eta$、$\beta_1$、$\beta_2$、$\beta_3$、$\beta_4$是空间回归系数。$Z$ 是 $N\times L$ 阶内生解释变量矩阵。$\mu$表示 $N\times1$ 阶特定空间效应。

$$\frac{\partial E(\boldsymbol{y})}{\partial x_{1k}} \cdots \frac{\partial E(\boldsymbol{y})}{\partial x_{Nk}} = (\boldsymbol{I} - \rho \boldsymbol{W})^{-1} \begin{bmatrix} \beta_k & w_{12}\theta_k & \cdots & w_{1N}\theta_k \\ w_{21}\theta_k & \beta_k & \cdots & w_{2N}\theta_k \\ \cdots & \cdots & \ddots & \cdots \\ w_{N1}\theta_k & w_{N2}\theta_k & \cdots & \beta_k \end{bmatrix}, k = 1, 2, \ldots, K$$

（1）特定解释变量的特定个体不仅会改变这个个体对应的被解释变量的值，还会改变被解释变量其他个体的值。改变该个体对应的被解释变量的值称作直接效应，改变被解释变量其他个体的影响称作间接效应。

偏导数矩阵的每一个主对角线上的元素代表的是直接效应。每一个非主对角线上的元素代表的是间接效应。因此，如果 $\rho = \theta_k = 0$，则间接效应不会发生。所有非主对角线上的元素都等于零。

（2）直接效应和间接效应对于个体来说是不同的。直接效应的不同是因为当 $\rho \neq 0$ 时，$(I_n - \rho W)^{-1}$ 的主对角线上的元素对于不同的个体是不同的。间接效应的不同是因为当 $\rho \neq 0$ 和（或）$\theta_k \neq 0$ 时，$(I_n - \rho W)^{-1}$ 和 $W$ 的非主对角线上的元素对于不同的个体是不同的。

（3）若 $\theta_k \neq 0$ 时，则发生的间接效应被称为局域效应，而 $\rho \neq 0$ 所发生的效应被称作全域效应。间接效应之所以被称为局域效应，是因为其仅仅影响该个体的附近的个体集合；当 $\rho \neq 0$ 时，对 $y$ 中的所有个体都会产生影响。当同时 $\rho \neq 0$，$\theta_k \neq 0$ 时，全域效应和局域效应都会存在，且无法区分。

（4）如果有 $N$ 个空间个体，$K$ 个解释变量，将得到 $K$ 个 $N \times N$ 维的直接和间接效应矩阵。即使 $N$ 和 $K$ 很小，简洁地报告这些结果也是困难的。于是建议使用概述指标。直接效应的概述指标是用右边矩阵主对角线上的元素的平均值度量第 $k$ 个解释变量全部个体对 $y$ 的平均影响。间接效应的概述指标是用右边矩阵剔除主对角线上的元素的行或列的平均值度量第 $k$ 个解释变量的个体对 $y$ 的影响。

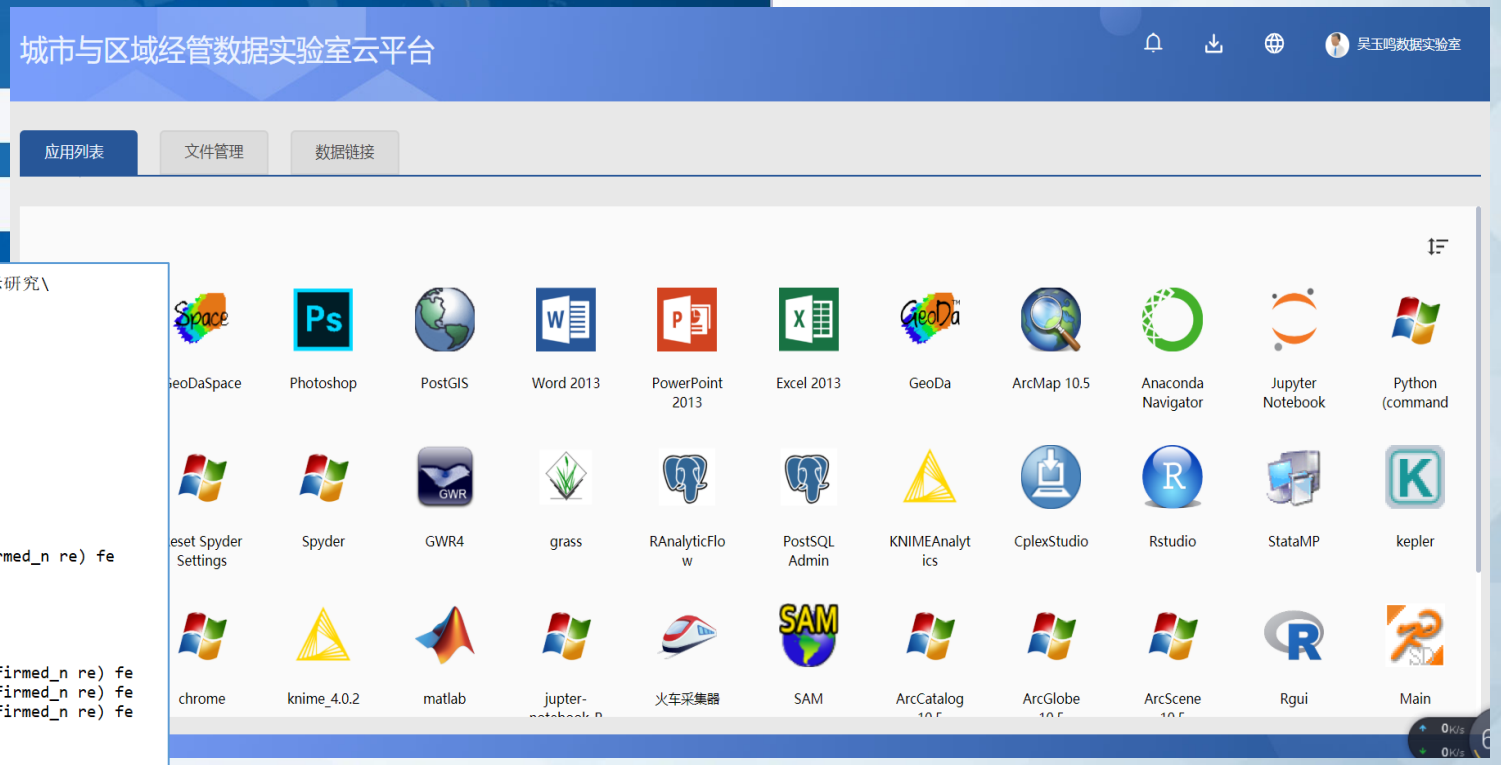
平均行效应表示第 $k$ 个解释变量的所有个体都发生一个单位的变化对被解释变量 $y$ 的某个特定个体的影响。

平均列效应表示第 $k$ 个解释变量的特定个体发生一个单位的变化对被解释变量 $y$ 所有个体的影响。

实际应用中间接效应的概述指标常使用列平均效应值。

# 方法与云计算平台Methodology & Platform

- GeoDA
- ArcGIS
- Stata




do file

```
cd C:\Users\user\Desktop\新冠肺炎疫情对中国经济增长的影响和重大突发公共卫生事件响应机制的启示研究\
新冠肺炎疫情对中国经济增长的影响和重大突发公共卫生事件响应机制的启示研究
use China_COVID-19_Data.dta,clear
xtset id month
spmat use winv using winv.spmat
spmat use queen using queen.spmat
spmat use idisswm using idisswm.spmat

. xtline confirmed_n
***(各省域新增新冠肺炎确诊病例的时间趋势图)

***变量的描述性统计
summarize eg confirmed_n deaths_n recovered_n fai re cpi rpi

xsmle eg confirmed_n deaths_n recovered_n fai cpi rpi, wmat(idisswm) model(sdm) durbin(confirmed_n re) fe

***交互项
gen intp1=confirmed_n*cpi
gen intp2=confirmed_n*rpi
gen intp3=confirmed_n*fai
xsmle eg confirmed_n deaths_n recovered_n cpi rpi intp1, wmat(idisswm) model(sdm) durbin(confirmed_n re) fe
xsmle eg confirmed_n deaths_n recovered_n cpi rpi intp2, wmat(idisswm) model(sdm) durbin(confirmed_n re) fe
xsmle eg confirmed_n deaths_n recovered_n cpi rpi intp3, wmat(idisswm) model(sdm) durbin(confirmed_n re) fe

xsmle eg confirmed_n cpi rpi intp1, wmat(idisswm) model(sdm) durbin(confirmed_n re) fe
xsmle eg confirmed_n cpi rpi intp2, wmat(idisswm) model(sdm) durbin(confirmed_n re) fe
xsmle eg confirmed_n cpi rpi intp3, wmat(idisswm) model(sdm) durbin(confirmed_n re) fe

***计算残差
predict e
```

培养知行合一的经管人才

AGERIAL TALENTS WITH ADVANCED KNOWLEDGE AND PRACTICAL ABILITY.

# 模型分析与结果Model Analysis & Results

## 动态空间杜宾面板计量经济模型

$$EG_{it} = \alpha + \tau EG_{it-1} + \rho \sum_{j=1}^{N} w_{ij} EG_{jt} + \beta_1 CONFIRMED_{it} \times FAI_{it} + \beta_2 CONFIRMED_{it} \times CPI_{it} + \beta_3 CONFIRMED_{it} \times RPI_{it}$$
$$+ \beta_4 X_{it} + \delta_1 \sum_{j=1}^{N} w_{ij} CONFIRMED_{jt} \times FAI_{jt} + \delta_2 \sum_{j=1}^{N} w_{ij} CONFIRMED_{jt} \times CPI_{jt}$$
$$+ \delta_3 \sum_{j=1}^{N} w_{ij} CONFIRMED_{jt} \times RPI_{jt} + \delta_4 \sum_{j=1}^{N} w_{ij} X_{jt} + \mu_i + \gamma_t + \varepsilon_{it} \quad (1)$$

由于部分省域的新冠肺炎确诊病例为0，为了减轻异方差性，进行了加1取对数的处理。

# 模型分析与结果Model Analysis & Results

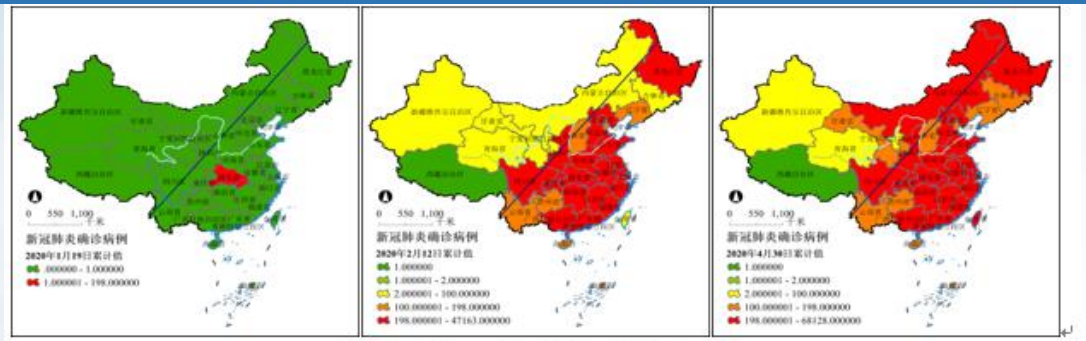


图1　累计新冠肺炎确诊病例空间分布图

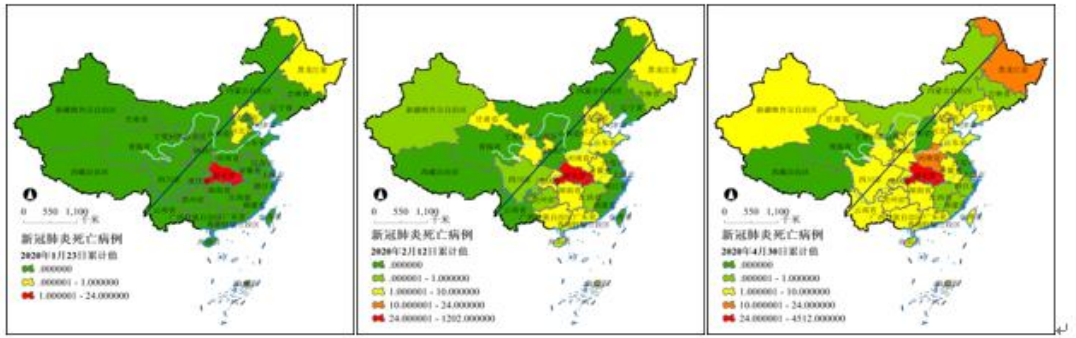


图2　累计新冠肺炎死亡病例空间分布图

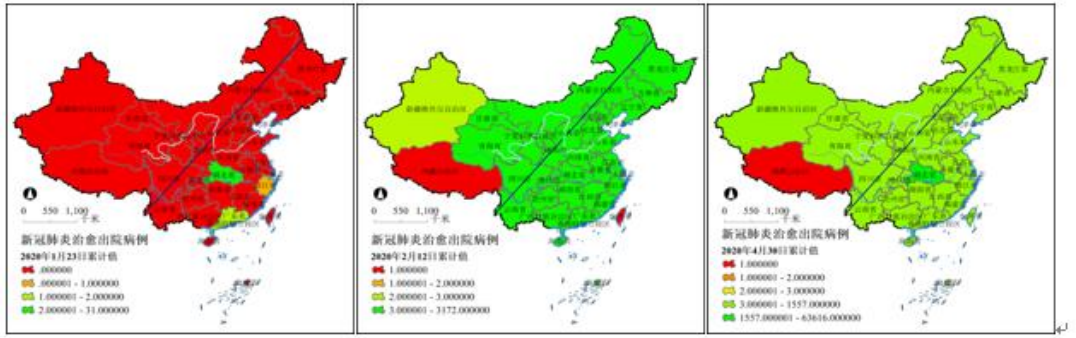


图3　累计新冠肺炎治愈出院病例空间分布图

图1-3显示：

- 累计新冠肺炎确诊病例在空间分布上呈现出逐渐增加的趋势，但其在空间分布上呈现出的逐渐增加趋势得到了明显遏制；
- 累计新冠肺炎死亡病例在空间分布上呈现出逐渐增加的趋势，同样其在空间分布上呈现出的逐渐增加趋势得到了明显遏制；
- 累计新冠肺炎治愈出院病例在空间分布上呈现出逐渐增加的趋势。

# 模型分析与结果Model Analysis & Results

图4-6显示：

- 新增新冠肺炎确诊病例在空间分布上呈现出先增加后减少的特征，表明在空间上增加趋势得到了明显遏制。
- 新增新冠肺炎死亡病例在空间分布上呈现出先增加后减少的特征，表明其空间上增加的趋势得到了明显遏制；
- 新增新冠肺炎治愈出院病例在空间分布上呈现出先增加后减少的特征，表明新冠肺炎病例存量在逐渐减少，防控措施初见成效。
- 总体上看，新冠肺炎疫情呈现出零星散发状态；在局域上，呈现出散发病例引起的聚集性疫情。

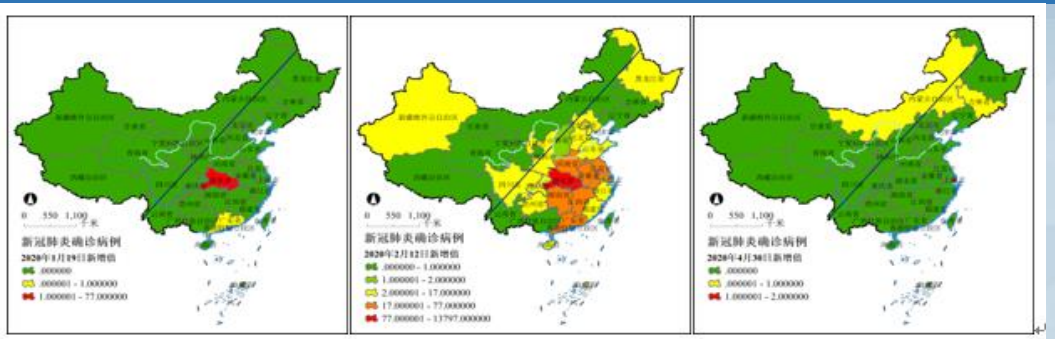


图 4　新增新冠肺炎确诊病例空间分布图

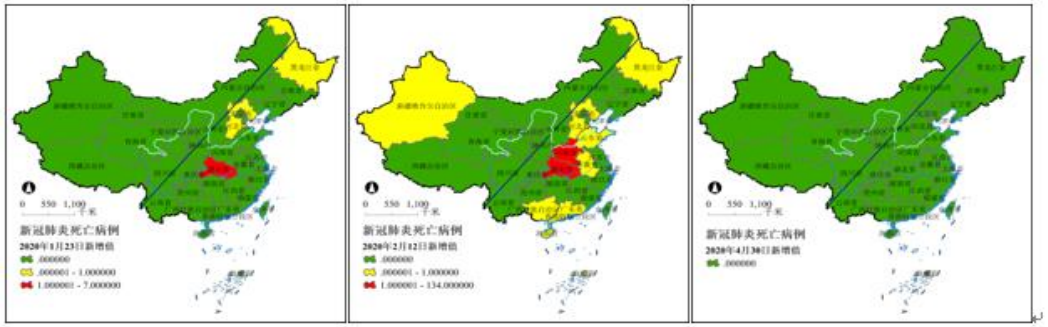


图 5　新增新冠肺炎死亡病例空间分布图

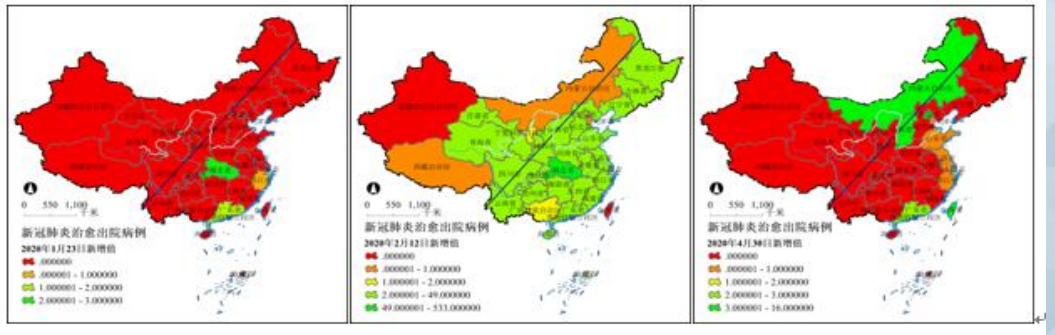


图 6　新增新冠肺炎治愈出院病例空间分布图

# 模型分析与结果Model Analysis & Results

表3 基于rook空间权重定义的被解释变量与核心解释变量的Moran's I

| 变量 | 2月的Moran's I | 3月的Moran's I | 4月的Moran's I |
|---|---|---|---|
| EG | 0.030***<br>（2.970） | 0.012***<br>（2.147） | 0.001*<br>（1.381） |
| CONFIRMED | -0.0247***<br>（-0.0345） | -0.0776***<br>（-0.0345） | -0.0123<br>（-0.0345） |
| CONFIRMED<br>（剔除武汉） | 0.0488***<br>（-0.0357） | 0.0488**<br>（0.221） | 0.0595**<br>（0.157） |

# 模型分析与结果Model Analysis & Results

表4 动态SDM模型回归结果

| 变量 | 系数 | 直接效应 | 间接效应 | 总效应 |
|---|---|---|---|---|
| $EG_{t-1}$（τ） | 0.175***（4.61） | — | — | — |
| W·EG (ρ) | 0.551**（2.00） | — | — | — |
| FAI | 0.018（0.44） | 0.032***（0.70） | 0.392**（1.98） | 0.424**（1.99） |
| CPI | 0.360（0.54） | 0.964（1.11） | 9.00*（1.73） | 9.964*（1.70） |
| RPI | -1.214**（-2.16） | -1.605**（-2.38） | -5.539（-1.39） | -7.144（-1.59） |
| W·FAI | 0.414***（3.03） | — | — | — |
| W·CPI | 8.990***（2.71） | — | — | — |
| W·RPI | -4.731*（-1.61） | | | |
| LNCONFIRMED·FAI | 0.072***（9.99） | 0.079***（8.41） | 0.108**（2.11） | 0.187***（3.22） |
| LNCONFIRMED·CPI | -1.083***（-6.66） | -1.408***（-5.07） | -5.283**（-2.40） | -6.691***（-2.72） |
| LNCONFIRMED·RPI | 1.101***（6.66） | 1.432***（5.06） | 5.381**（2.40） | 6.813***（2.72） |
| W·LNCONFIRMED·FAI | 0.080*（1.95） | — | — | — |
| W·LNCONFIRMED·CPI | -5.010***（-5.58） | — | — | — |
| W·LNCONFIRMED·RPI | 5.192***（5.59） | — | — | — |
| Variance（sigma2_e） | 0.256***（8.27） | — | — | — |
| Log-likelihood | -36.097 | — | — | — |
| R-sq | 0.951 | — | — | — |
| Number of Obs. | 93 | — | — | — |

注：.***、**、*分别表示在1%、5%、10%水平上通过了显著性检验。

# 模型分析与结果Model Analysis & Results

表4检验结果显示：

- 固定资产投资和居民消费价格指数对工业经济增长的作用不显著，而本地及邻近地区商品零售价格指数的作用显著为负，主要通过直接效应对工业经济增长产生抑制作用。
- 本地及邻地新冠肺炎疫情主要通过居民消费价格指数的渠道，对工业经济增长产生负面冲击，本地新冠肺炎疫情借助居民消费对工业经济增长的直接、间接及总效应均为显著为负。
- 原因可能有两点：一是邻近地区要为本地生产防控新冠肺炎疫情所必需的医疗用品；二是由于新冠肺炎疫情最大受害对象是老年群体，而且新冠肺炎老年群体重症率要远高于其他年龄群体，因而，邻近地区群众和政府的恐慌心理会相对较轻，邻近地区保持工业生产能力，这也解释了为什么本地新冠肺炎确诊病例增加，主要通过居民消费对邻近地区工业经济增长产生影响。

# 总结与讨论Summary and Discussions

- 值得注意的是，本地及邻地新冠肺炎疫情主要通过居民消费价格指数的渠道作用于工业经济增长，本地新冠肺炎病例增加借助居民消费对工业经济增长的直接、间接及总效应均为显著为负，这说明在抗击新冠疫情过程中，不仅要注意控制省域内部疫情产生的直接效应，以充分发挥其工业经济发展的直接效应，而且要重视对邻近省域疫情防控措施和刺激居民消费政策的统筹布局，以便于加强区域之间疫情防控合作和工业经济的协同，这对于通过扩大居民消费、促进工业经济发展具有现实意义。

# 总结与讨论Summary and Discussions

- Moran’s I 指数和回归检验出现的问题应都与武汉异常值有关。将湖北的确诊人数去掉武汉(不包含武汉的湖北指标)，然后再计算，或将所有确诊人数对数化然后再计算，结果会改进（为正相关），美国和其他国家的数据也可能会有类似问题(异常值)。
- 异常值的影响与处理。可以同时报告处理前后的结果对比，这是探索性空间数据分析的要点。模型部分则重点在疫情对各省影响的差异(分析模型残差中各省实际值与估计值差值)，如果进一步分析比较各省产业结构与主要进出口产品结构，在此基础上提出如何通过调整投资、消费、商品流通，控制风险，进而提高地区工业经济系统的风险抵抗能力和安全系数。
- 可能存在遗漏变量问题，需要进行残差的空间分析，进一步开展稳健性检验。

# 参考文献References


- Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China[J]. The Lancet, 395(10223): 497-506.
- 刘志彪. 新冠肺炎疫情对中国产业的影响: 特点、风险及政策建议[J]. 东南学术, 2020(3):42-47.
- 刘世锦, 韩阳, 王大伟. 基于投入产出架构的新冠肺炎疫情冲击路径分析与应对政策[J]. 管理世界, 2020, 36(5):1-12,51,263.
- 田盛丹. 新冠肺炎疫情及其应对政策对我国宏观经济的影响——基于可计算一般均衡模型的分析[J]. 消费经济, 2020, 36(3):42-52.
- 尹彦辉, 孙祥栋, 徐朝. 新冠肺炎疫情与宏观经济波动：基于DSGE模型的分析及启示[J]. 统计与决策, 2020, 36(7):85-90.
- 何诚颖, 闻岳春, 常雅丽, 耿晓旭. 新冠病毒肺炎疫情对中国经济影响的测度分析[J]. 数量经济技术经济研究, 2020, 37(5):3-22.

# 致谢Acknowledgement


- 感谢合作团队
  - ✓南通大学孙攀博士
  - ✓华东理工大学张晨峰博士
  - ✓中国数据研究所鲍曙明博士