防疫管制与经济维稳：基于改进SEIR模型的反事实分析与跨区域交通管制的必要性探究

张晓奇

2020 年 4 月 8 日

Structure

- Ch 1: 背景;
- Ch 2: 方法;
- Ch 3: 结论;
- Ch 4: 讨论.

Outline

1 背景
- 新冠病毒带来的挑战
- NP-Net-SIR模型

2 结果
- 拟合优度
- 估计结果的描述性统计
- 交通管制的反事实分析结果
- 交通管制、输入地隔离与输出地筛查的权衡

3 结论

Motivation I

- 此次新冠疫情对我国社会经济发展带来了严重的冲击，为了防止疫情扩散，各地均不同程度地实施了非常严厉的交通管制措施，导致跨区域的人流、物流出现不同程度的停摆。
- “一刀切”式地关停跨区域流动对于全产业链的负面影响巨大，同时其防疫效果是否真的如预期那样有效，针对以上问题，现有文献并没有给与充分的重视。
- 最新的文献已经开始关注“封城”等交通管制措施对疫情控制的效果（参见Tian et al. (2020); Li et al. (2020); Fang et al. (2020)）：
  - 充分性问题：**交通管制的实施在多大程度上降低了疫情的传播扩散？**
  - 必要性问题：**在达到当前防疫效果的前提下，是否一定要采取“一刀切”的跨区域交通管制，是否存在其他的效果相近但更加温和的替代措施？**

Motivation II

- 由于全面交通管制对经济社会伤害极大，在存在大量输入性病和疫情二次爆发风险的背景下，对于交通管制的替代策略的分析至关重要。文献(Tian et al., 2020; Li et al., 2020)及其他类似文献中对交通管制的有效性的论证均建立在改良SEIR模型的基础上，然而现有不同版本的SEIR模型并不适合刻画新冠病毒的一些重要的新特征，同时也难以很好地反应我国疫情早期从武汉、湖北向其他地方的扩散机制，在此基础上的政策评估会错估交通管制的实际效果。

方法上的挑战 I

此次新冠病毒在传染能力上的新特征对传统传染病模型（SEIR模型等）提出了挑战，主要体现在以下几方面：

第一 在新冠病毒潜伏期长达十四天（甚至更长）且感染者即使自身无症状仍能在潜伏期感染他人的情况下，新冠病毒的传播应当视为潜伏期内的一个随机过程，而基于确定性建模的SIR/SEIR模型忽略了这种长滞后性与其后期内感染的随机性，导致严重低估疫情的可能性。

第二 SEIR模型及其各改进版本均为总量模型，它们仅可用于对感染总数的刻画与模拟，而此次新冠疫情的超强传播能力使得全国各地的感染数彼此联动、密切相关，SEIR模型则忽略了病毒跨区域传播所带来的区域间的互动属性，这也可能导致对最终感染数字的低估。

方法上的挑战 II

第三 考虑到新冠疫情早期，各方面对疫情严重性认知不足，反应相对不够及时，导致感染数字存在大量低报、漏报的情况，严重失真，对于模型拟合和预测提出了挑战。

第四 病毒的跨区域扩散机制没有得到在SIR/SEIR模型中没有很好的体现，而对该扩散机制的分析在很大程度上决定了后续的防疫策略选择，因此有必要进行更深入的探讨（比如考虑疫情早期的恐慌效应以及后期强制隔离措施带来的瞒报风险等）。

方法上的挑战 III

第五 SEIR模型的设定相对简洁，依赖参数较少，这在一定程度上制约了SEIR模型在对不同防疫策略进行系统性评估时的实用性，在疫情防控已取得基本胜利、复工复产成为主要政策目标的当下，如何有序放松防疫管制，恢复正常的交通、物流、人流，同时确保疫情不会二次爆发成为亟待解决的问题，而基于SEIR模型的分析难以对具体管制措施进行精准评估，因此有待改进。

Outline

1 背景
- 新冠病毒带来的挑战
- NP-Net-SIR模型

2 结果
- 拟合优度
- 估计结果的描述性统计
- 交通管制的反事实分析结果
- 交通管制、输入地隔离与输出地筛查的权衡

3 结论

模型概述 I

$$\begin{aligned}\frac{d\mathbf{n}(t)}{dt} &= \int_{t-incub}^{t} p_I(\tau,t)W(t)\cdot\mathbf{n}(\tau)d\tau - r(t)\mathbf{n}(t)\\ \mathbf{m}(t) &= \int_{t-incub}^{t} p_B(\tau,t)\mathbf{n}(\tau)d\tau\end{aligned} \tag{1}$$

其中:

- 向量$\mathbf{n}_t = (n_1(t),\ldots,n_k(t))^\top$为k各地区截止t时刻的累计潜在感染数量（此处数量理解为连续变量，在连续数量的理解下，向量**n**可以被视为刻画了一个泊松积分过程的均值过程）;

模型概述 II

- 向量$\mathbf{m}_t = (m_1(t), \ldots, m_k(t))^\top$为k个地区截至t时刻可观测感染者数量，由于潜在感染者往往由于轻症甚至无症状而不能被即刻观测到，只有随着时间的推移，潜在感染者才能被慢慢识别，因此$\mathbf{m}$与$\mathbf{n}$有一定的滞后；
- 为反映潜在感染者的滞后性，我们设在潜伏期（incub=14天）内，假设潜在感染者会以一定的概率分布p_B逐渐显露症状而被识别。由于不同时期检测能力不同，潜在感染者的识别速度也会发生变化，这可以由p_B随时间的变化来体现，因此本文假设$p_B(\cdot, t)$以连续的方式依赖于时间t。

模型概述 III

- 由于新冠疫情的潜在感染者即使不出现症状也可以感染其他人，而且在整个潜伏期内均可感染，因此使用另一个概率分布p_I刻画潜在感染者在自己被感染后在不同时间点感染其他人的可能性，同样由于防疫措施以及公众对疫情认知的加深，这一感染概率也会发生变化，因此我们让$p_I(\cdot, t)$已连续方式依赖时间t；
- 沿用经典SIR/SEIR模型的假设，任何感染者都以一定的概率r从感染状态中恢复，而随着救治之手段的实施，这一概率也会发生变化，本文因此假设恢复概率$r(t)$已连续方式依赖时间t。

模型概述 IV

- 使用区域间的病毒传播网络（表示为邻接矩阵序列$W(t)$）刻画病毒感染动态的跨区域联系。注：近期关于新冠病毒的文献中，Li et al. (2020)也加入了类似本文的传播矩阵的设定。二者的区别在于，Li et al. (2020)的文章中，假设W是静态的和已知的，W中的ij元素由地区i与地区j之间的人流强度指数进行刻画。本文对$W(t)$采用非参数设定，其具体形式须从数据中反推，此外我们要求每个$W(t)$中元素的取值范围为[0,1]区间内，这一值域约束确保我们可以将$W_{ij}(t)$理解为t时刻地区j潜在感染者的前期存量对地区i增量的贡献度的比例，由于病毒传播过程一个总量增长的过程，因此这一比例加总可以超过1。

模型(1)在形式上等价于一个动态复发神经网络（T-RNN），T-RNN是一类深度神经网络模型，被广泛应用于各类时序数据的分类和预测问题中(Liu et al., 2066)。因此，模型(1)的估计与识别可以分别通过深度学习中的梯度下降方法与正则化方法实现。

损失函数

为了解决前期感染数低报的问题，我们特别考虑如下semi-supervised learning task中的损失函数形式(Zhou and Belkin, 2011)：

$$L(\mathbf{M},\mathbf{m})=\frac{1}{kn}\sum_{i=1}^{n}\sum_{j=1}^{k}\|(M_{t_i,j}-m_j(t_i))*I_{t^*}(t_i,m_j(t_i),M_{t_i,j})\|^2 \quad (2)$$

其中$\mathbf{m}(t_i)$代表给定参数W、p_I、p_B、r下，模型(1)在t_i观测时刻生成的可观测感染数，$\mathbf{M}_{t_i}$代表t_i时刻实际观测到的感染人数，而变换函数I如下所示

$$I_{t^*}(t,m,M)=\begin{cases}m-M & if\ t\geq t^*\ \text{or}\ m>M\\ 0 & else\end{cases}, \quad (3)$$

交通管制效果评估 I

- 随着我国疫情防控的基本胜利，各省的新增感染数陆续清零，各地开始将工作重点转移至复工复产，而复工复产的前提是恢复跨区域的人流物流，而在目前各地仍在执行不同程度的交通管制措施的背景下复工复产工作的开展并不顺利。
- 在疫情初期，进行全国范围内的交通管制（甚至交通阻断）对于组织输入性感染源以及降低地方感染爆发风险至关重要，这一基本结论也在近期国际上发表的多篇回溯性分析论文中的到了证实。因此，是否放松交通管制以及如何放松成为各个地方在复工复产与疫情防控之间进行选择时的一个难题。
- 与现有文献中的反事实分析不同，我们不再聚焦于实施交通管制相较于不管制的情况是否有效降低了潜在的新冠感染病例或延迟了疫情爆发高峰;
- 我们关注上述问题的反命题：**在不引起各地疫情二次爆发的前提下，是否存在放松或局部放松交通管制的可执行方案，以及在什么时间点以后，上述放松方案就已具备可行性?**
- In the other words: 除了目前的方案外，是否存在其他更宽松的替代方案，即可维持同等的疫情防控效果，同时减小对经济的伤害。

交通管制效果评估 II

为此，本节介绍一个评估跨区域交通管制的疫情防控效果反事实分析框架，在该框架中，上述必要性问题可以被转化为一个带约束的最优化问题，其中约束条件为若改变交通管制策略，则在策略调整时刻t^s（本文分别选取1月23日、2月2日、2月10日）之后，在新策略下的疫情防控效果应当至少不比继续维持当前管制策略的效果更糟，而为了度量防空效果，本文考虑如下三个不同度量指标：

$$R0(W^r(t_i), r(t_i)) \leq \min(R0(W(t_i), r(t_i)), 1), \forall t_i \geq t^s \tag{4}$$

$$\mathbf{m}^r(t_i) \leq \mathbf{m}(t_i), \forall t_i \geq t^s \tag{5}$$

$$\mathbf{D}^r(t_i) \leq \mathbf{D}(t_i), \forall t_i \geq t^s \tag{6}$$

其中，W^r、$\mathbf{m}^r$代表在更新的交通管制策略下计算的感染传播网络矩阵与可观测的感染人数，时变的向量$\mathbf{D}$、$\mathbf{D}^r$代表每个时刻各个地区在当前管制策略下的新冠病毒死亡病例总数和在更新的交通管制策略下哦的死亡总数。$R0$代表新冠病毒的基础再生数，在NP-Net-SIR该指标是一个动态变化的数值，它依赖于感染传播网络矩阵$W(t)$和恢复系数$r(t)$。

交通流量与感染连接强度的回归分析 I

为了将交通管制策略的强度与新冠病毒传播的动力学模型(1)进行连接，考虑如下回归方程：

$$W_{kj}(t_i) = \alpha + \beta \overline{\mathbf{T}}_{kj}(t_i) + \varepsilon_{kji} \tag{7}$$

其中，$\overline{T}(t)$矩阵为t时点之前时刻各地区间的人流强度矩阵依感染滞后概率$p_I(\cdot, t)$加权平均而成的人流矩阵，即：

$$\overline{\mathbf{T}}(t_i) = \sum_{j=0}^{incub-1} p_I(j, t_i)\mathbf{T}(t_i - incub + j). \tag{8}$$

其中，$\mathbf{T}(t)$为t时刻的瞬时人流强度矩阵（通过百度迁徙指数获得）。

反事实优化

在系数β与残差项ε估计值的基础上，本文反事实分析将设定一个人流放松矩阵$\mathbf{r}$，其中元素$\mathbf{r}_{ij} \geq 0$，对于调整时刻t^s之后的每一个观测点t_i，本文假设$\overline{T}^r(t_i) = \overline{T}(t_i) + \mathbf{r}$，因此$\mathbf{r}$将通过式(8)、(7)影响感染传播网络矩阵$W$进而通过模型(1)影响各期潜在的和可观测的感染数、基础再生数$R0$和死亡数。

综上，本文的反事实分析的目的是分别在约束条件(4)-(6)下求解以下最优化问题：

$$\max \sum_{1 \leq j,l \leq k} \mathbf{r}_{jl},\ \mathbf{r}_{jl} \geq 0 \tag{9}$$

如果问题(9)存在非零解，即最优的$\mathbf{r}$矩阵存在非零元素，则意味着当前的交通管控可以被合理放松。

交通管制与其他防疫策略的替代性 I

- 随着1月23日武汉封城措施的实施，全国各地均相继实施了不同程度的防疫管控措施，除了跨区域交通管制外，很多城市和地区还对外来人口自到达之日起实施强制居家观察/隔离、取消本地大型公共集会、关闭公共场所（如学校、餐馆、商场等），由于此类措施并不会直接作用于跨区域的人流，同时又会降低本地与其他地区的感染链接强度，因此这类措施相较于整体性的封城和交通管制对经济发展负面影响更小。
- 除了输入地拉长社交距离的防疫措施，输出地也不同程度地实施了对外流人员的筛查力度，典型措施包括在各个外出交通要塞（如高速、国道、省道的出/入省/市的出入口、长途汽车站、火车站、飞机场等）设立健康监测点，这类措施有效地降低了输出人口中潜在感染者的比例，但对输出的人流强度影响较小，因此对经济伤害较小。
- 以上两大类非全面交通管制的防疫措施既有助于疫情防控，同时可已经可能减小对经济的损伤，值得深入探讨。

交通管制与其他防疫策略的替代性 II

为了评估此类措施的影响，本文拓展了回归模型(7)，分别引入了代表在输入地采取的隔离、关闭公共场所等措施与在输出地实的施潜在病例筛查等措施的待估参数，其数值大小可以反映了相应措施的实施力度及其对感染链接强度的抑制效果。其中输入地措施参数由$in_k(t_i)$表示，输出地措施参数由$out_j(t_i)$表示。二者引入回归方程(7)后得到如下回归方程：

$$W_{kj}(t_i) = \alpha + (\beta - out_j(t_i)) \cdot \overline{T}_{kj}(t_i) - in_k(t_i) + \varepsilon'_{kji} \tag{10}$$

在(10)中，输入地与输出地的相关措施的实施起点分别选取为2020年1月26日与2020年1月30日Tian et al. (2020)。

在模型估计时，考虑两步法：首先考虑到β代表自然状态下人流强度对病毒感染强度的影响，因此本文先采用模型(7)的形式与武汉1月23日封城前的感染强度数据与人流数据估计β；此后，将β估计值带入模型(10)，此时(10)变为普通线性模型，采用最小二乘估计其余参数。

Outline

1 背景
- 新冠病毒带来的挑战
- NP-Net-SIR模型

2 结果
- 拟合优度
- 估计结果的描述性统计
- 交通管制的反事实分析结果
- 交通管制、输入地隔离与输出地筛查的权衡

3 结论

拟合优度 I

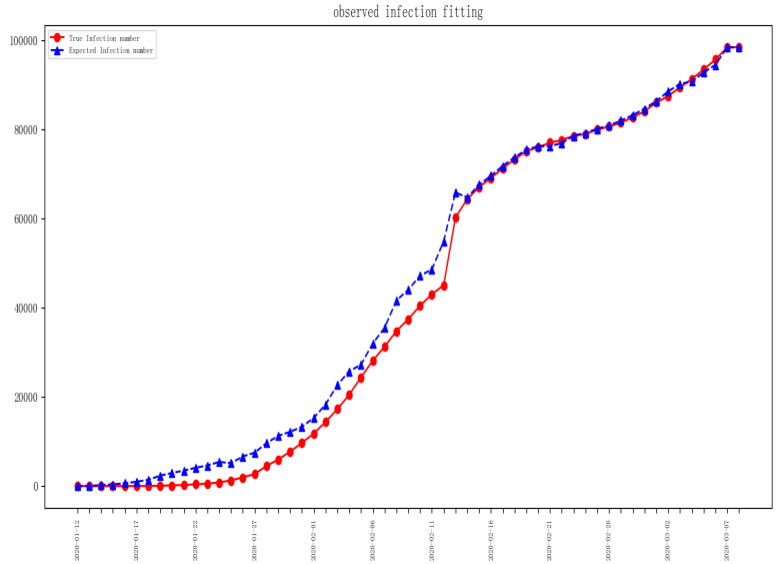
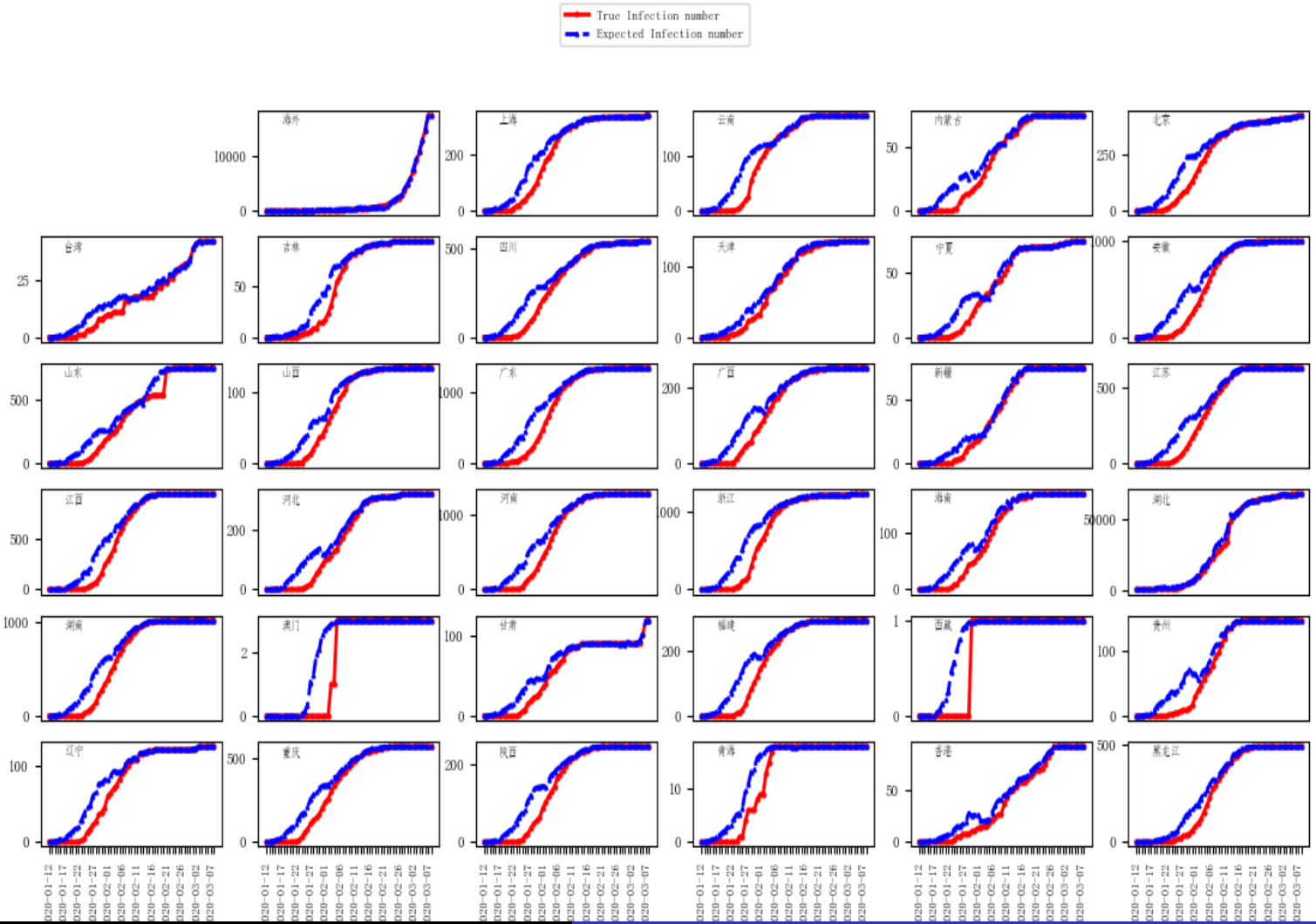


图: 全国加总的确诊人数（可观测感染数）拟合效果

拟合优度 II


True Infection number
Expected Infection number
海外
上海
云南
内蒙古
北京
台湾
吉林
四川
天津
宁夏
安徽
山东
山西
广东
广西
新疆
江苏
江西
河北
河南
浙江
海南
湖北
湖南
澳门
甘肃
福建
西藏
贵州
辽宁
重庆
陕西
青海
香港
黑龙江

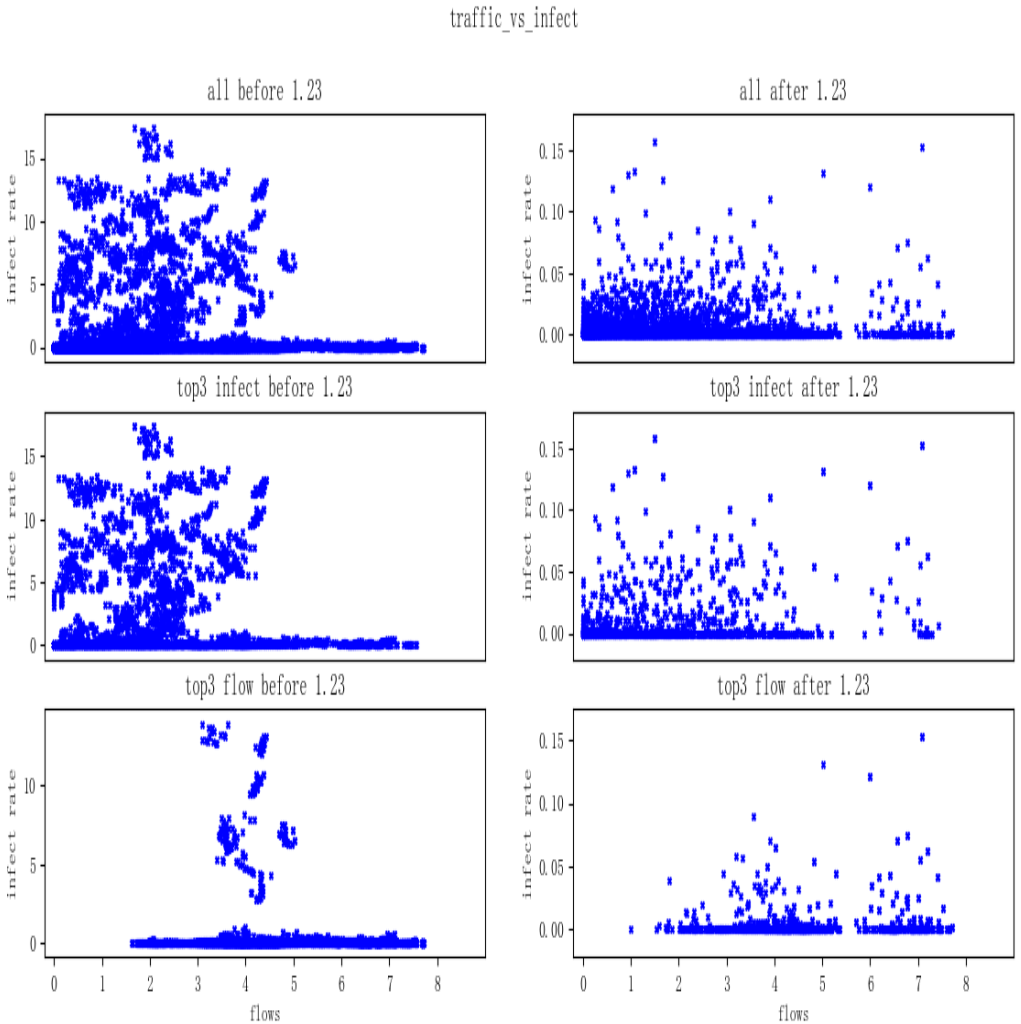
Outline

1 背景
- 新冠病毒带来的挑战
- NP-Net-SIR模型

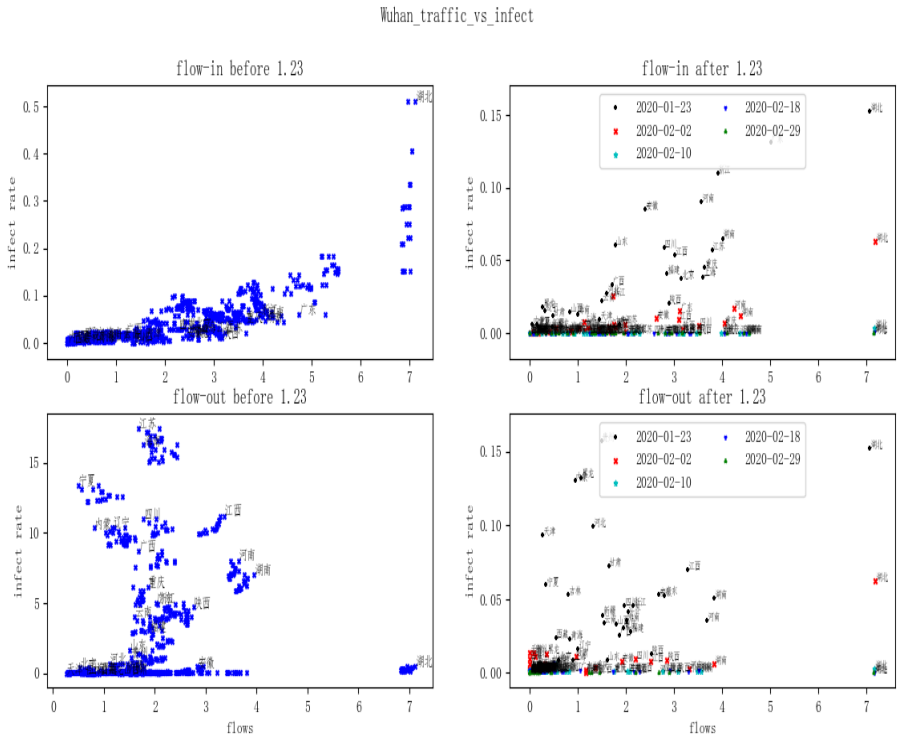
2 结果
- 拟合优度
- 估计结果的描述性统计
- 交通管制的反事实分析结果
- 交通管制、输入地隔离与输出地筛查的权衡

3 结论

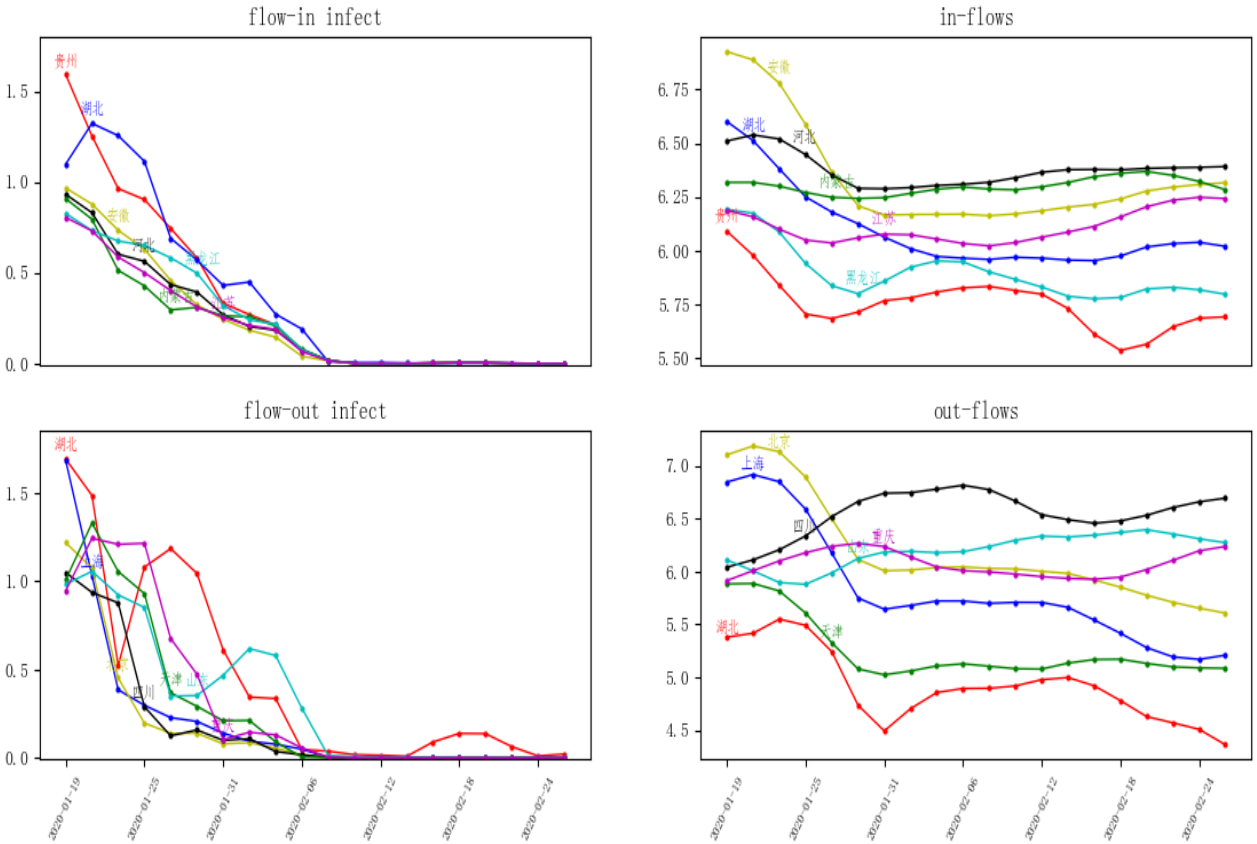
感染强度W与省际交通流量



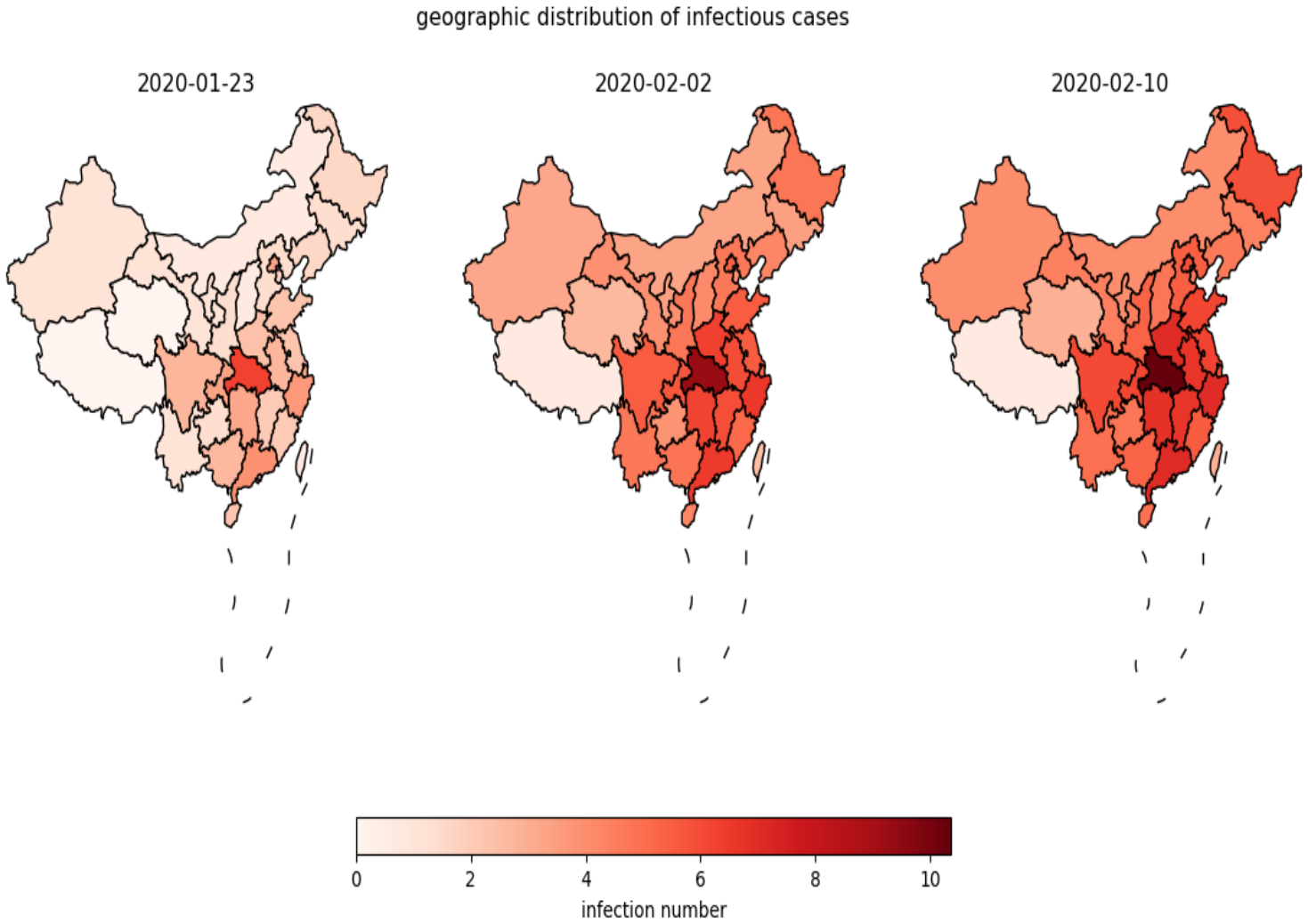
感染强度W与省际交通流量：以武汉为中心



感染强度W与省际交通流量：协变趋势（感染强度top 7）



三时点的感染数、总人流入、流出的地理分布



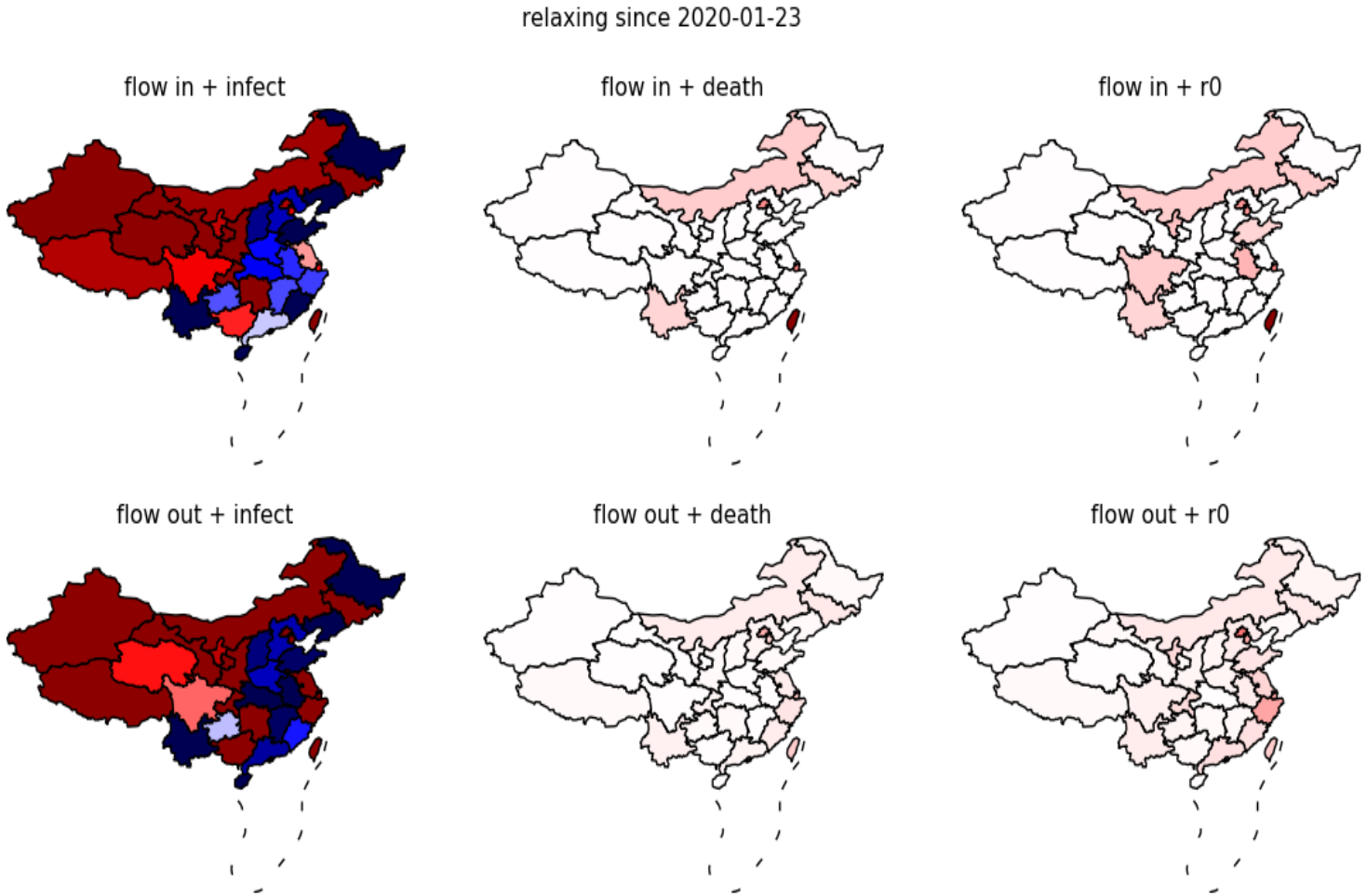
Outline

1 背景
- 新冠病毒带来的挑战
- NP-Net-SIR模型

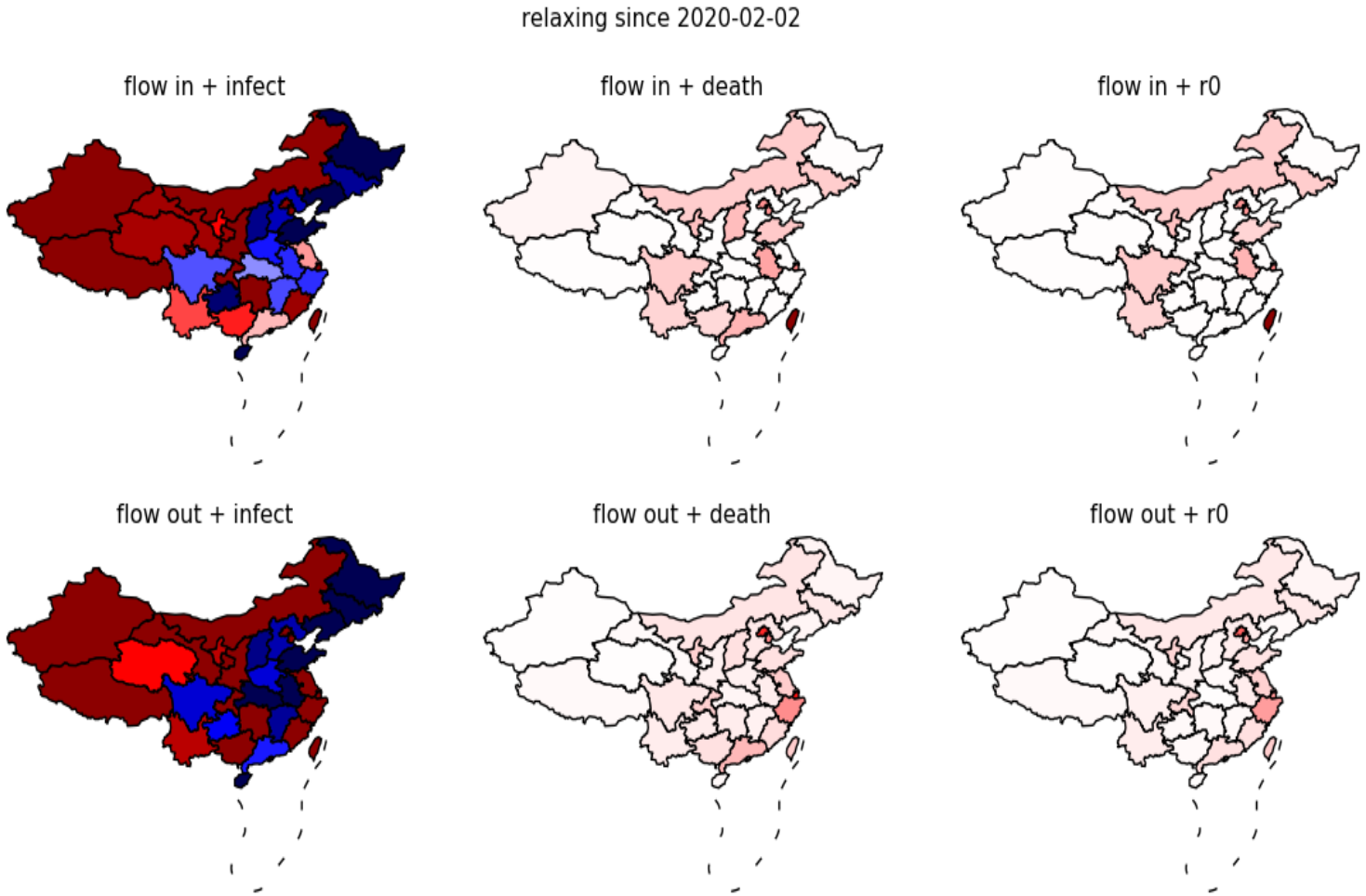
2 结果
- 拟合优度
- 估计结果的描述性统计
- 交通管制的反事实分析结果
- 交通管制、输入地隔离与输出地筛查的权衡

3 结论

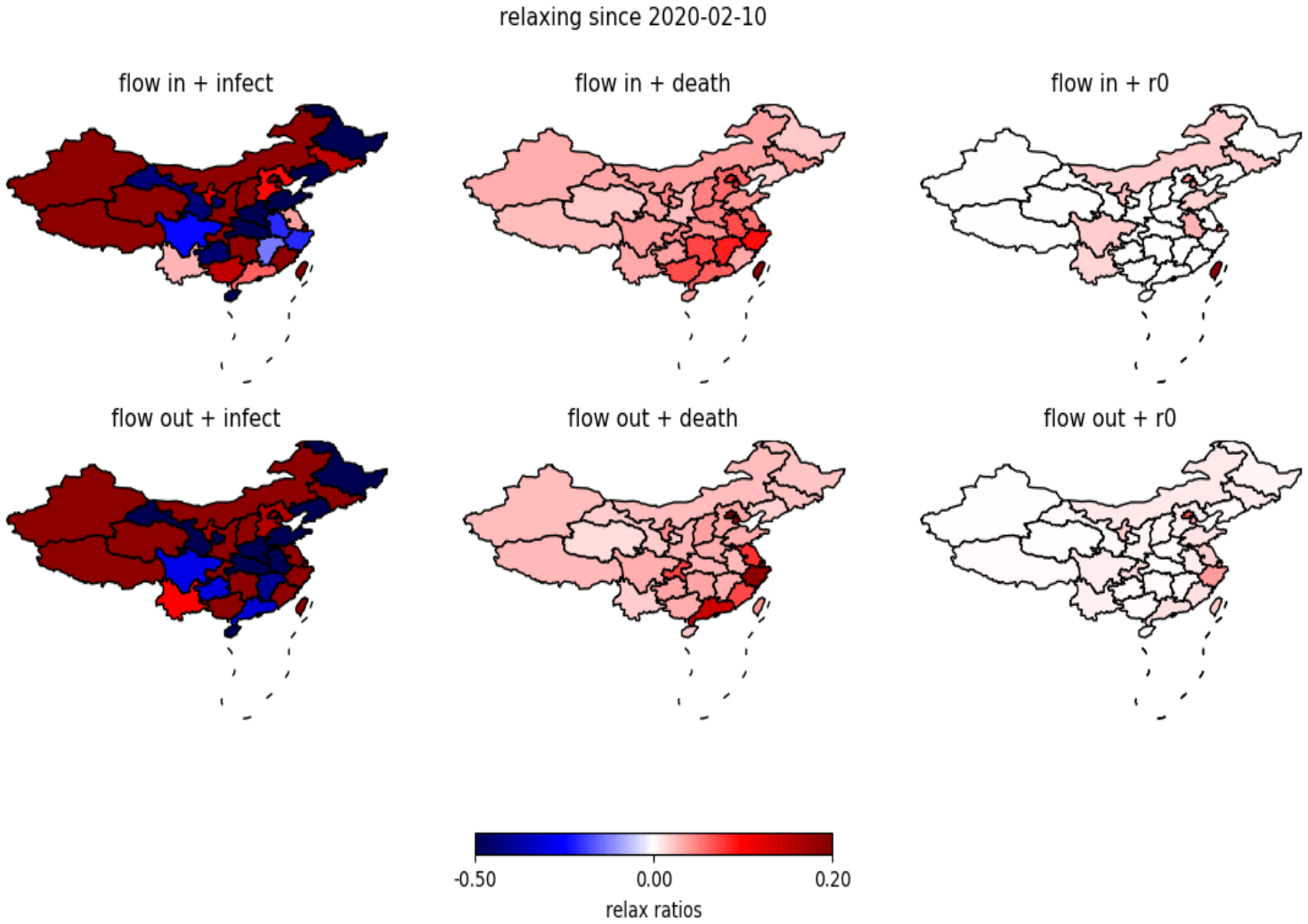
1月23日的放松管制方案



2月2日的放松管制方案



2月10日的放松管制方案



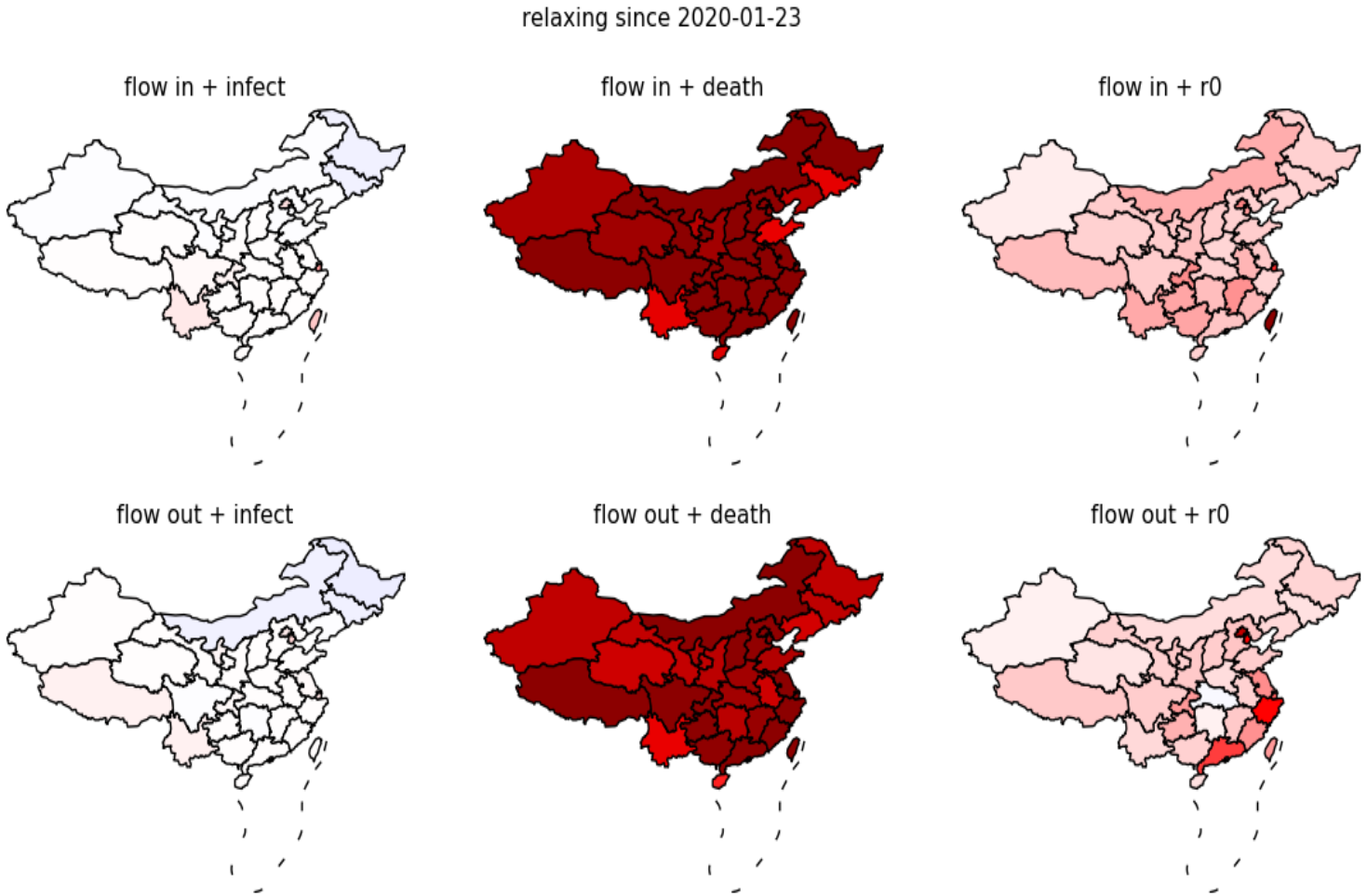
Outline

1 背景
- 新冠病毒带来的挑战
- NP-Net-SIR模型

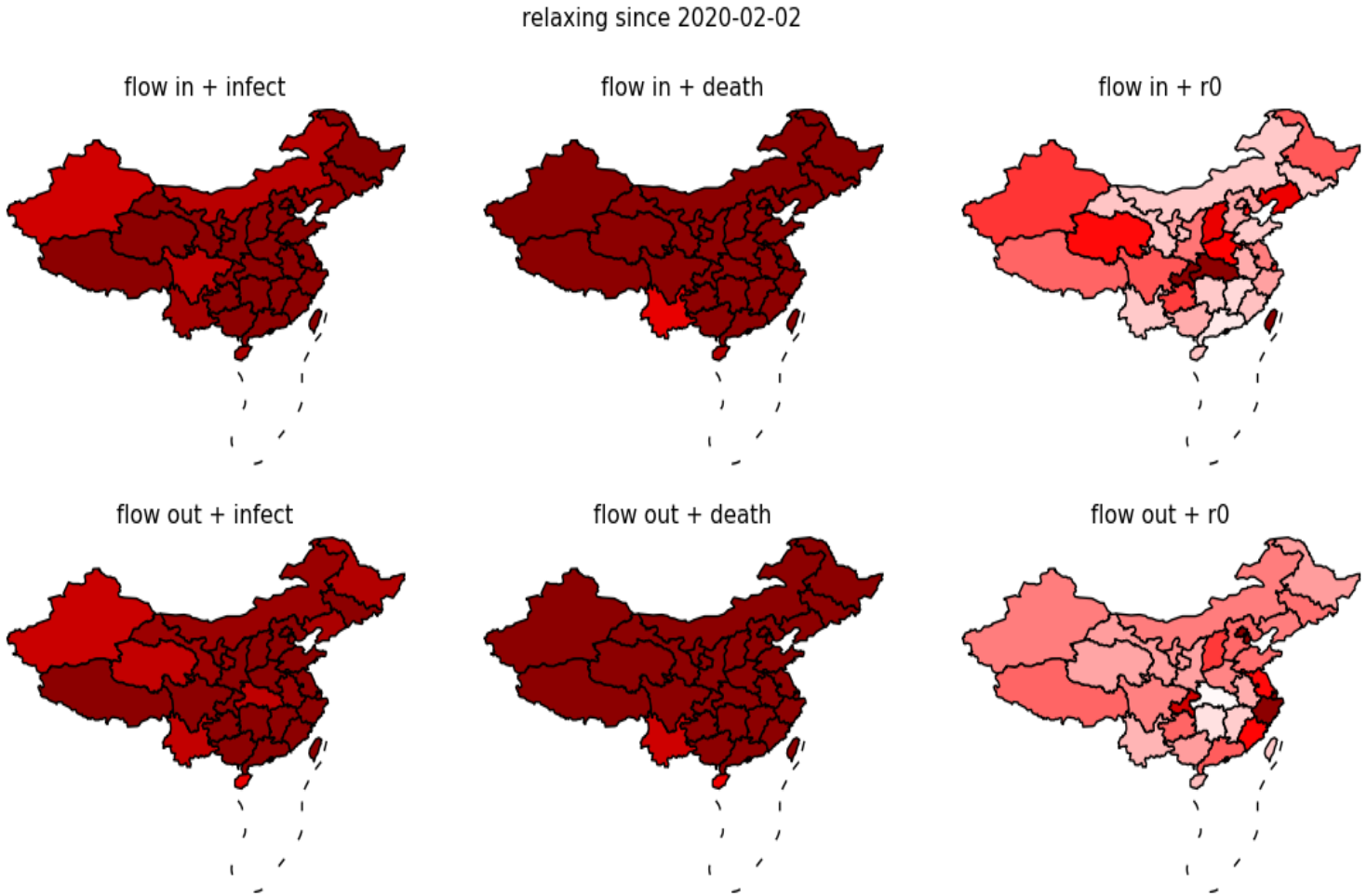
2 结果
- 拟合优度
- 估计结果的描述性统计
- 交通管制的反事实分析结果
- 交通管制、输入地隔离与输出地筛查的权衡

3 结论

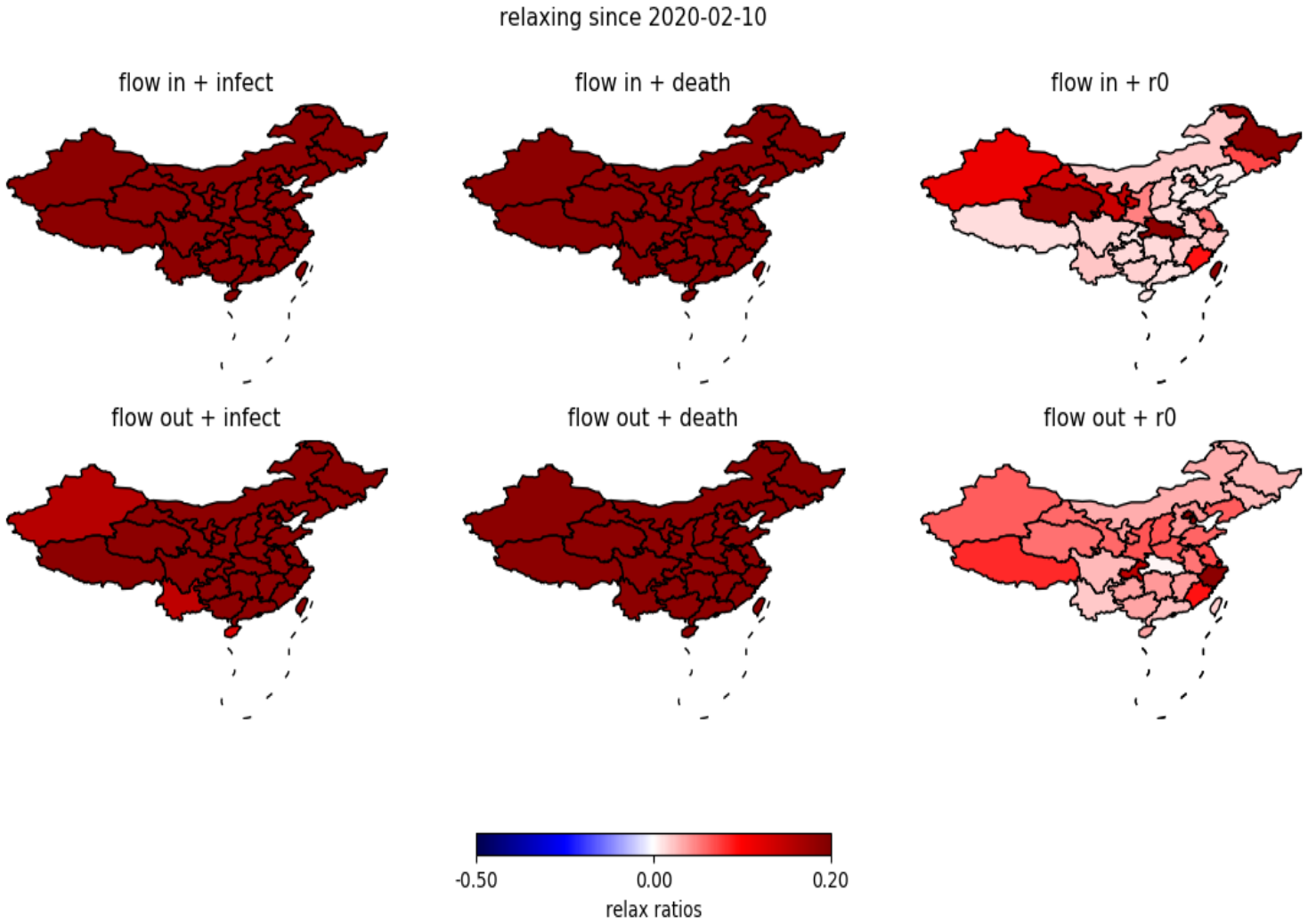
1月23日的放松管制方案



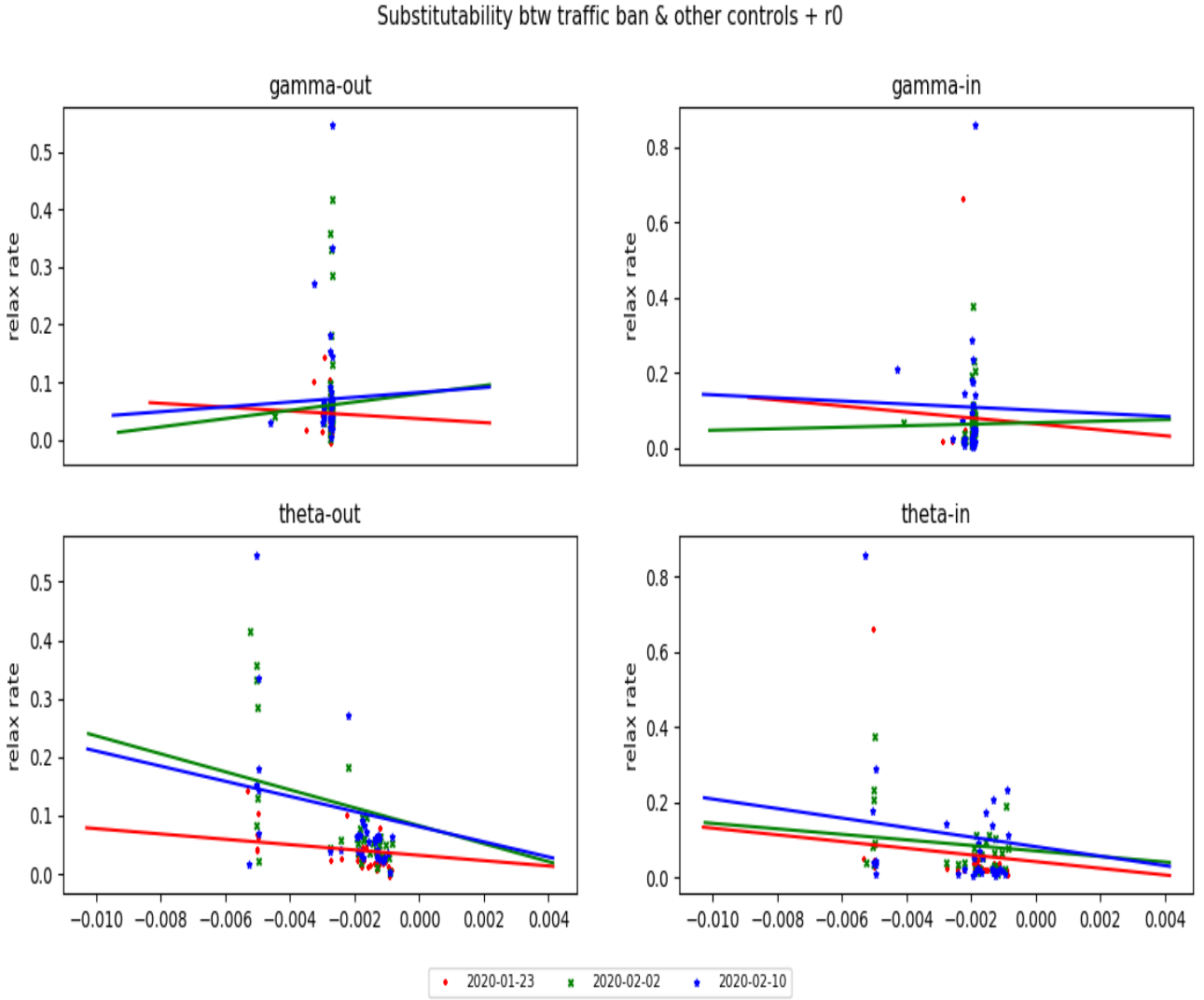
2月2日的放松管制方案



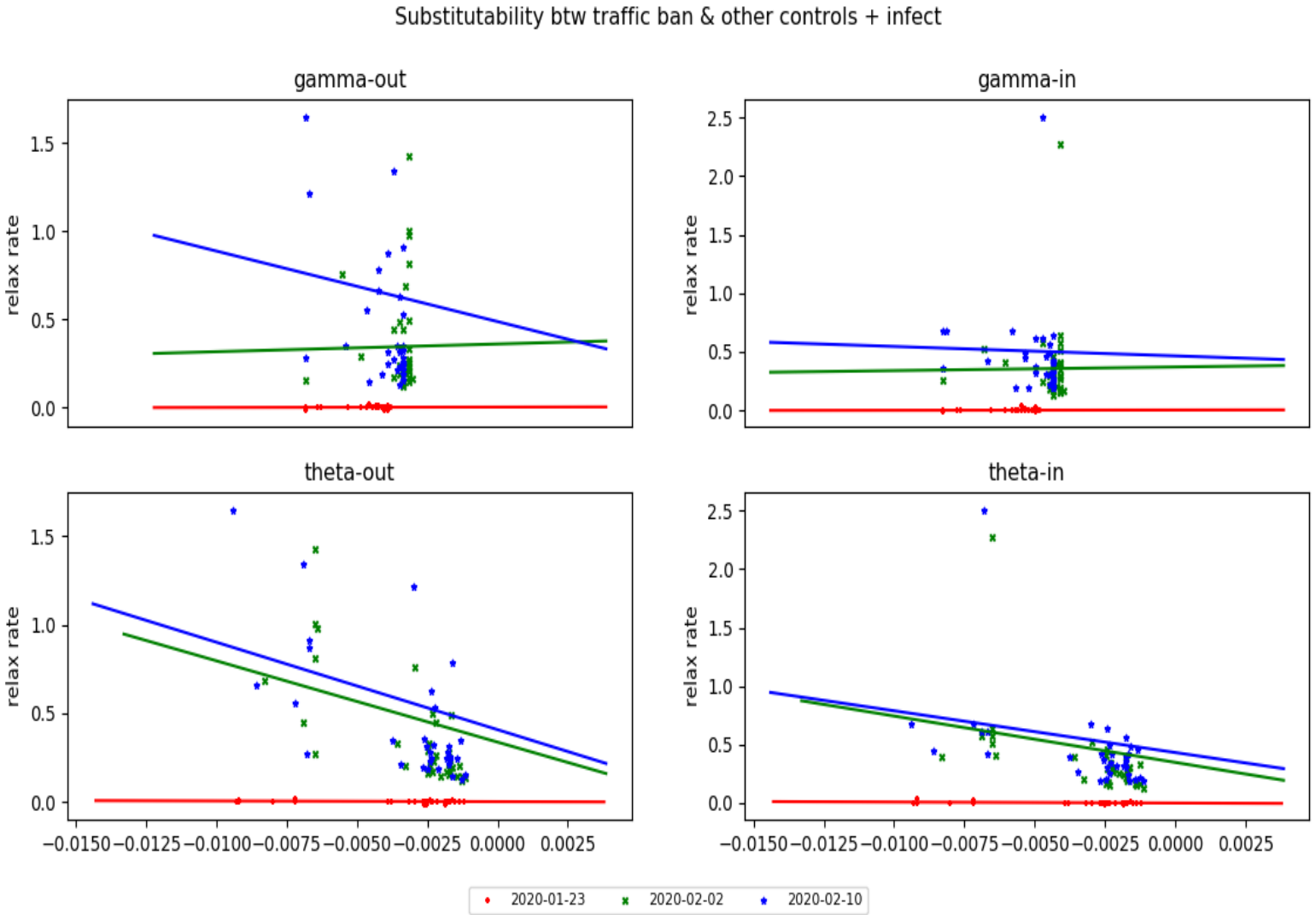
2月10日的放松管制方案



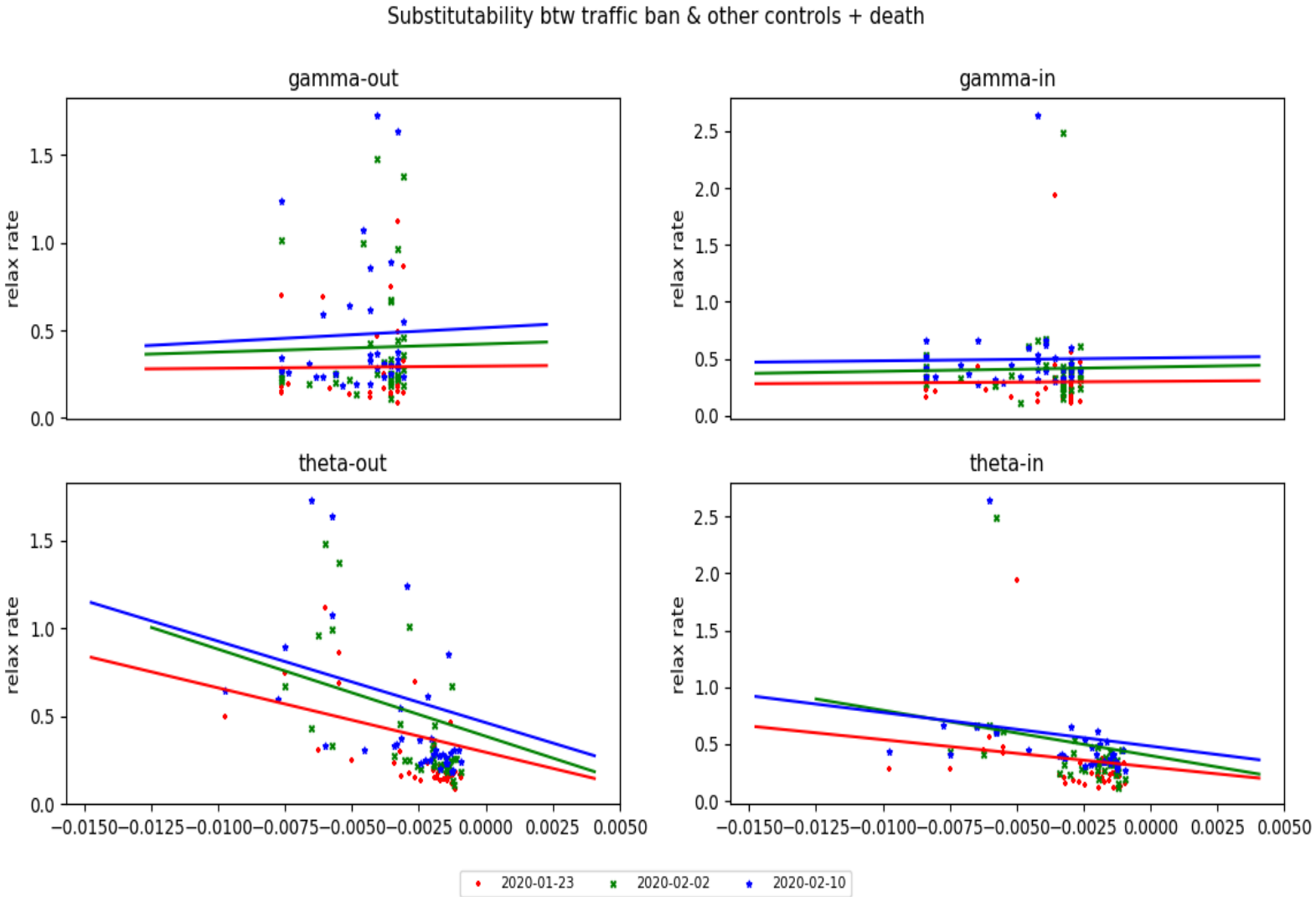
交通管制与其他非管制措施的替代性：R0



交通管制与其他非管制措施的替代性：感染总数



交通管制与其他非管制措施的替代性：死亡总数



结论

- “一刀切”式的交通管制/封城措施尽管在疫情早期的确起到了控制疫情扩散的作用，但长时间的交通封锁并不是控制防疫所必需的，存在着放松交通管制同时不影响疫情恶化的帕累托改进的防疫策略。
- 在输出地与输入地采取全面筛查和有针对性的social distancing策略，将在很大程度上替代全面交通管制的防疫效果，更严格地实施这些策略可以为全面放松交通管制赢取更大的空间。同时，相较于全面交通管制，前述策略对经济的负面影响较小，因此为了迎合更好的复工复产，可以考虑充分利用非管制策略对管制策略的替代性。

Bibliography: I

Liu Q., Wu S., Wang L. and Tan T. Predicting the next location: A recurrent model with spatial and temporal contexts. In *Thirtieth AAAI conference on artificial intelligence*, 2016.

Sadhanala V., Wang Y. and Tibshirani R. Total variation classes beyond 1d: Minimax rates, and the limitations of linear smoothers. *Advances in Neural Information Processing Systems*, 3513–3521, 2016.

Zhou X. and Belkin M. Semi-supervised learning by higher order regularization. *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, 892–900, 2011.

Tian H., Liu Y., Li Y. and others The impact of transmission control measures during the first 50 days of the COVID-19 epidemic in China. *Science*, 2020.

Li R., Pei S., Chen B., Song Y., Zhang T., Yang W. and Shaman J. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2). *Science*, 2020.

Fang H., Wang L. and Yang Y. Human Mobility Restrictions and the Spread of the Novel Coronavirus (2019-nCoV) in China. *National Bureau of Economic Research*, 2020.

- Thank You!