

案例分析：世界杯是否影响卡塔尔新冠疫情？

史长浩

2023 年 1 月 9 日

摘要

本文基于增广合成对照方法，研究了世界杯对卡塔尔新冠疫情的影响，结论为在 0.05 的显著性水平下，世界杯的举办确实加重了卡塔尔国内的疫情形势。但这一结论具有局限性，更严格的讨论需要基于更高质量的数据，并对不可观测混杂进行更为细致的分析。

关键字： 卡塔尔世界杯，新冠疫情，潜在结果模型，合成对照

1 案例背景

卡塔尔世界杯于 2022 年 11 月 20 日开幕，吸引了超 120 万球迷到场看球，为疫情以来首届不限观众人数的大型体育赛事。而一年前，同样在疫情期间举办的日本夏季奥运会、北京冬奥会则是“全封闭”的大型赛事。卡塔尔世界杯几乎全面放开，恢复了 2019 年之前的赛事形态。

卡塔尔当局自 11 月 1 日起取消了几几乎所有与新冠疫情相关的旅行限制措施，包括抵达前的强制核酸检测，不再提供疫苗证明、注册相关健康码的信息，以及在比赛时不需要佩戴口罩等。而球迷们主要来自德、意、法、日、巴西等疫情仍然高位运行的国家，他们的到来，是否会让卡塔尔这个仅有 250.5 万人的小国，变成新的疫情暴发地？

本文尝试基于增广合成对照方法研究这一问题。

2 数据说明

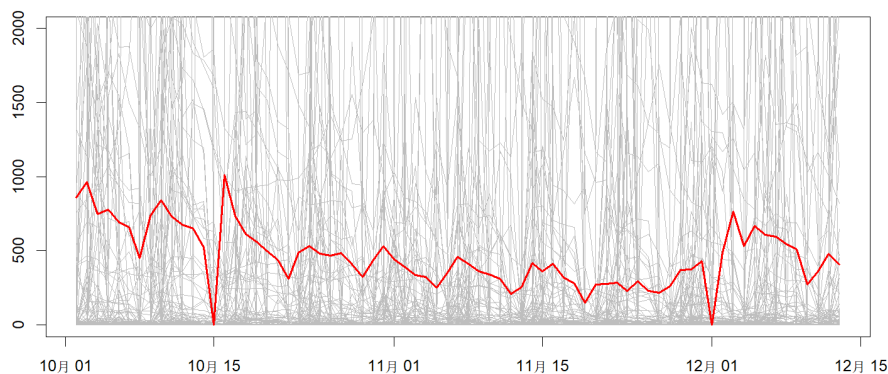


图 1: 128 个国家的日增病例数据可视化。横轴为日期，纵轴为确诊人数（个）。红色实线对应卡塔尔。

本案例所用数据来自公益项目 [Our World in Data](#)，本文从中截取了 2022 年 10 月 2 日到 12 月 13 日共 73 天的新增病例数据。因为卡塔尔地处连接欧、亚、非三洲的中东地区，故只选择地处这三洲的其他国家作为对照，剔除一些无效记录之后，得到 127 个国家作为对照组，如图1所示，红色实线对应卡塔尔新冠确诊人数随时间变化曲线，浅灰色实线则对应其他 127 个国家（未标准化）。由于不同国家新冠疫情强度不同、规律不同、政策不同，因此存在较大差异是正常的，这并不影响后续方法的使用。

卡塔尔世界杯于 11 月 20 日开幕，故认为自开幕日起，卡塔尔的新增确诊人数受到赛事干预。本文关心的问题即：若不举行世界杯，11 月 20 日之后卡塔尔的新增病例数是否与当前观测有显著区别？

表 1 展示了潜在结果变量的符号，其中 $Y_{i,j}(0)$ 表示未受世界杯干预的单元，而 $Y_{i,j}(1)$ 表示受到世界杯干预的单元。为了评估世界杯对卡塔尔国内疫情的影响，本质上需要对任意 $t > T_0$ 估计 $Y_{1,t}(1) - Y_{1,t}(0)$ ，核心问题在于估计缺失值 $Y_{1,t}(0)$ 。一种想法是基于其他国家的病例数“合成”卡塔尔 11.20 之后的病例数作为对照 $\hat{Y}_{1,t}(0) = \sum_{n=2}^N \gamma_n Y_{n,t}(0)$ ，这即是合成对照法的思想。

表 1: 10 月 2 日至 12 月 13 日 128 个国家新增新冠确诊病例数。

	10.2	...	11.19	11.20	...
Qatar	$Y_{1,1}(0)$...	$Y_{1,T_0}(0)$	$Y_{1,T_0+1}(1)$...
German	$Y_{2,1}(0)$...	$Y_{2,T_0}(0)$	$Y_{2,T_0+1}(0)$...
\vdots	\vdots	\ddots	\vdots	\vdots	...
Japan	$Y_{N,1}(0)$...	$Y_{N,T_0}(0)$	$Y_{N,T_0+1}(0)$...

3 方法介绍

基于一致性假定，建立观测数据与潜在结果之间的联系，得式 (1)，其中 $W_i = 1$ 当且仅当 $i = 1$ ，即只有卡塔尔可能受到世界杯干预。

$$Y_{it} = \begin{cases} Y_{it}(0) & \text{if } W_i = 0 \text{ or } t \leq T_0, \\ Y_{it}(1) & \text{if } W_i = 1 \text{ and } t > T_0. \end{cases} \quad (1)$$

为了强调预处理结果是被当作协变量对待的，之后使用 X_{it} 代替 Y_{it} ，对任意的 $t \leq T_0$ 。合成对照方法归结为求解约束优化问题 (2)[1]：

$$\begin{aligned} \min_{\gamma} & \left\| \mathbf{V}_{\mathbf{x}}^{1/2} (\mathbf{X}_1 - \mathbf{X}'_0 \gamma) \right\|_2^2 + \zeta \sum_{W_i=0} f(\gamma_i), \\ \text{subject to} & \sum_{W_i=0} \gamma_i = 1, \\ & \gamma_i \geq 0 \quad i : W_i = 0. \end{aligned} \quad (2)$$

其中 $\mathbf{V}_{\mathbf{x}} \in \mathbb{R}^{T_0 \times T_0}$ 是对称重要性矩阵 (symmetric importance matrix)，为了符号简单计，本文取其为单位阵； $f(\gamma_i)$ 是对权重的分散惩罚。

设 \hat{m}_{iT} 是 $Y_{iT}(0)$ 的估计量，则定义 $Y_{1T}(0)$ 的增广合成对照方法 (ASCM) 估计量为式 (3) 所示 [2]。

$$\begin{aligned} \hat{Y}_{1T}^{\text{aug}}(0) &= \sum_{W_i=0} \hat{\gamma}_i^{\text{scm}} Y_{iT} + \left(\hat{m}_{1T} - \sum_{W_i=0} \hat{\gamma}_i^{\text{scm}} \hat{m}_{iT} \right) \\ &= \hat{m}_{1T} + \sum_{W_i=0} \hat{\gamma}_i^{\text{scm}} (Y_{iT} - \hat{m}_{iT}), \end{aligned} \quad (3)$$

其中, $\hat{\gamma}_i^{\text{scm}}$ 是前文 SCM 方法得到的权重, 标准的 SCM 估计是该估计的一个特例, 只需令 \hat{m}_{iT} 为常数。原则上 SCM 只在处理前结果拟合较好时才可使用, 而 ASCM 通过权重的修正拓宽了 SCM 的使用场景。

若 $\hat{m}(\mathbf{X}_i)$ 通过岭回归确定, 即 $\hat{m}(\mathbf{X}_i) = \hat{\eta}_0^{\text{ridge}} + \mathbf{X}_i' \hat{\eta}^{\text{ridge}}$, 其中

$$\left\{ \hat{\eta}_0^{\text{ridge}}, \hat{\eta}^{\text{ridge}} \right\} = \arg \min_{\eta_0, \eta} \frac{1}{2} \sum_{W_i=0} (Y_i - (\eta_0 + \mathbf{X}_i' \eta))^2 + \lambda^{\text{ridge}} \|\eta\|_2^2, \quad (4)$$

则得到岭 ASCM 估计:

$$\hat{Y}_{1T}^{\text{aug}}(0) = \sum_{W_i=0} \hat{\gamma}_i^{\text{scm}} Y_{iT} + \left(\mathbf{X}_1 - \sum_{W_i=0} \hat{\gamma}_i^{\text{scm}} \mathbf{X}_i \right) \cdot \hat{\eta}^{\text{ridge}}, \quad (5)$$

岭 ASCM 是本案例分析中主要使用的方法。

4 结果分析

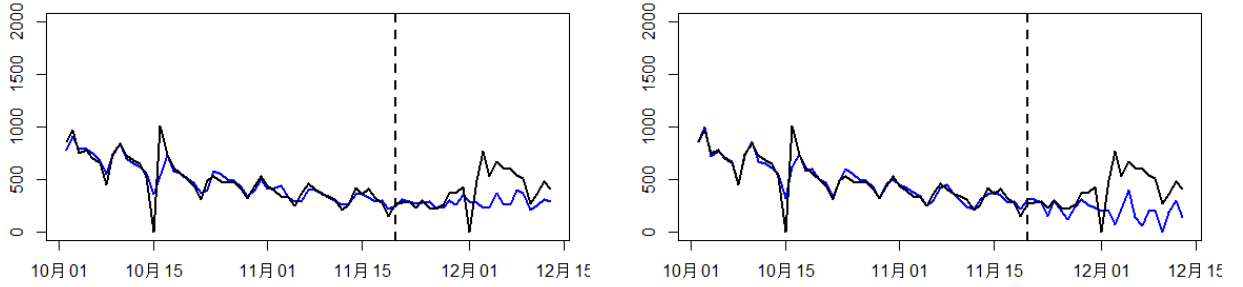


图 2: 卡塔尔确诊病例随时间变化曲线。黑色为实际观测, 蓝色为合成对照 (左: SCM; 右: Ridge ASCM)。

图2展示了卡塔尔确诊病例随时间变化曲线, 其中黑色实线为实际观测值, 蓝色实线为合成对照值。由图可见在 11 月 20 日之前合成对照方法较好地拟合了观测数据, 且 Ridge ASCM 相较于 SCM 略胜一筹。计算结果也表明, Ridge ASCM 相较于 SCM 实现了对处理前结果更好的拟合: $\text{MSE}^{\text{ascm}} = 6352.01$, $\text{MSE}^{\text{scm}} = 9421.58$ 。

进一步可以计算平均处理效应, Ridge ASCM 相较于 SCM 的处理组中的平均处理效应 $\text{ATT} = \frac{1}{T-T_0} \sum_{t=T_0+1}^T (Y_{1,t}(1) - Y_{1,t}(0))$ 更大: $\text{ATT}^{\text{ascm}} = 186.67$, $\text{ATT}^{\text{scm}} = 111.08$ 。在 sharp 零假设 ($H_0: Y_{1,t}(0) = Y_{1,t}(1), \forall t > T_0$) 下, 可以作置换检验, 检验的 p 值分别为 0.013 (SCM) 和 0.023 (ASCM), 即两种方法都表明处理组的处理效应是显著的, 世界杯确实使得卡塔尔的疫情更加严重。

图3展示了卡塔尔确诊病例实际观测与合成对照之差随时间的变化曲线, 从此图可更直观地看出 Ridge ASCM 相较于 SCM 对预处理结果的拟合更好。图中灰色区域为 95% 置信带。

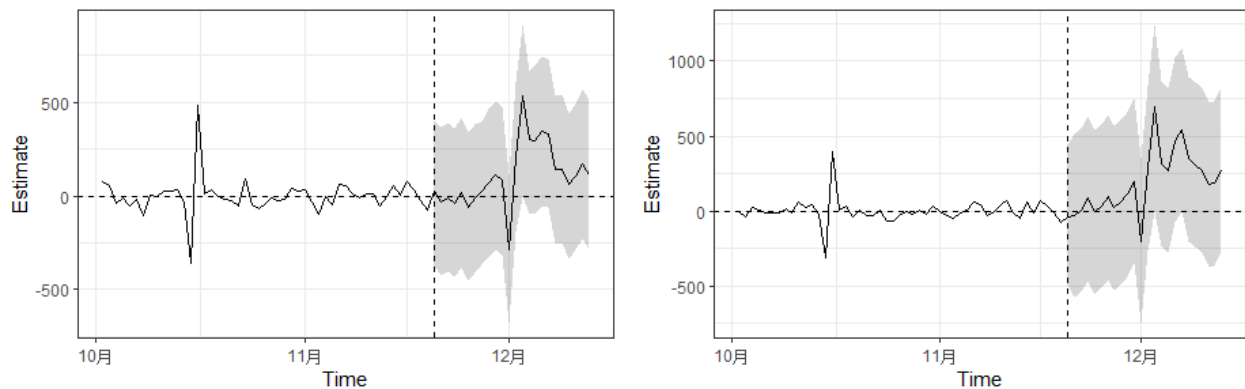


图 3: 卡塔尔确诊病例实际观测与合成对照之差随时间变化曲线。(左: SCM; 右: Ridge ASCM)。

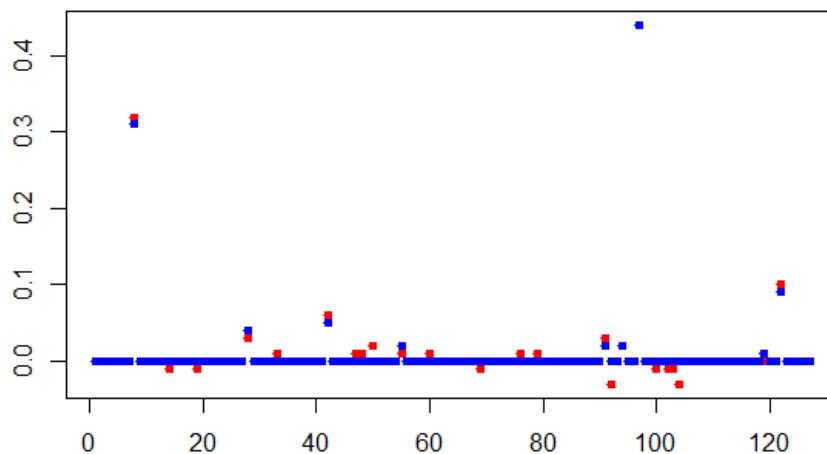


图 4: Ridge ASCM 与 SCM 合成对照组的权重对比。横轴为国家编号, 纵轴为权重大小。蓝色对应 SCM, 红色对应 ASCM。

图4展示了 Ridge ASCM 与 SCM 合成对照组的权重对比, 可见 SCM 的权重更加稀疏, 127 个国家之中, 只有 9 个国家权重不为 0, 且严格非负; Ridge ASCM 允许出现负权, 且有 23 个国家具有非零的权。

比较有趣的一点是, SCM 与 ASCM 赋予最高权重的 3 个国家, 恰好是卡塔尔的 3 个邻国: 沙特阿拉伯 (SCM:0.44, ASCM:0.44), 巴林岛 (0.31, 0.32), 阿拉伯联合酋长国 (0.09, 0.10)。这似乎表明, 对于新冠疫情, 地理上的相邻性确实会导致疫情的同步性, 这对未来研究或许会有所启发。

实验结果的代码可从<https://github.com/RUC-ChangHao/Scm-Qatar>获得。

5 讨论

合成对照方法还可通过引入额外的协变量信息来提高拟合的效果, 本文实验中考虑了疫苗接种率、人均 GDP、人均病床数等协变量, 但这些协变量的加入并没有使得实验效果有显著变化, 故不在论文中呈

现。这可能的原因有很多，有可能这些变量确实确实不会影响整体估计的趋势，也有可能由于各国统计能力的差异，这些数据存在较大的问题等。

关于世界杯是否真的对卡特尔当地的疫情防控产生消极影响，由于存在诸多不可控因素，该问题非常复杂，本文仅从数据层面给出一种尝试性回答，并不完善。更严格的讨论需要更谨慎地筛选数据，更细致地分析混杂因素，这是未来进步的方向。

参考文献

- [1] A. ABADIE, A. DIAMOND, AND J. HAINMUELLER, *Synthetic control methods for comparative case studies: Estimating the effect of california' s tobacco control program*, Journal of the American statistical Association, 105 (2010), pp. 493–505.
- [2] E. BEN-MICHAEL, A. FELLER, AND J. ROTHSTEIN, *The augmented synthetic control method*, Journal of the American Statistical Association, 116 (2021), pp. 1789–1803.