合成控制法 + 合成双重差分

Synthetic Control Method + Synthetic Diff-in-Diff

杨点溢

Department of Government
London School of Economics and Political Science

Spark 社科量化系列课程

本节课内容

1. 双重差分法的问题 Problems with Diff-in-Diff
2. 合成控制法 Synthetic Control Method
3. 合成双重差分 Synthetic Difference-in-Differences

DID 的问题：没有平行趋势

- 我们知道对于双重差分法 (DID) 来说，很重要的假设是平行趋势 (parallel trends)

DID 的问题：没有平行趋势

- 我们知道对于双重差分法 (DID) 来说，很重要的假设是平行趋势 (parallel trends)
  - 悲伤的是，现实的趋势不总是平行的

DID 的问题：没有平行趋势

- 我们知道对于双重差分法 (DID) 来说，很重要的假设是平行趋势 (parallel trends)
  - 悲伤的是，现实的趋势不总是平行的
  - 更悲伤的是，不平行的话很容易被看出来（易于检验）

DID 的问题：没有平行趋势

- 我们知道对于双重差分法 (DID) 来说，很重要的假设是平行趋势 (parallel trends)
  - 悲伤的是，现实的趋势不总是平行的
  - 更悲伤的是，不平行的话很容易被看出来（易于检验）
  - 所以符合平行假设时真的不能使用 DID

DID 的问题：没有平行趋势

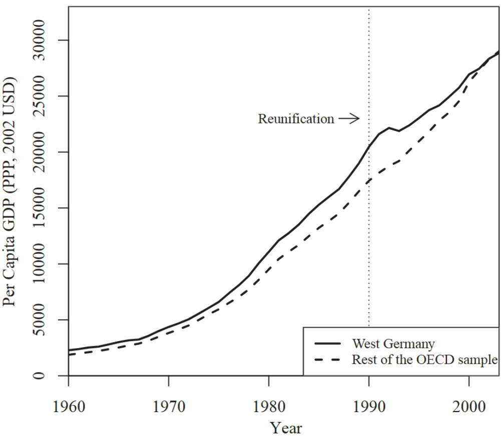
- 我们知道对于双重差分法 (DID) 来说，很重要的假设是**平行趋势** (parallel trends)
  - 悲伤的是，现实的趋势不总是平行的
  - 更悲伤的是，不平行的话很容易被看出来（易于检验）
  - 所以符合平行假设时真的不能使用 DID
- 案例：两德统一对西德的经济影响 (Abadie, Diamond, and Hainmueller 2015)
  - 假说：两德统一拖累了西德的经济
  - 结果变量：人均 GDP (PPP, 2002USD)
  - 处理组：西德
  - 控制组/对照组：OECD 其他国家

DID 的问题：没有平行趋势

- 我们知道对于双重差分法 (DID) 来说，很重要的假设是平行趋势 (parallel trends)
  - 悲伤的是，现实的趋势不总是平行的
  - 更悲伤的是，不平行的话很容易被看出来（易于检验）
  - 所以符合平行假设时真的不能使用 DID
- 案例：两德统一对西德的经济影响 (Abadie, Diamond, and Hainmueller 2015)
  - 假说：两德统一拖累了西德的经济
  - 结果变量：人均 GDP (PPP, 2002USD)
  - 处理组：西德
  - 控制组/对照组：OECD 其他国家
  - 悲伤：并不是很平行

案例：两德统一对西德的经济影响 (Abadie, Diamond, and Hainmueller 2015)

FIGURE 1 **Trends in per Capita GDP: West Germany versus Rest of the OECD Sample**



DID 的其它问题

- DID（或者说 TWFE）要求处理组和对照组的差距可被归为
  - Time-invariant, unit-specific measures, or 单位固定效应
  - time-varying in the same way for all units 时间固定效应
  - SCM（合成控制法）减弱了这个假设

DID 的其它问题

- DID（或者说 TWFE）要求处理组和对照组的差距可被归为
  - Time-invariant, unit-specific measures, or 单位固定效应
  - time-varying in the same way for all units 时间固定效应
  - SCM（合成控制法）减弱了这个假设
- DID 在样本量小的时候面临着较大的 se
  - SCM 没完全解决这个问题，但是我们可以用别的推断方法 (inference methods)

本节课内容

1. 双重差分法的问题 Problems with Diff-in-Diff
2. 合成控制法 Synthetic Control Method
3. 合成双重差分 Synthetic Difference-in-Differences

合成控制法的初衷

- 有时候我们很难找到和处理组真的相像的控制组

合成控制法的初衷

- 有时候我们很难找到和处理组真的相像的控制组
- 与此同时我们有**一些**和处理组比较像，但不完全一样的控制组 $\Rightarrow$ 在 SCM 里叫做 Donor pool
    - 不满足平行趋势假设

合成控制法的初衷

- 有时候我们很难找到和处理组真的相像的控制组
- 与此同时我们有一些和处理组比较像，但不完全一样的控制组 $\Rightarrow$ 在 SCM 里叫做 Donor pool
  - 不满足平行趋势假设
- 我们可以通过算法用 Donor pool “捏出”，即“合成”出一个和处理组相像的控制组 $\Rightarrow$ 合成控制法

合成控制法的初衷

- 有时候我们很难找到和处理组真的相像的控制组
- 与此同时我们有一些和处理组比较像，但不完全一样的控制组 ⇒ 在 SCM 里叫做 Donor pool
  - 不满足平行趋势假设
- 我们可以通过算法用 Donor pool“捏出”，即“合成”出一个和处理组相像的控制组 ⇒ 合成控制法
- 在两德统一的例子中，我们希望通过 OECD 国家“合成”出一个假想中从来没统一过的西德。

合成控制法基本信息

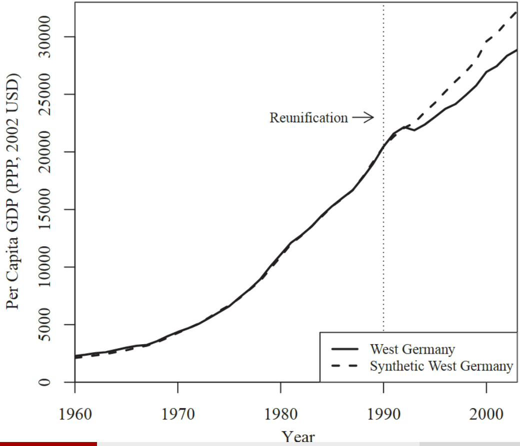
- 合成控制发 SCM 使用 Donor Pool 的加权平均来模拟处理单位（处理组一般只有 1 个单位）
  - 使用处理前 pre-treatment 时间段的各国的结果变量来计算出一个权重，使得“合成控制”单位的处理前趋势和处理单位最为相似
  - 权重是我们结果的一部分
  - 比如 $Germany = 0.5 France + 0.25 UK + 0.25 Poland$（非真实例子）

合成控制法基本信息

- 合成控制发 SCM 使用 Donor Pool 的加权平均来模拟处理单位（处理组一般只有 1 个单位）
  - 使用处理前 pre-treatment 时间段的各国的结果变量来计算出一个权重，使得“合成控制”单位的处理前趋势和处理单位最为相似
  - 权重是我们结果的一部分
  - 比如 $Germany = 0.5France + 0.25UK + 0.25Poland$（非真实例子）
- 合成控制法 SCM 是定性研究和定量研究之间的桥梁
  - 应该在 Donor Pool 里面涵盖哪些国家/地区？$\Rightarrow$ 对案例的了解
- 经常用来研究一个事件或者处理对国家/地区的影响

回到两德统一的例子 (Abadie, Diamond, and Hainmueller 2015)

FIGURE 2 Trends in per Capita GDP: West Germany versus Synthetic West Germany



SCM 的设定：时间

- 时间由 t 表示，$t \in \{1, ..., T_0, ..., T\}$
- 处理（事件)在 T_0 之后发生
- 我们试图计算 T_{0+1} 之后的处理效应(ATT)
  - 有 T_0 个处理前 (pre-treatment/pre-intervention) 时间段,$\{1, ..., T_0\}$
  - 有 T_1 个处理前 (pre-treatment/pre-intervention) 时间段,$\{T_0 + 1, ..., T\}$
  - $T = T_0 + T_1$

SCM 的设定：单位 Units

- 单位由 i 表示，$i \in \{1, ..., J+1\}$
  - 1 是被处理单位
  - J, 或者说第 $\{2, 3, ..., J+1\}$ 个单位，是控制组，或者说"donor pool"
- Donor pool 需要满足以下条件：

SCM 的设定：单位 Units

- 单位由 i 表示，$i \in \{1, ..., J+1\}$
  - 1 是被处理单位
  - J, 或者说第 $\{2, 3, ..., J+1\}$ 个单位，是控制组，或者说"donor pool"
- Donor pool 需要满足以下条件：
  - Untreated (during the observation period) （在此时间段内）没受到处理

SCM 的设定：单位 Units

- 单位由 i 表示，$i \in \{1, ..., J+1\}$
  - 1 是被处理单位
  - J, 或者说第 $\{2, 3, ..., J+1\}$ 个单位，是控制组，或者说"donor pool"
- Donor pool 需要满足以下条件：
  - Untreated (during the observation period)（在此时间段内）没受到处理
  - 不受处理影响 No spillover

SCM 的设定：单位 Units

- 单位由 i 表示，$i \in \{1, ..., J+1\}$
  - 1 是被处理单位
  - J, 或者说第 $\{2, 3, ..., J+1\}$ 个单位，是控制组，或者说"donor pool"
- Donor pool 需要满足以下条件：
  - Untreated (during the observation period) （在此时间段内）没受到处理
  - 不受处理影响 No spillover
  - No large "idiosyncratic shocks" 自身的结果变量 (Y) 不应该受大的、独有的因素影响

SCM 的设定：单位 Units

- 单位由 i 表示，$i \in \{1, ..., J+1\}$
  - 1 是被处理单位
  - J, 或者说第 $\{2, 3, ..., J+1\}$ 个单位，是控制组，或者说"donor pool"
- Donor pool 需要满足以下条件：
  - Untreated (during the observation period) （在此时间段内）没受到处理
  - 不受处理影响 No spillover
  - No large "idiosyncratic shocks" 自身的结果变量 (Y) 不应该受大的、独有的因素影响
  - 和处理单位 (1) 相像，从而避免插值偏差 (interpolation bias)
    - 尽管 SCM 在一定程度上会避免这个问题

SCM 的设定：单位 Units

- 单位由 i 表示，$i \in \{1, ..., J+1\}$
  - 1 是被处理单位
  - J, 或者说第 $\{2, 3, ..., J+1\}$ 个单位，是控制组，或者说"donor pool"
- Donor pool 需要满足以下条件：
  - Untreated (during the observation period) （在此时间段内）没受到处理
  - 不受处理影响 No spillover
  - No large "idiosyncratic shocks" 自身的结果变量 (Y) 不应该受大的、独有的因素影响
  - 和处理单位 (1) 相像，从而避免插值偏差 (interpolation bias)
    - 尽管 SCM 在一定程度上会避免这个问题
  - 处理单位不能高于/低于所有 donor pool 单位
    - 用前东欧国家捏不出刚改开的中国？($\Rightarrow$ 期待一波 SDID)
    - 权重必须非负，加起来等于 1（之后讲）
    - "The treated outcome (Y_{it}) is outside the convex hull of that of the donor pool(Y_{0t})"

SCM 的设定：结果 Outcomes (Y)

- $Y_t^{tr} = Y_{1t}$ =outcome for treated unit 处理单位的结果
- Y_t^{sc} =outcome for synthetic control unit 合成控制单位单位的结果
- Y_{0t} 为 donor pool 所有单位 ($j = 2, ..., J + 1$) 的结果 Y 在时间 t 的集合 (vector).

SCM 的设定：结果 Outcomes (Y)

- $Y_t^{tr} = Y_{1t}$ =outcome for treated unit 处理单位的结果
- Y_t^{sc} =outcome for synthetic control unit 合成控制单位单位的结果
- Y_{0t} 为 donor pool 所有单位 ($j = 2, ..., J + 1$) 的结果 Y 在时间 t 的集合 (vector).
- 我们只能观察到 Y_{it}^{tr}，不能观察到 Y_{it}^{sc}
- 我们假设两者在 $t \leq T_0$（处理前）时一样

SCM 的设定：结果 Outcomes (Y)

- $Y_t^{tr} = Y_{1t}$ =outcome for treated unit 处理单位的结果
- Y_t^{sc} =outcome for synthetic control unit 合成控制单位单位的结果
- Y_{0t} 为 donor pool 所有单位 ($j = 2, ..., J + 1$) 的结果 Y 在时间 t 的集合 (vector).
- 我们只能观察到 Y_{it}^{tr}，不能观察到 Y_{it}^{sc}
- 我们假设两者在 $t \leq T_0$（处理前）时一样
- 处理效应 $\tau_{it} = Y_{it}^{tr} - Y_{it}^{sc}$

SCM 的设定：结果 Outcomes (Y)

- $Y_t^{tr} = Y_{1t}$ =outcome for treated unit 处理单位的结果
- Y_t^{sc} =outcome for synthetic control unit 合成控制单位单位的结果
- Y_{0t} 为 donor pool 所有单位 ($j = 2, ..., J + 1$) 的结果 Y 在时间 t 的集合 (vector).
- 我们只能观察到 Y_{it}^{tr}，不能观察到 Y_{it}^{sc}
- 我们假设两者在 $t \leq T_0$（处理前）时一样
- 处理效应 $\tau_{it} = Y_{it}^{tr} - Y_{it}^{sc}$
    - 图像上来讲，为两条线在 $t > T_0$(处理后) 的竖直距离
    - 平均处理效应 ATT= $\frac{1}{T_1} \sum_{t=t_0+1}^{T} \tau_{it}$

SCM 的设定：结果 Outcomes (Y)

- $Y_t^{tr} = Y_{1t}$ =outcome for treated unit 处理单位的结果
- Y_t^{sc} =outcome for synthetic control unit 合成控制单位单位的结果
- Y_{0t} 为 donor pool 所有单位 ($j = 2, ..., J+1$) 的结果 Y 在时间 t 的集合 (vector).
- 我们只能观察到 Y_{it}^{tr}，不能观察到 Y_{it}^{sc}
- 我们假设两者在 $t \leq T_0$（处理前）时一样
- 处理效应 $\tau_{it} = Y_{it}^{tr} - Y_{it}^{sc}$
  - 图像上来讲，为两条线在 $t > T_0$(处理后) 的竖直距离
  - 平均处理效应 ATT= $\frac{1}{T_1} \sum_{t=t_0+1}^{T} \tau_{it}$
- Z_{it} 为协变量 covariates 的矩阵
  - 如果加控制变量的话

SCM：权重 weights

就如我们刚才讲的，合成控制的结果为 donor pool 的加权平均

$$Y_t^{sc} = \sum_{j=2}^{J+1} \omega_j^* Y_{jt}$$

- 其中 ω^* (omega star) 为第 j 个控制单位在合成控制中的最优权重
- $W^* = \{\omega_2, \omega_3, ..., \omega_{J+1}\} \Rightarrow$ 最优权重的集合 (vector).

那怎么求这个权重呢？

SCM：最优化求解 optimisation problem /1

在不加协变量 covariates 时，

$$\widehat{W}^* = \arg\min_{W}\{\sum_{t=1}^{T_0}(Y_{1t} - Y_{0t}W)^2\}.$$

- $Y_{0t}W = \widehat{Y}_t^{sc}$
- 翻译成人话：寻找到一个权重，让处理前的两条线捏到一起（线之间最小二乘）
- 限定条件：$\sum_{j=2}^{J+1}\omega_j^* = 1$
  - 权重加起来等于 1

SCM：最优化求解 optimisation problem /2

在加协变量 Z_{it} 时，

$$\widehat{W}^* = \arg\min_W \{\sum_{z=0}^{Z} v_z \sum_{t=1}^{T_0} (X_{zt}^{tr} - X_{zt}^{ct} W)^2\}$$

其中：

SCM：最优化求解 optimisation problem /2

在加协变量 Z_{it} 时，

$$\widehat{W}^* = \arg\min_W \{\sum_{z=0}^{Z} \upsilon_z \sum_{t=1}^{T_0} (X_{zt}^{tr} - X_{zt}^{ct} W)^2\}$$

其中：

- X_{zt}^{tr} 包含了处理单位的第 z 个特征在第 t 个时间段的数值.

SCM：最优化求解 optimisation problem /2

在加协变量 Z_{it} 时，

$$\widehat{W}^{*} = \arg\min_{W}\{\sum_{z=0}^{Z} v_z \sum_{t=1}^{T_0} (X_{zt}^{tr} - X_{zt}^{ct} W)^2\}$$

其中：

- X_{zt}^{tr} 包含了处理单位的第 z 个特征在第 t 个时间段的数值.
  - Y 本身也是特征之一 ($z = 0$ 代表 Y)$\Rightarrow X_{0t} = Y_t$

SCM：最优化求解 optimisation problem /2

在加协变量 Z_{it} 时，

$$\widehat{W}^* = \arg\min_{W}\{\sum_{z=0}^{Z} \upsilon_z \sum_{t=1}^{T_0} (X_{zt}^{tr} - X_{zt}^{ct} W)^2\}$$

其中：

- X_{zt}^{tr} 包含了处理单位的第 z 个特征在第 t 个时间段的数值.
  - Y 本身也是特征之一 ($z = 0$ 代表 Y)$\Rightarrow X_{0t} = Y_t$
- X_{zt}^{ct} 包含了 Donor pool 的第 z 个特征第 t 个时间段的 $J \times 1$ 一维矩阵 (vector).

在加协变量 Z_{it} 时，

$$\widehat{W}^* = \arg\min_W \{\sum_{z=0}^{Z} \upsilon_z \sum_{t=1}^{T_0} (X_{zt}^{tr} - X_{zt}^{ct} W)^2\}$$

其中：

- X_{zt}^{tr} 包含了处理单位的第 z 个特征在第 t 个时间段的数值.
  - Y 本身也是特征之一 ($z = 0$ 代表 Y)$\Rightarrow X_{0t} = Y_t$
- X_{zt}^{ct} 包含了 Donor pool 的第 z 个特征第 t 个时间段的 $J \times 1$ 一维矩阵 (vector).
- υ_z(upsilon) 为特征的权重
  - 越重要的（即越能预测 Y 的）协变量权重越高

计算影响

如同前文所讲，结果计算可以通过：

$$\tau_{it} = Y_{it}^{tr} - Y_{it}^{sc}$$

$$ATT = \bar{\tau}_{it} = \frac{1}{T_1} \sum_{t=t_0+1}^{T} \tau_{it}$$

Arkhangelsky et al. (2021) 认为 SCM 的平均处理效应 ATT 的计算也可以用回归问题 Regression problem 的形式写出：

$$\left(\widehat{\Delta\tau}^{sc}, \ \widehat{\mu}, \ \widehat{\beta}\right) = \arg\min_{\tau,\mu,\beta} \left\{ \sum_{i=1}^{N} \sum_{t=1}^{T} \left(Y_{it} - \mu - \beta_t - W_{it}\tau\right)^2 \widehat{w}_i^{sc} \right\}.$$

衡量拟合度：RMSPE

- The Root Mean Squared Prediction Error (RMSPE) 衡量的是模型的拟合度.
  - 处理前应该拟合度高，处理后应该拟合度低（由于处理效应）。
- 处理前的 RMSPE：

$$RMSPE_{pre} = \sqrt{\frac{1}{T_0}\sum_{t=1}^{T_0}(Y_{1t} - \sum_{j=2}^{J+1}\omega_j^* Y_{jt})^2}$$

- 处理之后的 RMSPE：

$$RMSPE_{post} = \sqrt{\frac{1}{T_1}\sum_{t=T_0+1}^{T}(Y_{1t} - \sum_{j=2}^{J+1}\omega_j^* Y_{jt})^2}$$

Hypothesis Testing 假说检验

Placebo Method （安慰剂检验）:

1. 计算处理单位的处理后 RMSPE 和处理前 RMSPE 的比例 (Ratio of post-treatment RMSPE to pre-treatment RMSPE)

$$Ratio_1 = \frac{RMSPE_{post}}{RMSPE_{pre}}$$

2. 重复的为 donor pool 的每个单位做一次合成控制
3. 然后为 donor pool 的每个单位也按照上述公式计算一次比例 Ratio
4. 按照从高到低 (desending order) 将得到的比例 Ratio 结果排序
5. 计算处理单位在排行中的靠前程度:

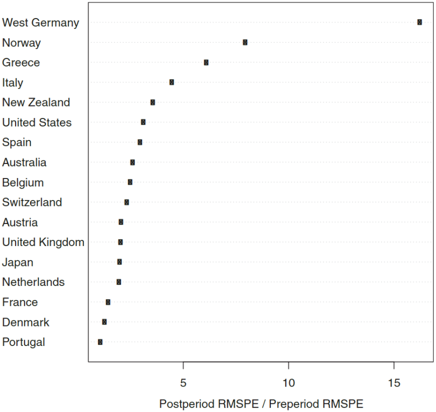
$$\text{P-value} = \frac{Rank}{Total}$$

原理: 由于 donor pool 没有处理效应, Ratio 应该接近 1

- p-value: 如果我们从列表中随机抽取一个单位, 他的比例 ratio 至少和处理单位一样高的几率

假说检验案例：两德统一 (Abadie, Diamond, and Hainmueller 2015)

FIGURE 5 Ratio of Postreunification RMSPE to Prereunification RMSPE: West Germany and Control Countries

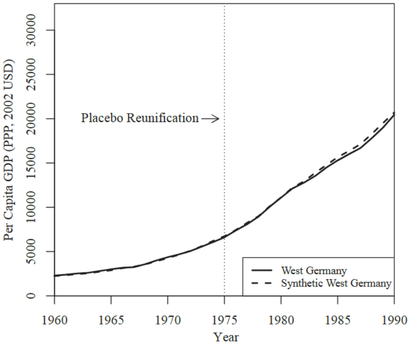


$$p = 1/17 \simeq 0.059$$

In-time placebo test

In-time placebo test 检验了模型的可信度。（使用处理时间前的一个时间作为 Placebo treatment date，来看模型的预测度）。理论上来讲，预测线应该和实际线重合：

FIGURE 4 Placebo Reunification 1975–Trends in per Capita GDP: West Germany versus Synthetic West Germany

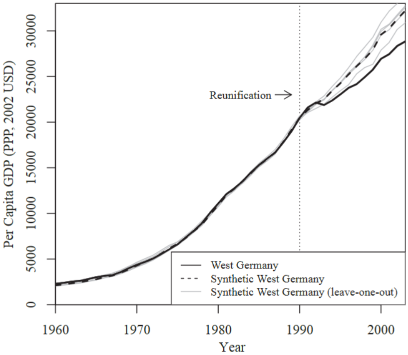


稳健性检验 Robustness Test

Leave-one-out test:

- 将 donor pool 中有正权重的单位每次踢出去一个，再跑 SCM
- 损失了一些准确性，但是能看到 ATT 的变化幅度
- 如果变化幅度小，证明结果稳健

FIGURE 6 Leave-One-Out Distribution of the Synthetic Control for West Germany



SCM 的变种

由于 SCM 的一些问题（convex hull 等），也有一些魔改的 SCM 版本可以参考

- 合成双重差分 SDID
- Xu (2017) 的 Generalized Synthetic Control Method
  - 权重加起来可以不 $= 1$
  - 个体固定效应和随时间变化的系数交互

本节课内容

SDID：解决了 SCM 和 DID 的痛点

合成双重差分 (SDID) 方法由 Arkhangelsky et al. (2021) 发明，继承了 SCM 和 DID 的优点，部分解决了两者的痛点。

- 解决了 DID 平行趋势假设不满足的问题（SCM 也解决）
- 处理组可以多于一个单位（相对于 SCM）
- 不怕处理组的结果比 Donor Pool 高/低太多（相对于 SCM）
- 惩罚了过度拟合 (overfitting) (相对于 SCM)
  - 没有 Convex Hull 的限制
- 权重较为均匀（相对于 SCM）
  - 不杀 Donor Pool（之后降到）
- 也可以多期应用 (staggered)
  - 不过需要额外 R 包支持
  - 不过 SCM 魔改版也可以多期
- 双重稳健性 (Double robustness property)
  - 单位权重 (ω)+**时间权重**(λ)

但是 SDID 也有自己的问题

- 需要强平衡面板 (Balanced Panel)
  - 一点 missing data 不能有
- 加协变量 (covariates) 的机制不完善
  - 和 SCM 的处理方式不太一样
  - Arkhangelsky et al. (2021) 原版给的 covariates 支持很垃圾
  - 网友有一些野生改进（我没试过，也没看人用过）
- 计算 SE 的方式不太行
  - Placebo (大，且慢)
  - Jackknife 快，但需要处理组有 1 个以上单位
  - Bootstrap 超级慢，且需要处理组有 1 个以上单位
  - 弥补办法：转换成 weighted regression

三者的区别可以用回归问题形式体现:

$$(\widehat{\Delta\tau}^{sdid}, \widehat{\mu}, \widehat{\alpha}, \widehat{\beta}) = \arg\min_{\tau,\mu,\alpha,\beta} \{\sum_{i=1}^{N}\sum_{t=1}^{T}(Y_{it} - \mu - \alpha_i - \beta_t - W_{it}\tau)^2 \widehat{\omega}_i^{sdid}\widehat{\lambda}_t^{sdid}\}. \tag{1}$$

$$(\widehat{\Delta\tau}^{did}, \widehat{\mu}, \widehat{\alpha}, \widehat{\beta}) = \arg\min_{\tau,\mu,\alpha,\beta} \{\sum_{i=1}^{N}\sum_{t=1}^{T}(Y_{it} - \mu - \alpha_i - \beta_t - W_{it}\tau)^2\}. \tag{2}$$

$$(\widehat{\Delta\tau}^{sc}, \widehat{\mu}, \widehat{\beta}) = \arg\min_{\tau,\mu,\beta} \{\sum_{i=1}^{N}\sum_{t=1}^{T}(Y_{it} - \mu - \beta_t - W_{it}\tau)^2 \widehat{\omega}_i^{sc}\}. \tag{3}$$

- SDID 和 SC 的式子都有单位权重($\widehat{\omega}_i$)$\Rightarrow$ “合成”
- DID 和 SDID 都有个体固定效应 (α_i)$\Rightarrow$ “平行趋势”和“双重差分”
- 时间权重($\widehat{\lambda}_t$) 是 SDID 独有的 $\Rightarrow$ 双重稳健性

SDID 的权重计算

1. 先计算正则化参数 (regularization parameter ζ)
   - 用来惩罚过度拟合 (over-fitting)⇒ 过于技术性，不细讲
2. 计算单位权重 unit weight($\widehat{\omega}^{sdid}$)
   - 跟 SC 差不多，但是加入了正则化参数 ζ 来惩罚过度拟合和不均匀权重
   - 多计算一个 $\widehat{\omega}_0$ 作为合成趋势和处理趋势之间的平行距离
3. 计算时间权重 time weight($\hat{\lambda}^{sdid}$)

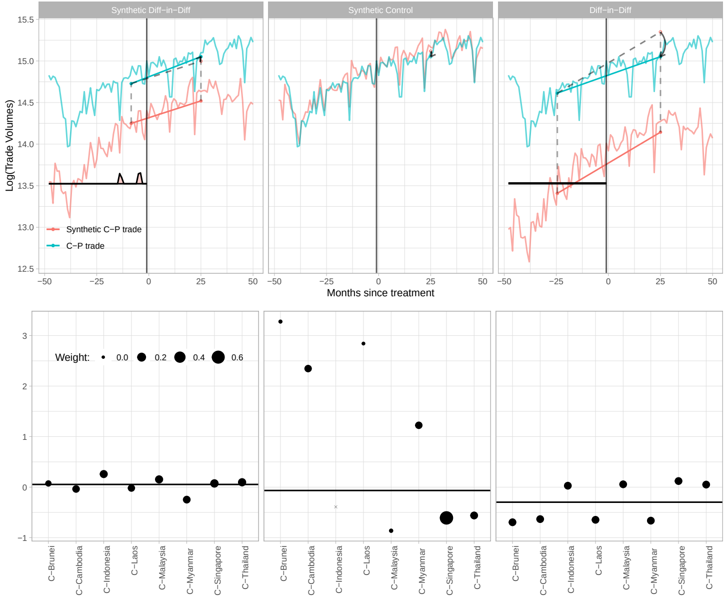
$$\left(\hat{\lambda}_0, \hat{\lambda}^{\text{sdid}}\right) = \underset{\lambda_0 \in \mathbb{R}, \lambda \in \Lambda}{\arg\min}\, \ell_{time}(\lambda_0, \lambda) \quad \text{where}$$

$$\ell_{time}(\lambda_0, \lambda) = \sum_{i=1}^{N_{\text{co}}} \left(\lambda_0 + \sum_{t=1}^{T_{\text{pre}}} \lambda_t Y_{it} - \frac{1}{T_{\text{post}}} \sum_{t=T_{\text{pre}}+1}^{T} Y_{it} \right)^2,$$

$$\Lambda = \left\{ \lambda \in \mathbb{R}_+^T : \sum_{t=1}^{T_{\text{pre}}} \lambda_t = 1,\ \lambda_t = T_{\text{post}}^{-1} \text{ for all } t = T_{\text{pre}} + 1, \ldots, T \right\}.$$

   - 直觉理解：（在 donor pool 中）挑选前后平行的时间段赋予高权重

Huang and Yang (2023)：黄岩岛冲突对中菲贸易的影响



Inference Methods 检验方法

就像我们之前说的，官方给的 se 不大靠谱（特别是在只有一个处理单位的时候，多个还好一些）:

- Placebo (大，且慢)
- Jackknife 快，但需要处理组有 1 个以上单位
- Bootstrap 超级慢，且需要处理组有 1 个以上单位

我们可以通过（未经专业学者正是）转换成 weighted regression 的方法来解决这个问题

- clustered SE （单位够多）
- wild-clustered bootstrap (单位不够多)
- Driscoll-Kraay (DK) (时间够多，解决 cross-sectionally and time-correlated errors)

R 实操

来练习一下吧！

References I

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller (2015). "Comparative Politics and the Synthetic Control Method". en. In: *American Journal of Political Science* 59.2, pp. 495–510. ISSN: 1540-5907. DOI: 10.1111/ajps.12116. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/ajps.12116 (visited on 08/12/2023).

Arkhangelsky, Dmitry et al. (Dec. 2021). "Synthetic Difference-in-Differences". In: *American Economic Review* 111 (12), pp. 4088–4118. ISSN: 0002-8282. DOI: 10.1257/aer.20190159.

Xu, Yiqing (2017). "Generalized Synthetic Control Method: Causal Inference with Interactive Fixed Effects Models". In: *Political Analysis* 25.1, pp. 57–76. DOI: 10.1017/pan.2016.2.