

- ① 因果推断中的内生性
- ② DID 中的协变量
- ③ DID 中的机器学习
- ④ 内生的处理配置
- ⑤ DID-IV 与 IV 型 DID
- ⑥ 参考文献

1 因果推断中的内生性

2 DID 中的协变量

3 DID 中的机器学习

4 内生的处理配置

5 DID-IV 与 IV 型 DID

6 参考文献

因果推断回顾

- 回忆：线性回归方程的斜率系数 $\beta = \frac{\partial Y}{\partial T}$ 测度了处理变量 T 对结果变量 Y 的因果效应
- β 的无偏估计要求“外生性”：

$$\text{cov}(T, u) = 0$$

当随机扰动项 u 中包含遗落变量 W ，外生性就不满足，“内生性”问题就出现了

- 计量经济学从以下几个方面来尝试解决内生性问题：
 - 做实验，在实验中，我们可以控制所有其它因素为常数（即 $u = c$ ），从而使得 $\text{cov}(T, c) = 0$
 - 在 RCT 中，随机配置处理 T ($T = \text{rand}$)，从而使得 $\text{cov}(\text{rand}, c) = 0$
 - 从 u 中把遗落遗漏变量 W “拿”出来，例如，回归调整或者匹配，控制住 W ，比较相同 W 下的处理组与控制组结果差异
 - 准自然实验方法，IV、RD、panel data、SC 和 DID 等

- A set of small navigation icons typically found in Beamer presentations, including symbols for back, forward, search, and other slide controls.

1 因果推断中的内生性

2 DID 中的协变量

3 DID 中的机器学习

4 内生的处理配置

处理配置讨论的重要性

处理配置机制的正式定义

内生处理配置机制引起的偏误及其应对方法

应用案例: Hollingsworth et al.(2024,AER)

实践建议

5 DID-IV 与 IV 型 DID

1 因果推断中的内生性

2 DID 中的协变量

3 DID 中的机器学习

4 内生的处理配置

处理配置讨论的重要性

处理配置机制的正式定义

内生处理配置机制引起的偏误及其应对方法

应用案例: Hollingsworth et al.(2024,AER)

实践建议

5 DID-IV 与 IV 型 DID

Ref 许文立 (2024): 《听 Rubin 和 Imbens 的话——DID 中处理配置机制讨论的实践指导》

- ◀ ◻ ▶ ◀ ◻ ▶ ◀ ≡ ▶ ◀ ≡ ▶ ≡ ↺ 🔍 ↻

- 政策评估非常具有挑战性
- 鲁宾因果模型 (Rubin causal model, RCM) 是思考政策因果效应最常用的工具之一

- 政策评估非常具有挑战性
- 鲁宾因果模型 (Rubin causal model, RCM) 是思考政策因果效应最常用的工具之一
 - 1 潜在结果 (potential outcome)

- 政策评估非常具有挑战性
- 鲁宾因果模型 (Rubin causal model, RCM) 是思考政策因果效应最常用的工具之一

1 潜在结果 (potential outcome)

2 处理配置机制 (assignment mechanism)

This perspective was called the 'Rubin causal model (RCM)' by Paul Holland [2] for a sequence of papers I wrote in the 1970s [3–9]. The RCM can be seen as having two essential parts, together called the 'potential outcomes with assignment mechanism' perspective (Rubin [10, p. 476]),

- 政策评估非常具有挑战性
- 鲁宾因果模型 (Rubin causal model, RCM) 是思考政策因果效应最常用的工具之一

- 1 潜在结果 (potential outcome)

- 2 处理配置机制 (assignment mechanism)

This perspective was called the 'Rubin causal model (RCM)' by Paul Holland [2] for a sequence of papers I wrote in the 1970s [3–9]. The RCM can be seen as having two essential parts, together called the 'potential outcomes with assignment mechanism' perspective (Rubin [10, p. 476]),

- 前者将因果效应定义为同一个体不同处理下的潜在结果的差异

- 政策评估非常具有挑战性
- 鲁宾因果模型 (Rubin causal model, RCM) 是思考政策因果效应最常用的工具之一
 - 1 潜在结果 (potential outcome)
 - 2 处理配置机制 (assignment mechanism)

This perspective was called the 'Rubin causal model (RCM)' by Paul Holland [2] for a sequence of papers I wrote in the 1970s [3–9]. The RCM can be seen as having two essential parts, together called the 'potential outcomes with assignment mechanism' perspective (Rubin [10, p. 476]),
- 前者将因果效应定义为同一个体不同处理下的潜在结果的差异
- 后者关注的是什么因素决定了处理配置, 或者因果效应的研究设计

- 政策评估非常具有挑战性
- 鲁宾因果模型 (Rubin causal model, RCM) 是思考政策因果效应最常用的工具之一
 - 1 潜在结果 (potential outcome)
 - 2 处理配置机制 (assignment mechanism)

This perspective was called the ‘Rubin causal model (RCM)’ by Paul Holland [2] for a sequence of papers I wrote in the 1970s [3–9]. The RCM can be seen as having two essential parts, together called the ‘potential outcomes with assignment mechanism’ perspective (Rubin [10, p. 476]),
- 前者将因果效应定义为同一个体不同处理下的潜在结果的差异
- 后者关注的是什么因素决定了处理配置，或者因果效应的研究设计
- Rubin(2007,p21) believe that “the careful implementation of this step is absolutely essential for drawing objective inferences for causal effects in practice, whether in randomized experiments or observational studies, yet is often basically ignored in observational studies relative to methods of analysis for causal effects”

- 对处理配置机制的讨论在 1975 年之前就出现了
- Roy (1951) 认为，每个个体都会选择带来最优结果的处理

- ◀ ◻ ▶ ◀ ◻ ▶ ◀ ≡ ▶ ◀ ≡ ▶ ≡ ▶ ↺ 🔍 ↻

- 对处理配置机制的讨论在 1975 年之前就出现了
- Roy (1951) 认为, 每个个体都会选择带来最优结果的处理
- Haavelmo (1944) 的供给与需求行为理论
- 这些早期的讨论并没有形成“处理配置机制”的正式概念和框架, 也没有基于处理配置机制来进行因果效应推断
- 经济学界开始用回归模型从协变量和处理变量中预测可观测结果, 并且对一些回归系数和误差项施加一些限制 (属于处理配置机制的范畴)
- 直到 1975 年, Rubin (1975) 给出了处理配置机制的一般化定义, 即处理配置可能依赖于潜在结果

- The assignment mechanism $k(W|X, Y, \pi)$ can be under of experimenter's control since he can assign treatments to experimental units.

- The assignment mechanism $k(W|X, Y, \pi)$ can be under of experimenter's control since he can assign treatments to experimental units.

- 基础：Fisher (1925) 和 Neyman (1923) 等提出的随机配置机制假设

- 20 世纪 80 年代后，越来越多的经济学文献开始对处理配置机制或者选择机制施加明确的限制，以此来发展并比较处理效应的不同估计方法
- Ashenfelter 和 Card(1985) 假设基于固定效应选择进入处理
- 其它的文献则基于未处理潜在结果、基于处理效应、基于其它经济模型等等选择机制
- Rubin (2007) 指出，“在实践中，处理配置机制的分析对于因果效应推断绝对是必要的步骤”

- 20 世纪 80 年代后，越来越多的经济学文献开始对处理配置机制或者选择机制施加明确的限制，以此来发展并比较处理效应的不同估计方法
- Ashenfelter 和 Card(1985) 假设基于固定效应选择进入处理
- 其它的文献则基于未处理潜在结果、基于处理效应、基于其它经济模型等等选择机制
- Rubin (2007) 指出，“在实践中，处理配置机制的分析对于因果效应推断绝对是必要的步骤”
- Imbens 和 Xu(2024) 更强调，“因果效应分析的第一步是理解处理配置机制”

- ◀ ◻ ▶ ◀ ◻ ▶ ◀ ≡ ▶ ◀ ≡ ▶ ≡ ▶ ↺ 🔍 ↻

- 20 世纪 80 年代后，越来越多的经济学文献开始对处理配置机制或者选择机制施加明确的限制，以此来发展并比较处理效应的不同估计方法
- Ashenfelter 和 Card(1985) 假设基于固定效应选择进入处理
- 其它的文献则基于未处理潜在结果、基于处理效应、基于其它经济模型等等选择机制
- Rubin (2007) 指出，“在实践中，处理配置机制的分析对于因果效应推断绝对是必要的步骤”
- Imbens 和 Xu(2024) 更强调，“因果效应分析的第一步是理解处理配置机制”
- 但在经验研究中，研究者通常忽略了处理配置机制的分析 (Robin, 2007)
- 研究者会用“准随机”来指代处理配置，以此来支持所选研究设计 (Ghanem et al., 2024)

- ◀ ◻ ▶ ◀ ◻ ▶ ◀ ≡ ▶ ◀ ≡ ▶ ≡ ▶ ↺ 🔍 ↻

- ◀ ◻ ▶ ◀ ◻ ▶ ◀ ≡ ▶ ◀ ≡ ▶ ≡ ▶ ↺ 🔍 ↻

- A set of small navigation icons typically found in Beamer presentations, including symbols for back, forward, search, and other slide controls.

1 因果推断中的内生性

2 DID 中的协变量

3 DID 中的机器学习

4 内生的处理配置

处理配置讨论的重要性

处理配置机制的正式定义

内生处理配置机制引起的偏误及其应对方法

应用案例: Hollingsworth et al.(2024,AER)

实践建议

5 DID-IV 与 IV 型 DID

- 基于 Rubin(1975,2007)、Ghanem et al.(2024) 的潜在因果框架
- 考察 2×2 DID 设定, D_{it} 为处理变量, Y_{it} 为结果变量, 个体 i 基于一定的原因做出是否进入处理的决策。如果个体 i 在第 1 期未处理, 在第 2 期选择处理, 它就属于处理组 ($G_i = 1$), 如果两期均未处理, 就属于控制组 ($G_i = 0$)

- 基于 Rubin(1975,2007)、Ghanem et al.(2024) 的潜在因果框架
- 考察 2×2 DID 设定, D_{it} 为处理变量, Y_{it} 为结果变量, 个体 i 基于一定的原因做出是否进入处理的决策。如果个体 i 在第 1 期未处理, 在第 2 期选择处理, 它就属于处理组 ($G_i = 1$), 如果两期均未处理, 就属于控制组 ($G_i = 0$)
- 个体 i 处于处理状态或者未处理状态所对应的潜在结果分别为 $Y_{it}(1)$ 和 $Y_{it}(0)$, $t = (1, 2)$

- 基于 Rubin(1975,2007)、Ghanem et al.(2024) 的潜在因果框架
- 考察 2×2 DID 设定, D_{it} 为处理变量, Y_{it} 为结果变量, 个体 i 基于一定的原因做出是否进入处理的决策。如果个体 i 在第 1 期未处理, 在第 2 期选择处理, 它就属于处理组 ($G_i = 1$), 如果两期均未处理, 就属于控制组 ($G_i = 0$)
- 个体 i 处于处理状态或者未处理状态所对应的潜在结果分别为 $Y_{it}(1)$ 和 $Y_{it}(0)$, $t = (1, 2)$
- 个体 i 的处理效应 (δ_i) 可以定义为:

$$\delta_i = Y_{i2}(1) - Y_{i2}(0)$$

- 基于 Rubin(1975,2007)、Ghanem et al.(2024) 的潜在因果框架
- 考察 2×2 DID 设定, D_{it} 为处理变量, Y_{it} 为结果变量, 个体 i 基于一定的原因做出是否进入处理的决策。如果个体 i 在第 1 期未处理, 在第 2 期选择处理, 它就属于处理组 ($G_i = 1$), 如果两期均未处理, 就属于控制组 ($G_i = 0$)
- 个体 i 处于处理状态或者未处理状态所对应的潜在结果分别为 $Y_{it}(1)$ 和 $Y_{it}(0)$, $t = (1, 2)$
- 个体 i 的处理效应 (δ_i) 可以定义为:

$$\delta_i = Y_{i2}(1) - Y_{i2}(0)$$

- 对于总体来说, 平均处理效应 (ATE) 为:

$$ATE = E[\delta_i] = E[Y(1)] - E[Y(0)]$$

上式可以明确地看出, ATE 并不依赖于哪个个体选择进入处理, 哪个个体没有选择进入处理

- 但是 ATT ($ATT = E[Y_{i2}(1) - Y_{i2}(0) | G_i = 1]$) 则是通过 DID 估计量:

$$ATT = E[Y_{i2} - Y_{i1} | G_i = 1] - E[Y_{i2} - Y_{i1} | G_i = 0]$$

- 但是 ATT ($ATT = E[Y_{i2}(1) - Y_{i2}(0) | G_i = 1]$) 则是通过 DID 估计量:

$$ATT = E[Y_{i2} - Y_{i1} | G_i = 1] - E[Y_{i2} - Y_{i1} | G_i = 0]$$

- 平行趋势假设成立:

$$E[Y_{i2}(0) - Y_{i1}(0) | G_i = 1] = E[Y_{i2}(0) - Y_{i1}(0) | G_i = 0]$$

- 但是 ATT ($ATT = E[Y_{i2}(1) - Y_{i2}(0) | G_i = 1]$) 则是通过 DID 估计量:

$$ATT = E[Y_{i2} - Y_{i1} | G_i = 1] - E[Y_{i2} - Y_{i1} | G_i = 0]$$

- 平行趋势假设成立:

$$E[Y_{i2}(0) - Y_{i1}(0) | G_i = 1] = E[Y_{i2}(0) - Y_{i1}(0) | G_i = 0]$$

- 为了识别 ATT , 研究者通常会隐含着对未处理的潜在结果施加一些限制, 例如,

$$Y_{it}(0) = f(\alpha_i, \alpha_t, \epsilon_{it}), i = 1, \dots, n; t = 1, 2$$

Ashenfelter 和 Card(1985) 就是基于固定效应的处理配置机制, 即 α_i

- 定义一个一般化的处理配置机制：个体选择进入处理是基于 $(\alpha_i, \epsilon_{i1}, \epsilon_{i2})$ 和其它一些时间不变和时变因素 $(\nu_i, \mu_{i1}, \mu_{i2})$ ：

$$G_i = g(\alpha_i, \epsilon_{i1}, \epsilon_{i2}, \nu_i, \mu_{i1}, \mu_{i2})$$

- 定义一个一般化的处理配置机制：个体选择进入处理是基于 $(\alpha_i, \epsilon_{i1}, \epsilon_{i2})$ 和其它一些时间不变和时变因素 $(\nu_i, \mu_{i1}, \mu_{i2})$ ：

$$G_i = g(\alpha_i, \epsilon_{i1}, \epsilon_{i2}, \nu_i, \mu_{i1}, \mu_{i2})$$

- 上述选择机制包括许多不同类型的选择：随机选择、基于固定效应的选择、基于未处理潜在结果的选择、基于处理效应的选择、其它经济模型的选择（例如，Heckman and Robb, 1985; Chabe-Ferret, 2015; Marx et al., 2023）

- 定义一个一般化的处理配置机制：个体选择进入处理是基于 $(\alpha_i, \epsilon_{i1}, \epsilon_{i2})$ 和其它一些时间不变和时变因素 $(\nu_i, \mu_{i1}, \mu_{i2})$ ：

$$G_i = g(\alpha_i, \epsilon_{i1}, \epsilon_{i2}, \nu_i, \mu_{i1}, \mu_{i2})$$

- 上述选择机制包括许多不同类型的选择：随机选择、基于固定效应的选择、基于未处理潜在结果的选择、基于处理效应的选择、其它经济模型的选择（例如，Heckman and Robb, 1985; Chabe-Ferret, 2015; Marx et al., 2023）
- 这就意味着， G_i 是实际处理配置机制。注意，有六个元素：两个元素只有下标 i (α_i, ν_i)，另外四个元素与两个时期（1 期和 2 期）、个体差异有关

1 α_i : 时间不变的特征, 例如, 国有企业带头低碳转型

- 1 α_j : 时间不变的特征, 例如, 国有企业带头低碳转型
- 2 ϵ_{i1} : 处理前的时变冲击, 例如, 只有那些 $\epsilon_{i1} > 0$ 的企业才会选择进入处理

- 1 α_i : 时间不变的特征, 例如, 国有企业带头低碳转型
- 2 ϵ_{i1} : 处理前的时变冲击, 例如, 只有那些 $\epsilon_{i1} > 0$ 的企业才会选择进入处理
- 3 ϵ_{i2} : 处理后的时变冲击

- 1 α_i : 时间不变的特征, 例如, 国有企业带头低碳转型
- 2 ϵ_{i1} : 处理前的时变冲击, 例如, 只有那些 $\epsilon_{i1} > 0$ 的企业才会选择进入处理
- 3 ϵ_{i2} : 处理后的时变冲击
- 4 ν_i : 时间不变的约束, 后三项可归类于其它因素

- 1 α_i : 时间不变的特征, 例如, 国有企业带头低碳转型
- 2 ϵ_{i1} : 处理前的时变冲击, 例如, 只有那些 $\epsilon_{i1} > 0$ 的企业才会选择进入处理
- 3 ϵ_{i2} : 处理后的时变冲击
- 4 ν_i : 时间不变的约束, 后三项可归类于其它因素
- 5 μ_{i1} : 处理前时变约束, 它与 ϵ_{i1} 不同, 因为它影响的是企业参与碳交易市场的约束条件, 而不是直接影响企业的碳排放/结果

- 1 α_i : 时间不变的特征, 例如, 国有企业带头低碳转型
- 2 ϵ_{i1} : 处理前的时变冲击, 例如, 只有那些 $\epsilon_{i1} > 0$ 的企业才会选择进入处理
- 3 ϵ_{i2} : 处理后的时变冲击
- 4 ν_i : 时间不变的约束, 后三项可归类于其它因素
- 5 μ_{i1} : 处理前时变约束, 它与 ϵ_{i1} 不同, 因为它影响的是企业参与碳交易市场的约束条件, 而不是直接影响企业的碳排放/结果
- 6 μ_{i2} : 处理后时变约束

处理配置机制的类型

1 基于处理前结果的处理

$$G_i = 1\{E[Y_{i1}(0)|\omega_i] \geq E[C_i|\omega_i]\}$$

处理配置机制的类型

1 基于处理前结果的处理

$$G_i = 1\{E[Y_{i1}(0)|\omega_i] \geq E[C_i|\omega_i]\}$$

2 基于处理前结果变化的处理

$$G_i = 1\{E[Y_{i1}(0)|\omega_i] \geq E[C_i|\omega_i]\}$$

这种处理配置机制是著名的“阿申菲尔特沉降 (The Ashenfelter dip)” (Ashenfelter, 1978)。这种现象在经济社会非常常见，例如，工人参加就业培训可能是因为以前经历了收入下降，球队换教练可能是因为前几场比赛都输球

处理配置机制的类型

1 基于处理前结果的处理

$$G_i = 1\{E[Y_{i1}(0)|\omega_i] \geq E[C_i|\omega_i]\}$$

2 基于处理前结果变化的处理

$$G_i = 1\{E[Y_{i1}(0)|\omega_i] \geq E[C_i|\omega_i]\}$$

这种处理配置机制是著名的“阿申菲尔特沉降 (The Ashenfelter dip)” (Ashenfelter, 1978)。这种现象在经济社会非常常见，例如，工人参加就业培训可能是因为以前经历了收入下降，球队换教练可能是因为前几场比赛都输球

3 基于可观测的个体特征

$$G_i = 1\{E[X_{i1}|\omega_i] \geq E[C_i|\omega_i]\}$$

为了与个体固定效应区分，假设处理前个体特征对潜在结果有异质性效应

处理配置机制的类型

4 基于固定效应的选择

$$G_i = 1\{\alpha_i \leq C\}$$

这种选择机制对应于经典的 Ashenfelter and Card (1985) 劳动就业培训的效应

处理配置机制的类型

4 基于固定效应的选择

$$G_i = 1\{\alpha_i \leq C\}$$

这种选择机制对应于经典的 Ashenfelter and Card (1985) 劳动就业培训的效应

5 基于处理效应的选择。如果来自处理的预期碳排放量 $E[Y_{i2}(1) - Y_{i2}(0)|\omega_i]$ 超过了处理的预期成本 $E[C_{i2}|\omega_i]$,

$$G_i = 1\{E[Y_{i2}(1) - Y_{i2}(0)|\omega_i] \geq E[C_{i2}|\omega_i]\}$$

1 因果推断中的内生性

2 DID 中的协变量

3 DID 中的机器学习

4 内生的处理配置

处理配置讨论的重要性

处理配置机制的正式定义

内生处理配置机制引起的偏误及其应对方法

应用案例: Hollingsworth et al.(2024,AER)

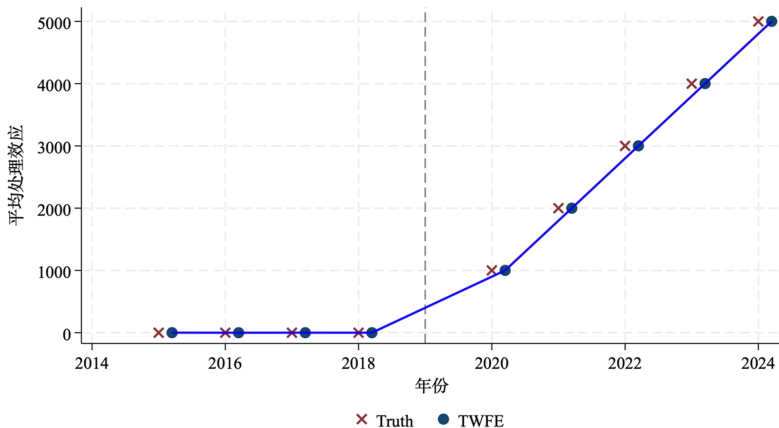
实践建议

5 DID-IV 与 IV 型 DID

1、基于不可观测固定效应的处理配置机制

- 基于固定效应的选择机制可能引起的内生性问题并不会打破 DID 的平行趋势假设 (Ghanem et al., 2024), 因为差分可以消除时间不变的固定效应, 这也是面板数据中的因果效应推断广泛使用固定效应模型最重要的原因之一 (Angrist and Pischke, 2009; Miller et al., 2023)。

1、基于不可观测固定效应的处理配置机制



DGPs基于固定效应的选择机制。1000次蒙特卡洛模拟。

图 1 基于固定效应的蒙特卡洛模拟：真实处理效应和 TWFE

2、基于未处理潜在结果的处理配置机制

Table 1: 处理前后结果变量均值与 DID 估计量				
Panel A: 平行趋势假设				
组群	年份	均值	一阶差分	平行趋势
处理组	2023	2554.232	-1	0
	2024	2553.232		
控制组	2023	854.7286	-1	
	2024	853.7286		
Panel B:DID 估计量				
组群	年份	均值	一阶差分	双重差分
处理组	2023	2554.232	8999	9000
	2024	11553.23		
控制组	2023	854.7286	-1	
	2024	853.7286		
Panel C:TWFE 的 DID 估计量				
$\hat{\mu}$		9000(23.46821)		
注：括号中数值为稳健标准误				

3、基于未处理潜在结果变化的处理配置机制

- 企业选择进入处理是基于处理前潜在结果的变化，这种选择机制也是著名的“阿申菲尔特沉降 (Ashenfelter Dip)”

3、基于未处理潜在结果变化的处理配置机制

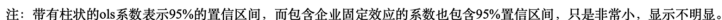
- 企业选择进入处理是基于处理前潜在结果的变化，这种选择机制也是著名的“阿申菲尔特沉降 (Ashenfelter Dip)”
- 它是由著名劳动经济学家 Ashenfelter 在 1978 年研究就业培训项目的效应时发现的，它的本质是个体选择进入处理是基于处理前结果的变化

3、基于未处理潜在结果变化的处理配置机制

- 企业选择进入处理是基于处理前潜在结果的变化，这种选择机制也是著名的“阿申菲尔特沉降 (Ashenfelter Dip)”
- 它是由著名劳动经济学家 Ashenfelter 在 1978 年研究就业培训项目的效应时发现的，它的本质是个体选择进入处理是基于处理前结果的变化
- 实际上，在“阿申菲尔特沉降”选择机制下，研究者应该尤其关注应用研究实践中最常用的处理前趋势检验——TWFE 事件研究的处理前系数

3、基于未处理潜在结果变化的处理配置机制

- 企业选择进入处理是基于处理前潜在结果的变化，这种选择机制也是著名的“阿申菲尔特沉降 (Ashenfelter Dip)”
- 它是由著名劳动经济学家 Ashenfelter 在 1978 年研究就业培训项目的效应时发现的，它的本质是个体选择进入处理是基于处理前结果的变化
- 实际上，在“阿申菲尔特沉降”选择机制下，研究者应该尤其关注应用研究实践中最常用的处理前趋势检验——TWFE 事件研究的处理前系数
- 在“阿申菲尔特沉降”选择机制下，处理后 TWFE 事件研究系数可能是无偏的，反而是处理前系数有偏



◀ ◻ ▶ ◀ ◻ ▶ ◀ ≡ ▶ ◀ ≡ ▶ ≡ ↺ 🔍 ↻

3、基于未处理潜在结果变化的处理配置机制

- 之所以“阿申菲尔特沉降”选择机制不会使得处理后系数产生偏误，而使得经典的处理前趋势检验方法无效，是因为经典的 TWFE 事件研究每个时期的系数都是相对基期的相对效应 (Huntington-Klein, 2021)

3、基于未处理潜在结果变化的处理配置机制

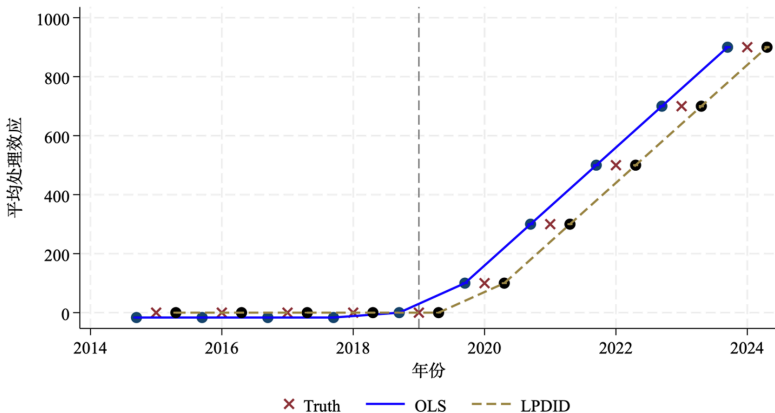
- 之所以“阿申菲尔特沉降”选择机制不会使得处理后系数产生偏误，而使得经典的处理前趋势检验方法无效，是因为经典的 TWFE 事件研究每个时期的系数都是相对基期的相对效应 (Huntington-Klein, 2021)
- 既然是由于“阿申菲尔特沉降”选择机制在处理前结果变量的变化导致经典 TWFE 事件研究的处理前趋势检验无效，那么，研究者可以选择“处理前所有时期的平均”作为基期，或者控制住处理前结果变量的变化趋势

3、基于未处理潜在结果变化的处理配置机制

- 之所以“阿申菲尔特沉降”选择机制不会使得处理后系数产生偏误，而使得经典的处理前趋势检验方法无效，是因为经典的 TWFE 事件研究每个时期的系数都是相对基期的相对效应（Huntington-Klein，2021）
- 既然是由于“阿申菲尔特沉降”选择机制在处理前结果变量的变化导致经典 TWFE 事件研究的处理前趋势检验无效，那么，研究者可以选择“处理前所有时期的平均”作为基期，或者控制住处理前结果变量的变化趋势
- Dube et al. (2023) 提出一种新的局部投影法 DID 估计量 (LPDID)，它的一个显著特征是它允许控制处理前的时变特征，包括结果变量变化

$$Y_{it+h}-Y_{it-1} = \mu_h^{LPDID} \underbrace{\Delta D_{it}}_{treatment} + \underbrace{\sum_{k=-1}^K \gamma_k^h \Delta Y_{it+k}}_{change\ in\ Y} + \underbrace{\alpha_t^h}_{time\ effect} + \epsilon_{it}, h =$$

3、基于未处理潜在结果变化的处理配置机制



注: DGPs使用基于处理前潜在结果变化的平行趋势假设;所有回归均包含企业固定效应; LPDID使用处理前三期作为基期,且控制处理前结果变化;1000次蒙特卡洛模拟。

图 6: “阿申菲尔特沉降”机制的蒙特卡洛模拟: 真实处理效应、TWFE 和 LPDID

4、基于处理前可观测协变量的处理配置机制

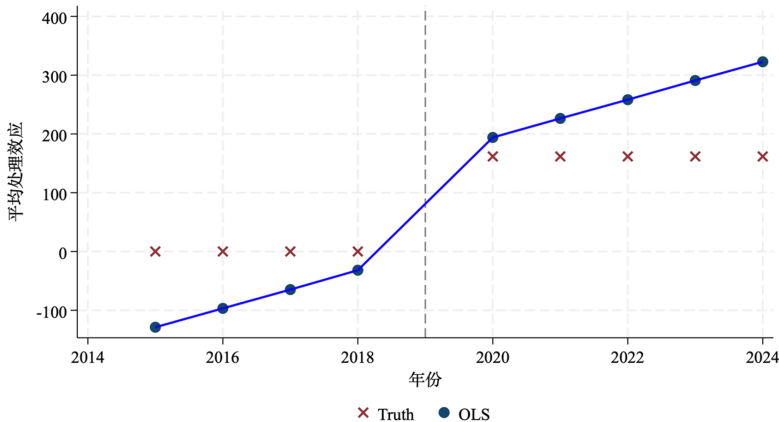
- 当选择机制基于可观测的协变量时，经验研究者通常会在 TWFE 中直接控制协变量

4、基于处理前可观测协变量的处理配置机制

- 当选择机制基于可观测的协变量时，经验研究者通常会在 TWFE 中直接控制协变量
- TWFE 的方程：

$$Y_{it} = \alpha_p + \lambda_t + \alpha_i + \sum_{s \neq 2019} 1[s = t] \times treat_i \times \mu_s + \gamma_1 age_i \times t + \lambda_1 size_i \times t -$$

4、基于处理前可观测协变量的处理配置机制



DGPs使用条件于协变量的平行趋势假设，且TWFE包含所有的协变量。1000蒙特卡洛模拟。

图 7：蒙特卡洛模拟：真实处理效应、带协变量的 TWFE

4、基于处理前可观测协变量的处理配置机制

- 因为在 TWFE 回归中控制协变量主要考虑的是“遗漏变量偏误”——基于可观测协变量选择机制可能带来的遗漏变量偏误，但是却忽略了“条件平行趋势”问题

4、基于处理前可观测协变量的处理配置机制

- 因为在 TWFE 回归中控制协变量主要考虑的是“遗漏变量偏误”——基于可观测协变量选择机制可能带来的遗漏变量偏误，但是却忽略了“条件平行趋势”问题
- “条件平行趋势”主要是指不同的企业在协变量方面具有不同的特征，而具有相似（相同）协变量的企业在没有发生处理时的潜在结果具有相似（相同）的变化趋势 [见第一讲“协变量”]

4、基于处理前可观测协变量的处理配置机制

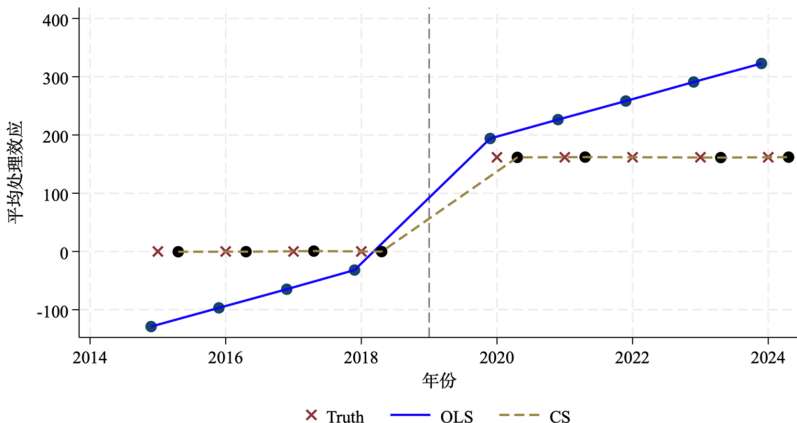
- 因为在 TWFE 回归中控制协变量主要考虑的是“遗漏变量偏误”——基于可观测协变量选择机制可能带来的遗漏变量偏误，但是却忽略了“条件平行趋势”问题
- “条件平行趋势”主要是指不同的企业在协变量方面具有不同的特征，而具有相似（相同）协变量的企业在没有发生处理时的潜在结果具有相似（相同）的变化趋势 [见第一讲“协变量”]
- “条件平行趋势假设”的本质是处理组和控制组的构成在协变量方面存在差异。因此，在基于可观测协变量的选择机制下，可以在 DID 中引入匹配策略

4、基于处理前可观测协变量的处理配置机制

- 因为在 TWFE 回归中控制协变量主要考虑的是“遗漏变量偏误”——基于可观测协变量选择机制可能带来的遗漏变量偏误，但是却忽略了“条件平行趋势”问题
- “条件平行趋势”主要是指不同的企业在协变量方面具有不同的特征，而具有相似（相同）协变量的企业在没有发生处理时的潜在结果具有相似（相同）的变化趋势 [见第一讲“协变量”]
- “条件平行趋势假设”的本质是处理组和控制组的构成在协变量方面存在差异。因此，在基于可观测协变量的选择机制下，可以在 DID 中引入匹配策略
- Callaway and Sant' Anna (2021) 将 Sant' Anna and Zhao (2020) 的双重稳健估计量引入 DID 中

$$ATT(g, t) = E \left[\left(\frac{G_g}{E[G_g]} - \frac{\frac{p_g(X)C}{1-p_g(X)}}{E \left[\frac{p_g(X)C}{1-p_g(X)} \right]} \right) (Y_t - T_{g-1}) \right]$$

4、基于预期处理效应的处理配置机制



DGPs使用条件于协变量的平行趋势假设，且TWFE包含所有的协变量，CSDID使用双重稳健估计量。
1000蒙特卡洛模拟。

图 8：蒙特卡洛模拟：真实处理效应、带协变量的 TWFE 和 CSDID

5、基于处理前可观测协变量的处理配置机制

- 这就是著名的罗伊 (Roy) 自选择机制 (Roy,1951)。

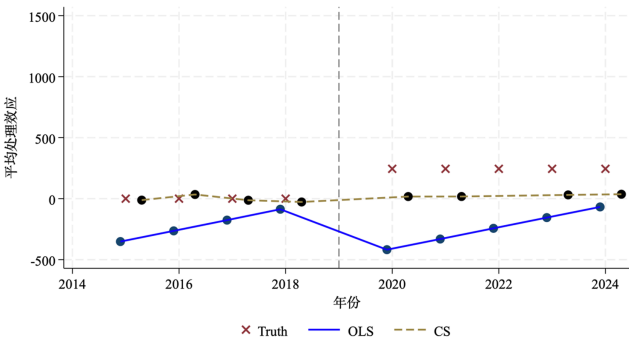
5、基于处理前可观测协变量的处理配置机制

- 这就是著名的罗伊 (Roy) 自选择机制 (Roy, 1951)。
- 罗伊研究了人们根据未来的工资来最优地选择职业

5、基于处理前可观测协变量的处理配置机制

- 这就是著名的罗伊 (Roy) 自选择机制 (Roy,1951)。
- 罗伊研究了人们根据未来的工资来最优地选择职业
- Borjas (1987) 通过研究移民对工资收入的影响正式理论化了罗伊自选择机制。在移民自选择机制中, 移民可以预期到移入国家的工资收入, 进而做出移民决策。从这个意义上来看, 个体根据其处理效应, 选择最有利于自身的处理状态

5、基于处理前可观测协变量的处理配置机制



DGPs使用条件于协变量的平行趋势假设，且TWFE包含所有的协变量，CSDID使用双重稳健估计量。
1000蒙特卡洛模拟。

图 9：蒙特卡洛模拟：真实处理效应、TWFE 和 CSDID

- 处理组和控制组之间的平行趋势假设并不成立
- 因此，要求平行趋势假设的估计量可能均不适用于“基于处理效应的选择机制”下因果效应的估计

1 因果推断中的内生性

2 DID 中的协变量

3 DID 中的机器学习

4 内生的处理配置

处理配置讨论的重要性

处理配置机制的正式定义

内生处理配置机制引起的偏误及其应对方法

应用案例: Hollingsworth et al.(2024,AER)

实践建议

5 DID-IV 与 IV 型 DID

- We further demonstrate that our results are not driven by the two primary threats to identification, differential pre-trends and selection into treatment

1 因果推断中的内生性

2 DID 中的协变量

3 DID 中的机器学习

4 内生的处理配置

处理配置讨论的重要性

处理配置机制的正式定义

内生处理配置机制引起的偏误及其应对方法

应用案例: Hollingsworth et al.(2024,AER)

实践建议

5 DID-IV 与 IV 型 DID

- ◀ ◻ ▶ ◀ ◻ ▶ ◀ ≡ ▶ ◀ ≡ ▶ ≡ ▶ ↺ 🔍 ↻

- A set of small navigation icons typically found in Beamer presentations, including symbols for back, forward, search, and other slide controls.

- 在研究设计阶段，研究者应该仔细地收集、整理关于选择进入处理的经济社会制度等背景信息，或者根据理论知识来充分理解个体选择进入处理的动机与过程（在制度背景中阐述）
- Pedro Sant' Anna(2024) 建议研究者要尽可能弄清楚“谁决定处理”“决策者知道关于处理的什么信息”“处理类型是什么”
- 并没有讨论选择机制的所有类型，例如，序列相关的处理，即一个个体收到处理，那么，下一期他也可能收到处理。这可能需要一个有别于传统一次性处理效应的“处理效应乘数”概念和新的估计方法（Jorda, 2023）

- 在研究设计阶段，研究者应该仔细地收集、整理关于选择进入处理的经济社会制度等背景信息，或者根据理论知识来充分理解个体选择进入处理的动机与过程（在制度背景中阐述）
- Pedro Sant' Anna(2024) 建议研究者要尽可能弄清楚“谁决定处理”“决策者知道关于处理的什么信息”“处理类型是什么”
- 并没有讨论选择机制的所有类型，例如，序列相关的处理，即一个个体收到处理，那么，下一期他也可能收到处理。这可能需要一个有别于传统一次性处理效应的“处理效应乘数”概念和新的估计方法（Jorda, 2023）
- 虽然针对每种选择机制，上文给出了一种稳健估计量，但这并非意味着只有一种稳健估计量可以纠正选择机制可能带来的偏误

- ◀ ◻ ▶ ◀ ◻ ▶ ◀ ≡ ▶ ◀ ≡ ▶ ≡

- 1 因果推断中的内生性
- 2 DID 中的协变量
- 3 DID 中的机器学习
- 4 内生的处理配置
- 5 DID-IV 与 IV 型 DID
- 6 参考文献

45 / 46

◀ ◻ ▶ ◀ ◻ ▶ ◀ ≡ ▶ ◀ ≡ ▶ ≡ 🔍 ↺ ↻