

DIDの計量経済手法の近年の展開 ～ 問題と対策 ～

小西祥文

慶應義塾大学

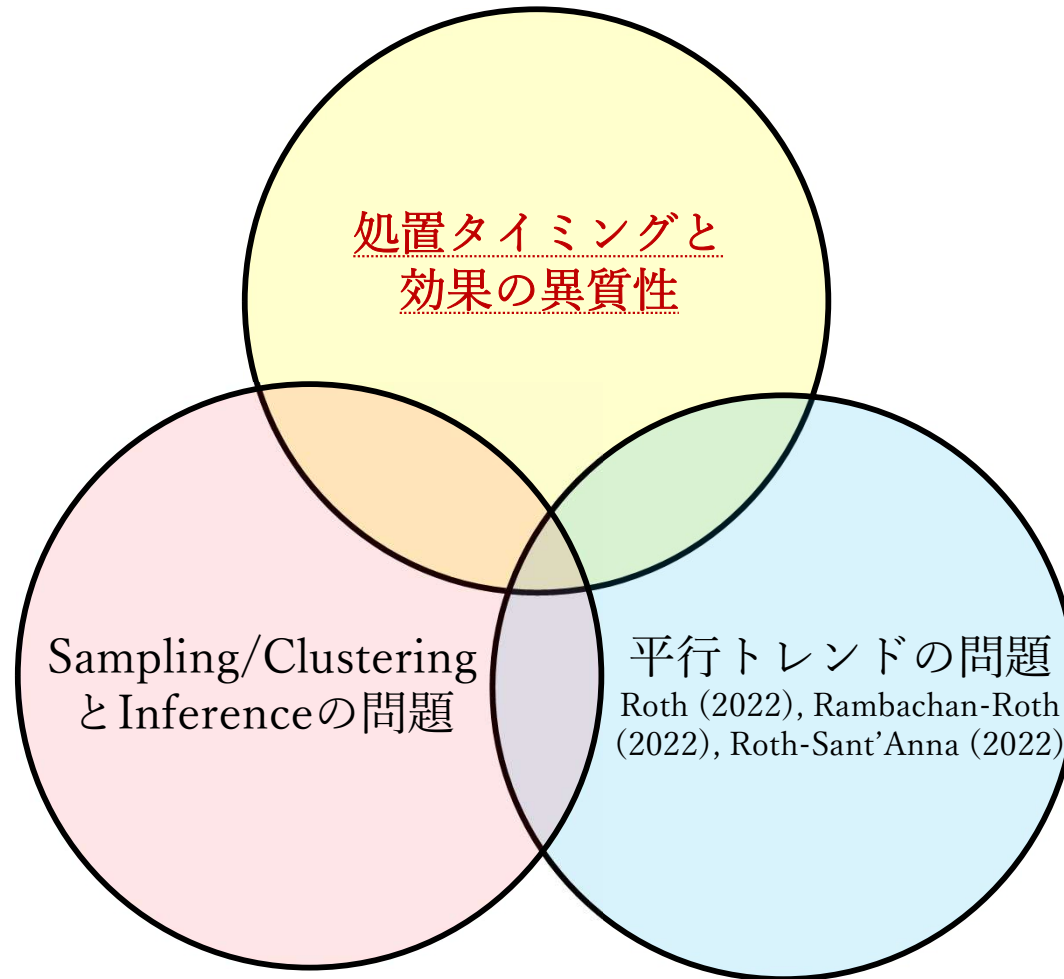
日本経済学会 春季大会

共催：日本学術会議 数量的経済・政策分析分科会

Tutorial Session

2023.5.28

What's Trending in DID?

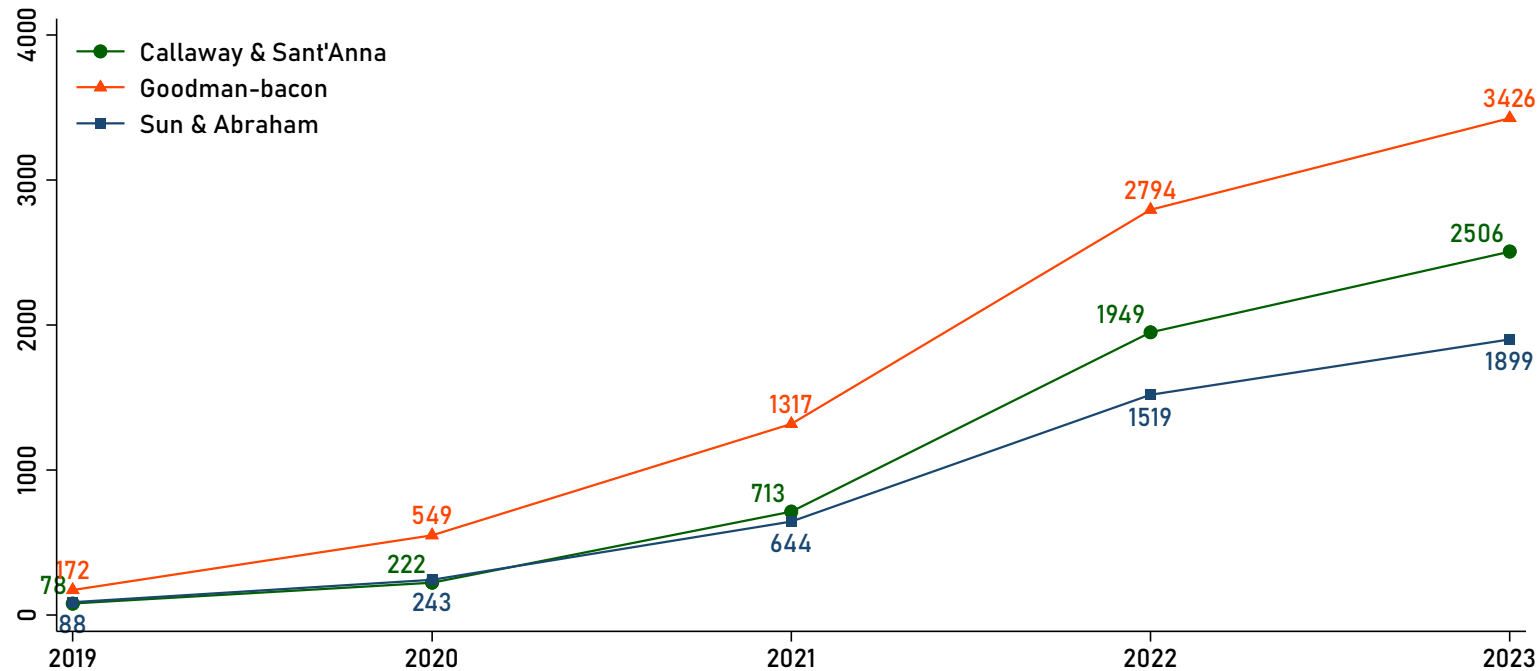


参考: Roth, Sant'Anna, Blinski, and Poe (2023)

“What's Trending in Difference-in-Differences?”, forthcoming at *Journal of Econometrics*

Scope of This Talk

- Dennis J. Aigner Award in 2023 (Best Empirical Theory Papers published in *Journal of Econometrics* 2021-22)
 - Callaway and Sant'Anna (JE, 2021)
 - Goodman-Bacon (JE, 2021)
 - Sun and Abraham (JE, 2021)



What's New? Should I Stay and Listen to Your Talk?



出所: Khoa Vu on Twitter (2021.6.13)

What's New? Should I Stay and Listen to Your Talk?



出所: Pixabay with photo credit to PublicDomainPictures.

Outline

- 問題設定
 - Event-study/Staggered DID design
 - Static or Dynamic TWFE specification
- 何が問題か？なぜ重大なバイアスが生じるのか？
 - “Forbidden Comparisons”の問題
 - Goodman-Bacon/dCDH decomposition
- 対処法は？
 - Fully-saturated TWFE
 - Rolling methods
 - Imputation methods
- Stata (& R) による実践演習
 - 基礎編：シミュレーション
 - 実践編：Cicala (AER, 2022)

問題設定

- 異なる観測体が異なる時点で処置を受けることを外生的変分として利用した Two-way Fixed Effects (TWFE)

$$Y_{it} = \alpha_i + \lambda_t + \delta X_{it} + \epsilon_{it} \quad (1)$$

- ※1. X_{it} : 処置変数(連続変数の場合も含む).
- ※2. 他の制御変数を入れる場合も含む.

- 多くの分野の実証研究で使われている研究デザインの一つ
 - 労働：最低賃金の引上げの雇用への影響
 - 開発経済：CCTの教育・健康効果
 - 国際経済：貿易ショックの移民への影響
 - 経済地理：市場アクセスの都市への影響
 - 環境：電力自由化の“交易の利”への効果

Clarifying Our Setup

1. 問題設定(1)には様々な設定・推定式が含まれる：

- 例1. RCTフィールド実験を複数回異なる時期に行い, poolしたデータでTWFE推定を行う.

$$Y_{it} = \alpha_i + \lambda_t + \delta D_{it} + \epsilon_{it} \quad (1a)$$

- 例2. 異なる地域の異なる時期における最低賃金の引上げが雇用に与えた効果をFirst-difference TWFEで推定する.

$$\Delta y_{it} = \alpha_i + \lambda_t + \delta \Delta x_{it} + \Delta \epsilon_{it} \quad (1b)$$

- 例3. 中国からの輸出が, Exposureの異なる産業の雇用に与えた効果をTWFE(first-differenced or not)で(誘導形)推定.

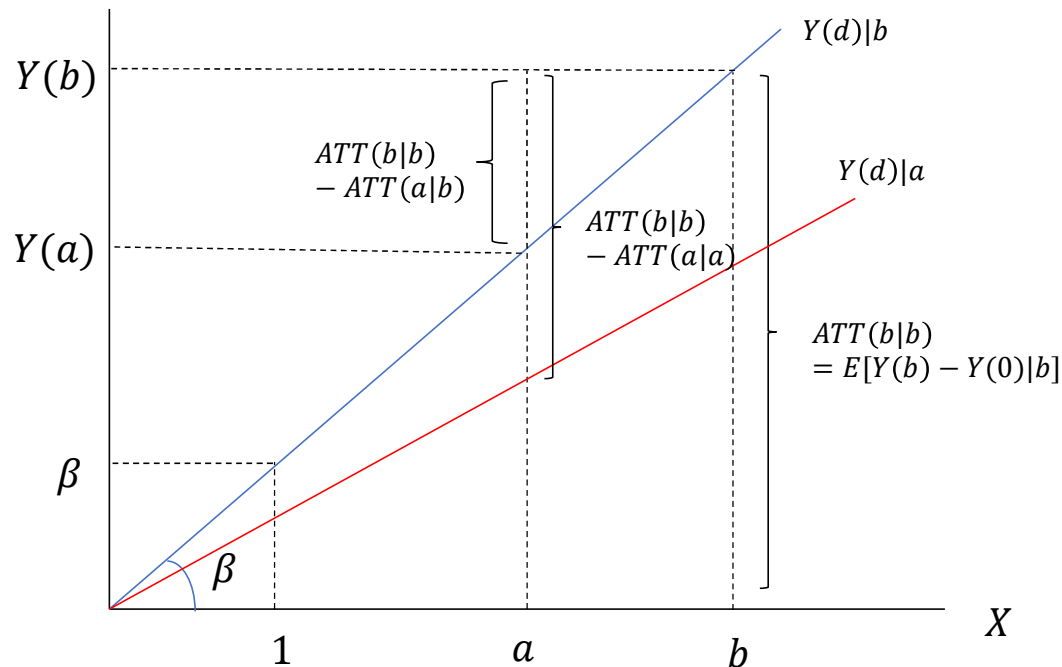
$$\Delta y_{it} = \alpha_i + \lambda_t + \delta \Delta z_{it} + \Delta \epsilon_{it} \quad (1c)$$

Clarifying Our Setup

2. X_{it} が連続変数でも、問題の本質は同じ(より対処が困難).

- 但し, 連続変数の場合, Level Effect (Treatment) と Slope Effect (Causal Response) が異なる可能性(Angrist-Imbens, 1995)

⇒ 研究デザインに応じた適切な特定化が前提
(例: $\delta = E[\Delta Y_{it} | \Delta X_{it} > 0] - E[\Delta Y_{it} | \Delta X_{it} = 0]$)



適切な変数処理・特定化が行われていなければ, そもそも識別できない; 追加的な仮定が必要.
(Callaway *et al.*, 2021)

Common, but “Misleading” Analogy

- “良い”実証研究では, 処置タイミングが異なることを “DID/Event-studyの研究デザイン”として用いる:

$$Y_{it} = \alpha_i + \lambda_t + \delta D_{it} + \epsilon_{it} \quad (1)'$$

$$\Leftrightarrow \delta = \frac{E[Y_{i1} - Y_{i0} | i \in T]}{-E[Y_{i1} - Y_{i0} | i \in C]} \quad (2)$$

- Q. これは(計量経済学的に)正しいAnalogyか?
(TWFE estimatorは δ の正しい推定量か?)

⇒ A. **Yes and No!**

TWFE推定量は, 処置効果 δ が異質性(横断的/時系列的)を持つ場合, (時に重大な)バイアスを持つ.

Preview

Q1. 平行トレンドの仮定の下では, TWFEはATTを正しく推定する?

⇒ × : 処置効果に異質性がある場合, 常にバイアスが生じる. **真の係数とは逆に推定されてしまうケースも (バイアスの重大性).**

Q2. Dynamic TWFEを行えば, こういう問題は生じない?

⇒ × : 処置効果に異質性がある場合, 同様に重大なバイアスが生じる. **Event-study によるPre-trend Checkが信頼できなくなる.**

Q3. 問題が生じているかどうかの診断 (Diagnosis) は可能か?

⇒ Yes : **GB (JE, 2021)/dCDH decomp (AER, 2020)/Jakiela (2021).**

Q4. Diagnosisで問題を発見したら, 対処可能か?

⇒ Yes : **3つのアプローチ. Stata/R有り.**

Q5. 三つのアプローチに違いはある? 実際の実証分析ではどうすべき?

⇒ Yes : **Stata/Rによる実践演習.**

問題の整理

- これまで(1)' (or (1))のATTパラメターを正しく推定するための条件として主に議論されていたのは次の三つ：

$$Y_{it} = \alpha_i + \lambda_t + \delta D_{it} + \epsilon_{it} \quad (1)'$$

Parallel Trend：処置が無かった場合，処置群と対照群は(観測不可能な要因に関して) 同じトレンドを持つ。

No Anticipation：処置を予測して処置群への効果が生じない（≡ SUTVAの第二の仮定）。

No Spillover：処置群と対照群の間に相互作用が存在しない（≡ SUTVAの第一の仮定）。

問題の整理

- 新しい問題：処置効果に異質性が存在し，処置タイミングと相関していたら？（例：規制緩和は効果が高い地域から？）

$$Y_{kt} = \alpha_k + \lambda_t + \delta_k D_{kt} + \epsilon_{kt} \quad (3)$$

$$\begin{aligned} \Rightarrow E[Y_{kt}|D_{kt}] &= \alpha_k + \lambda_t + E[\delta_k|D_{kt} = 1]D_{kt} \\ &\quad + E[\epsilon_{kt}|D_{kt}] + \{\delta_k - E[\delta_k|D_{kt} = 1]\}D_{kt} \end{aligned}$$

$$\Rightarrow E[\epsilon_{kt}|D_{kt}] = 0 \text{ であったとしても，最後の項が残ってしまう.}$$

- ATT(= $E[\delta_k|D_{kt} = 1]$)が正しく推定される条件は？
①効果が同質的 or ②処置が異質性と無相関 (Gardner, 2021).

なぜ重大なバイアスが生じるのか？

- TWFE推定量(Estimator)は 2×2DID推定量の加重和となる(但し, 重みは全て正).

定理 (Theorem 1, Goodman-Bacon, JE, 2021)

式(1)'のTWFE推定量 $\hat{\delta}^{TWFE}$ は全ての2×2DIDの加重和となる：

$$\hat{\delta}^{TWFE} = \sum_{k \neq u} s_{ku} \hat{\delta}_{ku}^{2 \times 2} + \sum_{k \neq u} \sum_{l > k} \left[s_{kl}^k \hat{\delta}_{kl}^{2 \times 2, k} + s_{kl}^l \hat{\delta}_{kl}^{2 \times 2, l} \right] \quad (4)$$

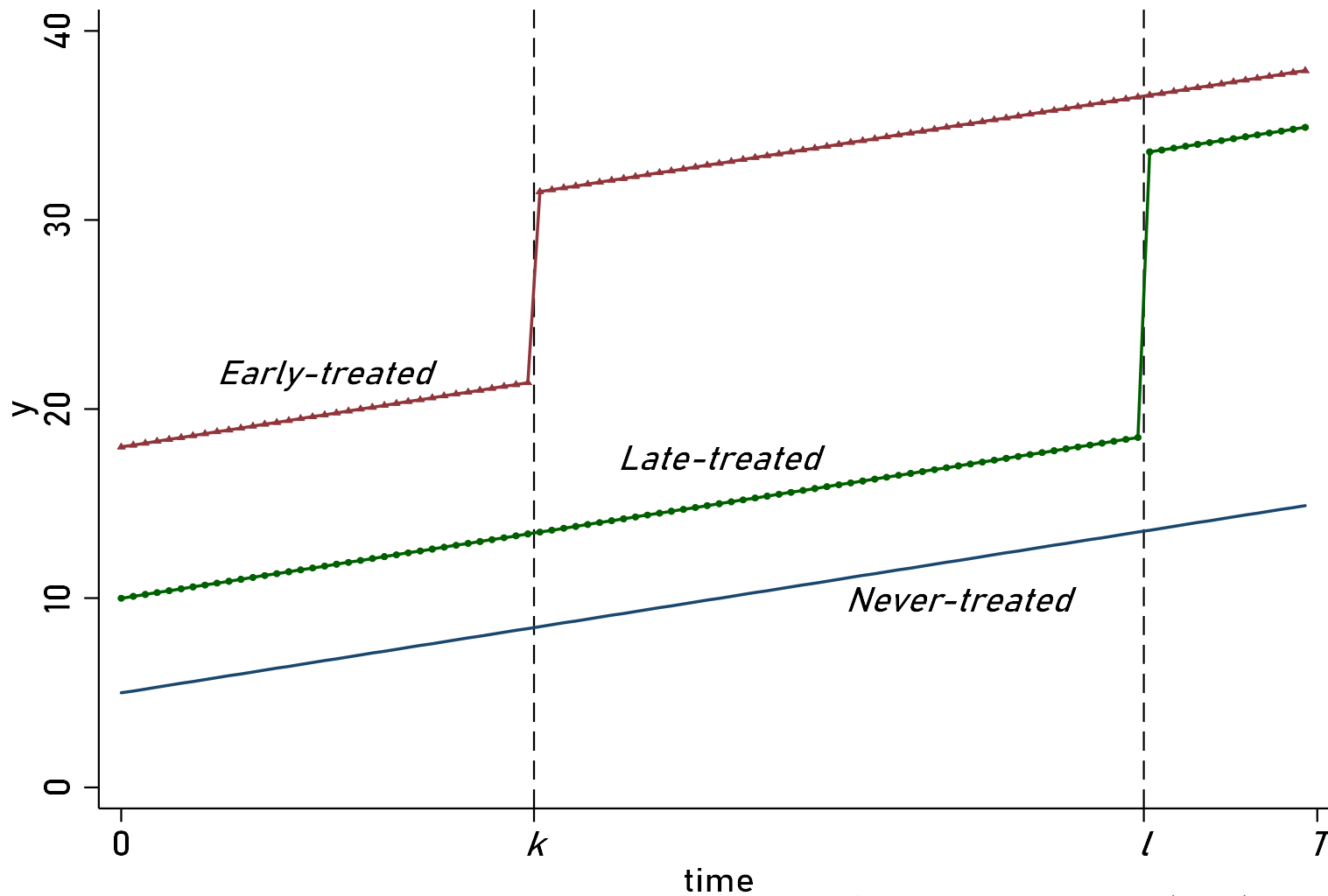
- TWFE被推定量(Estimand)は 真の処置効果の加重和となる(但し, 重みは負になる場合も).

定理 (Theorem 1, de Chaisemartin-D'Haultfoeulle (dCDH), AER, 2020)

式(1)'のTWFE被推定量 δ^{TWFE} は全ての処置効果の加重和となる：

$$\delta^{TWFE} = \sum_{k, t: D=1} \eta_{kt} \delta_{kt} \quad (5)$$

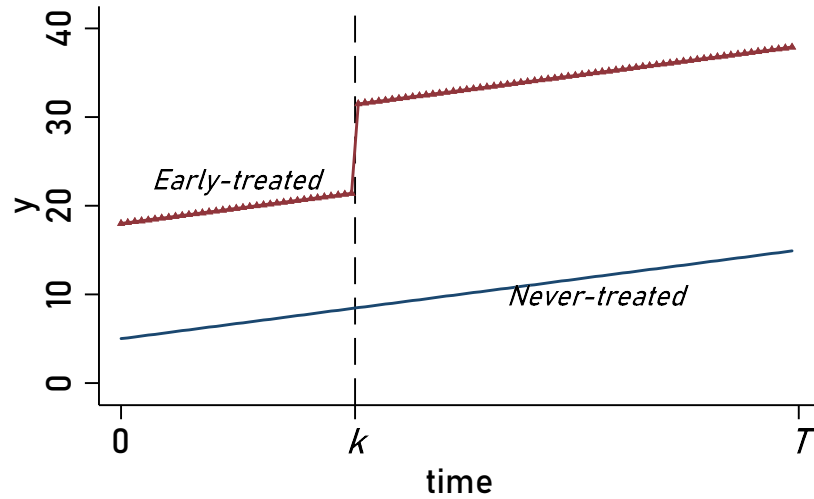
なぜ重大なバイアスが生じるのか？ （“Forbidden Comparisons”の問題）



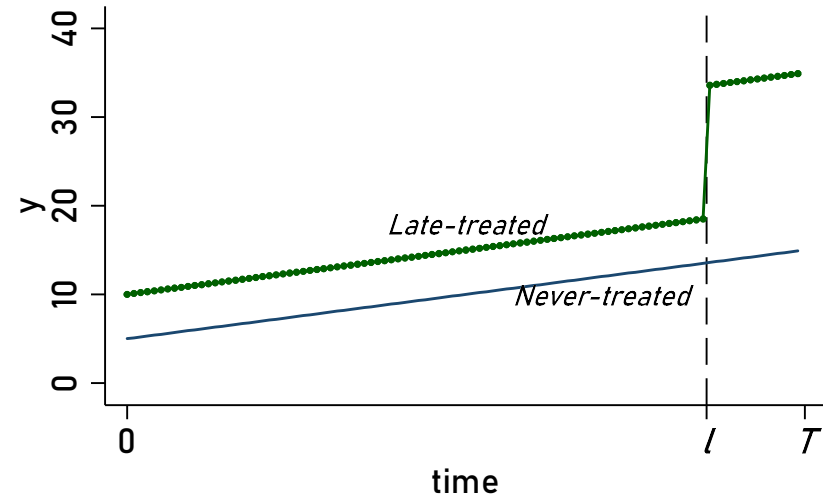
出所: Goodman-Bacon (2021)をベースに作成.

なぜ重大なバイアスが生じるのか？

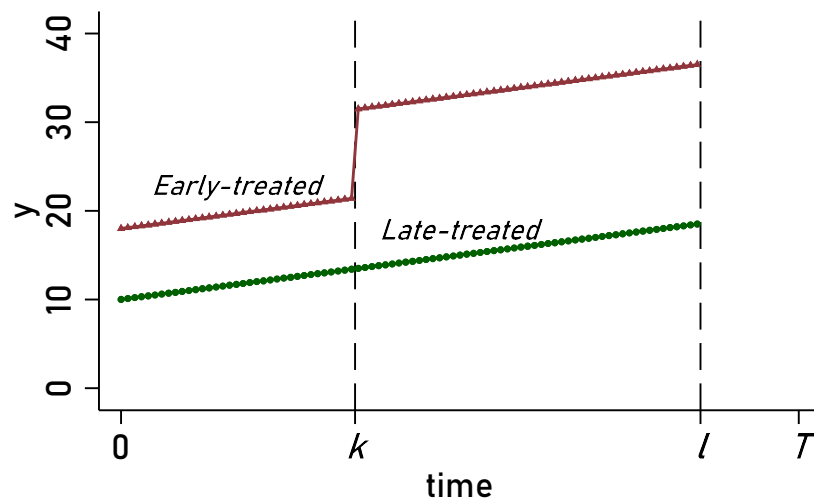
A. Early vs. Untreated Groups



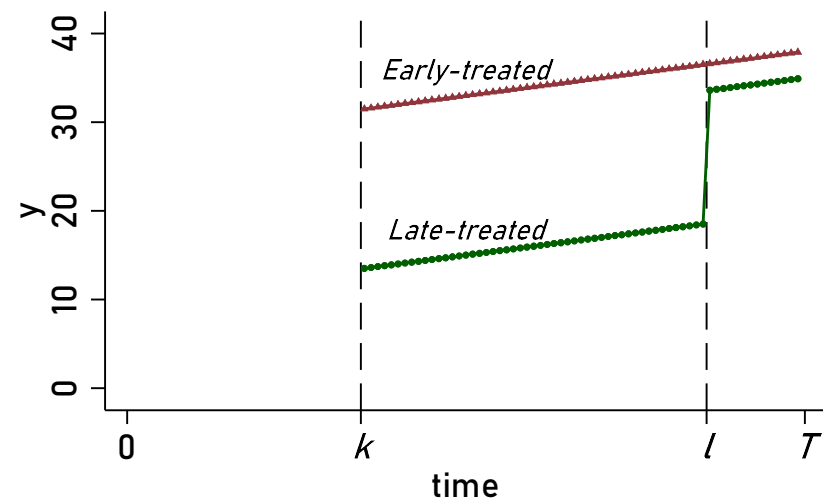
B. Late vs. Untreated Groups



C. Early vs. Late Groups



D. Late vs. Early Groups



計量理論で理解してみよう

Frisch–Waugh–Lovell (FWL) theorem より :

$$\delta^{TWFE} = \frac{E(Y_{it}\tilde{D}_{it})}{E(\tilde{D}_{it}^2)} \quad (6)$$

$$\left(= \frac{E(\tilde{D}_{it}\sum\delta_{kt}\mathbb{I}(k,t))}{E(\tilde{D}_{it}^2)} = \sum_{k,t:D=1} \frac{E(\tilde{D}_{it}\mathbb{I}(k,t))}{E(\tilde{D}_{it}^2)}\delta_{kt} = \sum_{k,t:D=1} \eta_{kt}\delta_{kt} \right)$$

\tilde{D}_{it} : 回帰式 $D_{it} = \hat{\alpha}_i + \hat{\lambda}_t + u_{it}$ の残差

$$\tilde{D}_{it} = D_{it} - \hat{D}_{it} = D_{it} - \bar{D}_i - \bar{D}_t + \bar{D}$$

- 分析期間中の処置確率が高いグループほど $\bar{D}_i \uparrow$, 処置確率の高い期ほど $\bar{D}_t \uparrow$.
- 分析期間の後期には early adopter の weight がマイナスになる.
- 全ての units が処置を受けている期間は, weights は確実にマイナスになる.

どの程度のバイアスが生じるのか？

- Baker *et al.* (2022) の仮想例
 - $T = 36$: 1980~2015
 - 処置タイミング : 1989, 1998, 2007
 - $N = 1000$ firms $\Rightarrow i$ randomly assigned to $\{G_{89}, G_{98}, G_{07}\}$
 - $\alpha_i, \lambda_t, \epsilon_{it} \sim N(0, 0.5)$

- Case 1 : 処置効果は同質的, No Dynamic Treatment Effect

$$\text{DGP: } y_{it} = \alpha_i + \lambda_t + \tau D_{it} + \epsilon_{it}, \tau \sim N(3, 0.2^2)$$

- Case 2 : 処置効果は異質的, Dynamic Treatment Effect

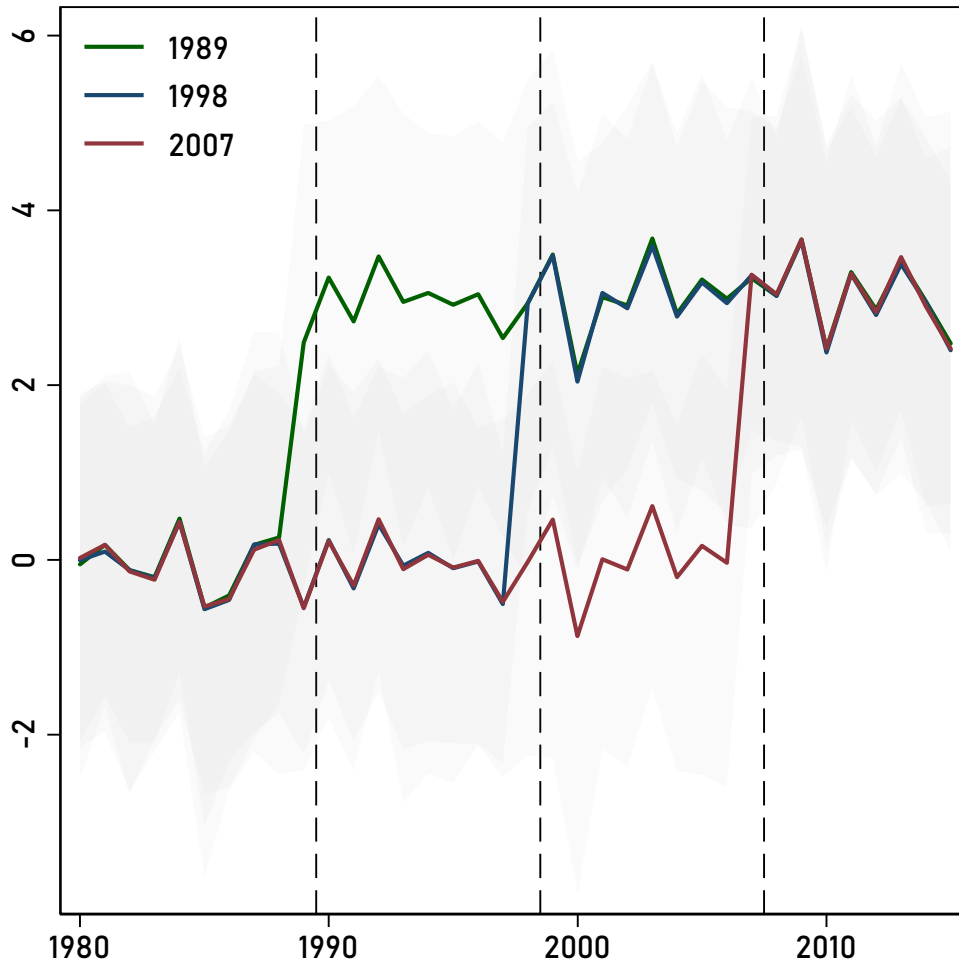
$$\text{DGP: } y_{it} = \alpha_i + \lambda_t + \tau_k D_{it} + \epsilon_{it}$$

$$\tau_k \sim N(\delta_k \times [T - k], 0.2^2 \times [T - k])$$

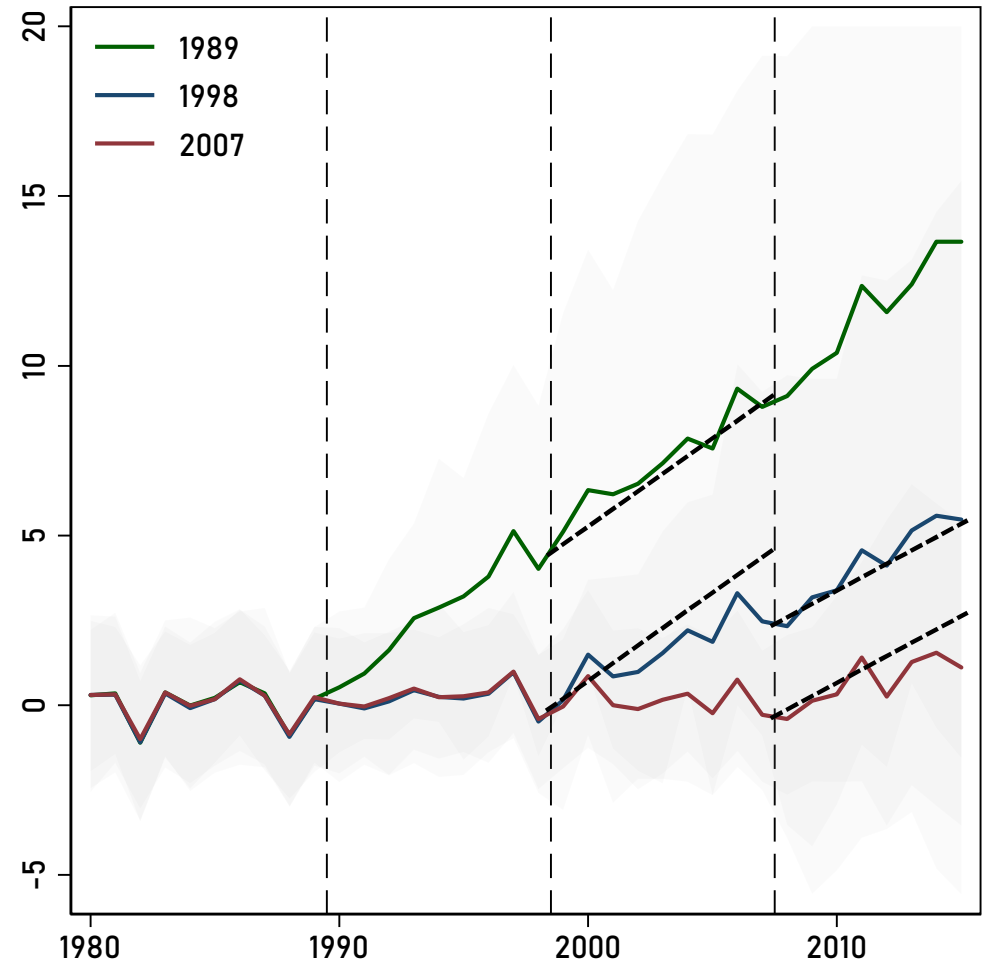
$$\delta_{89} = 0.5, \delta_{98} = 0.3, \delta_{07} = 0.1$$

どの程度のバイアスが生じるのか？

Case 1. Homogeneous/No Dynamic Effect
(Simulation 3 in Baker *et al.*)

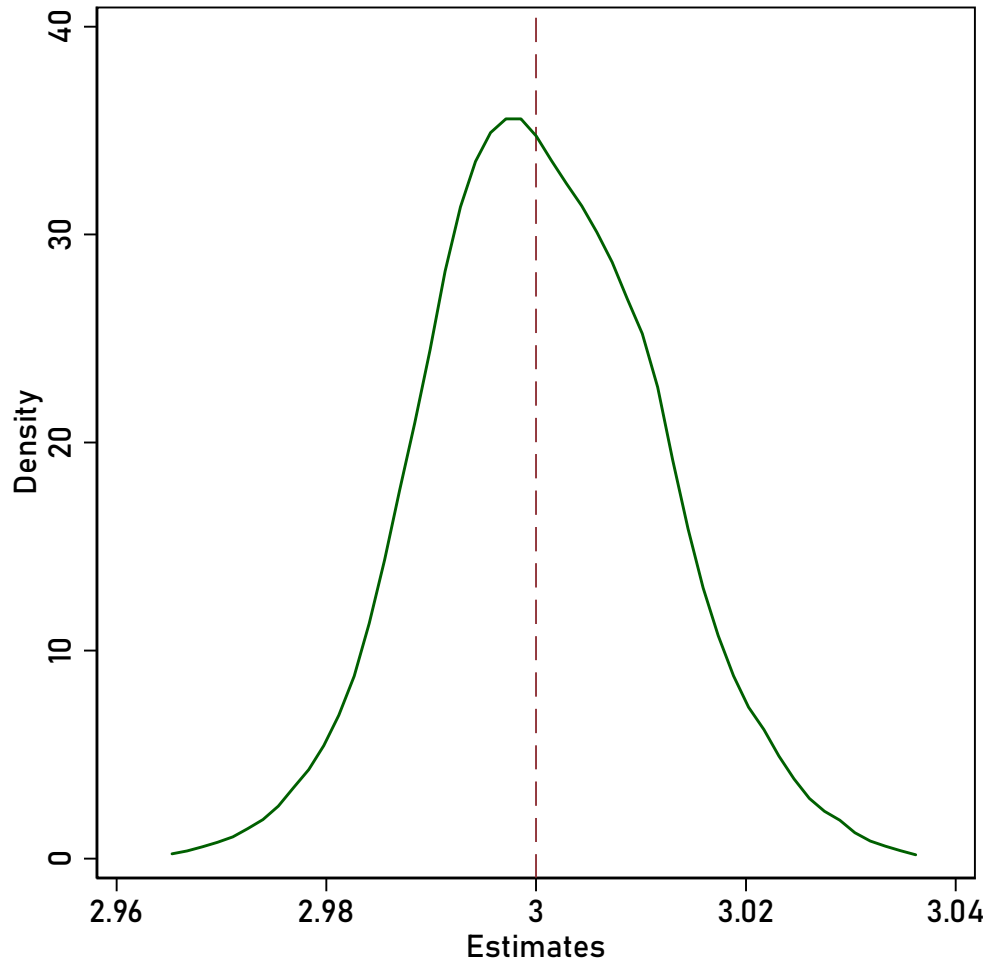


Case 2. Heterogeneous/Dynamic Effect
(Simulation 6 in Baker *et al.*)



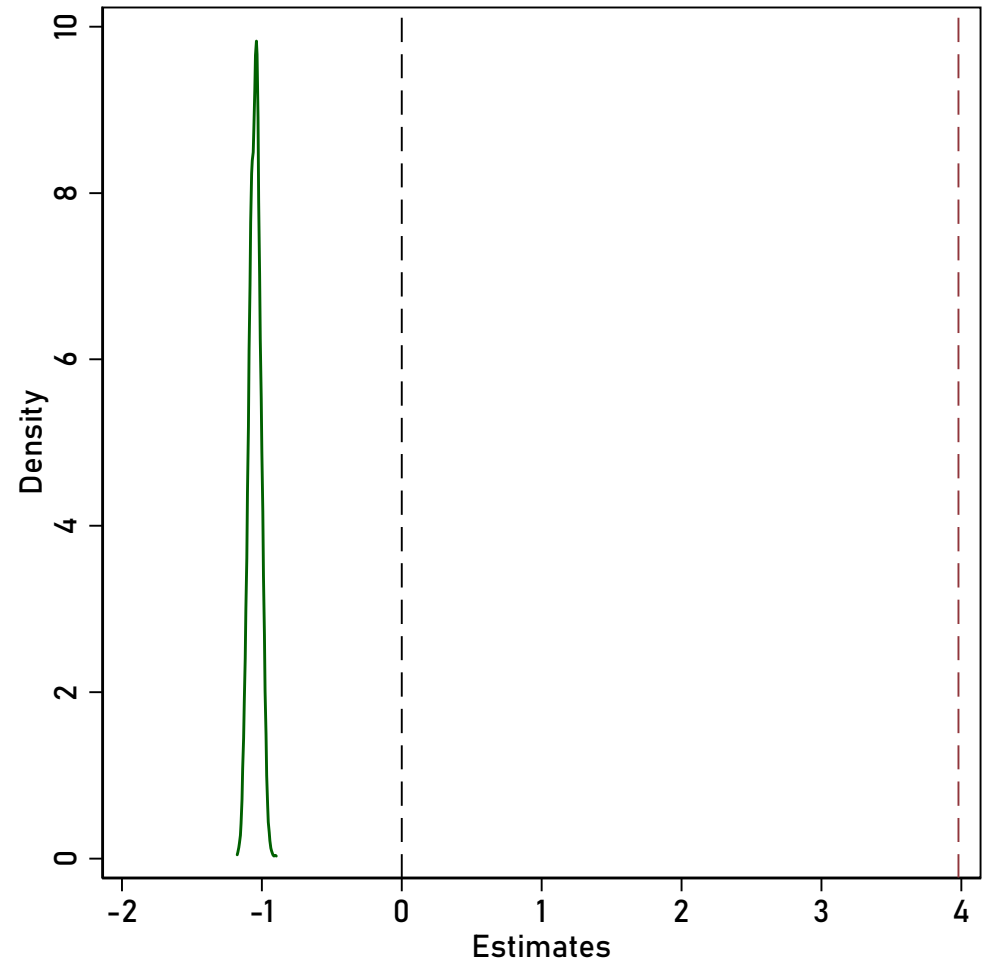
真のATTと逆になるケースも

Case 1. Homogeneous/No Dynamic Effect
(Simulation 3 in Baker *et al.*)



kernel = epanechnikov, bandwidth = 0.0024

Case 2. Heterogeneous/Dynamic Effect
(Simulation 6 in Baker *et al.*)



kernel = epanechnikov, bandwidth = 0.0089

註: Baker *et al.* (2022) Figure 3 (ii)をベースに作成.

Event Studyで問題は解決する？

- 式(1)'のTWFE-DIDの問題は, (i) 単一の推定量を推定, (ii) 処置タイミングを基準化していない. Dynamic TWFEであれば問題を解決する？

$$Y_{it} = \alpha_i + \lambda_t + \sum_{l \neq -1, l \in L} \delta_l \mathbb{I}[t - k_i = l] + \epsilon_{it} \quad (7)$$

命題 (Proposition 1, Sun-Abraham, 2021)

各期 l のEvent study TWFEのEstimand (被推定量) δ_l^{ES} は, 全ての期, 全ての処置グループの処置効果(dynamic effects)の加重和として表される (詳細は資料編).

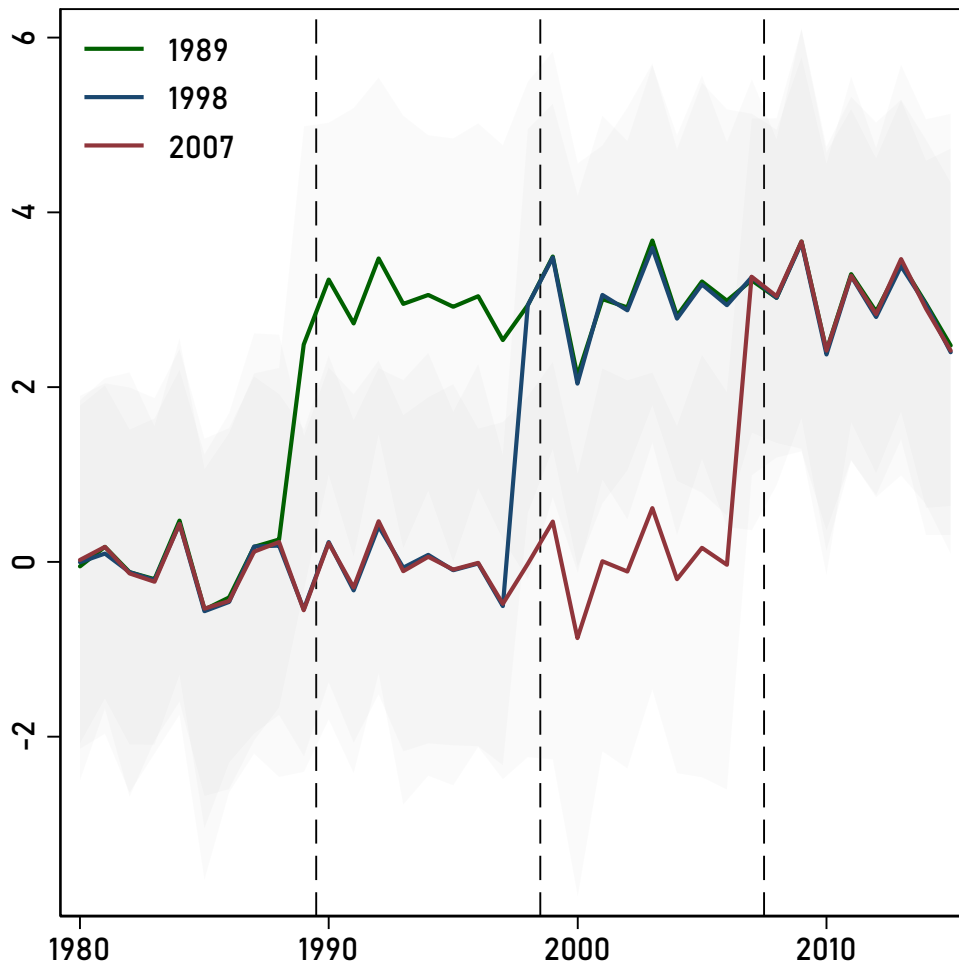
- ※1. 「全ての期」には Event study の 対象期間から外した期間 も含まれてしまう.
- ※2. 線形結合の「weight」にはマイナスも含む. 他の期間の処置効果が差引かれる“negative weight”となる場合も.
- ※3. 論文では, 各期 l では無く各window g に対する命題として示されているが, 式(7)に合わせて解釈している.

問題の診断は可能？

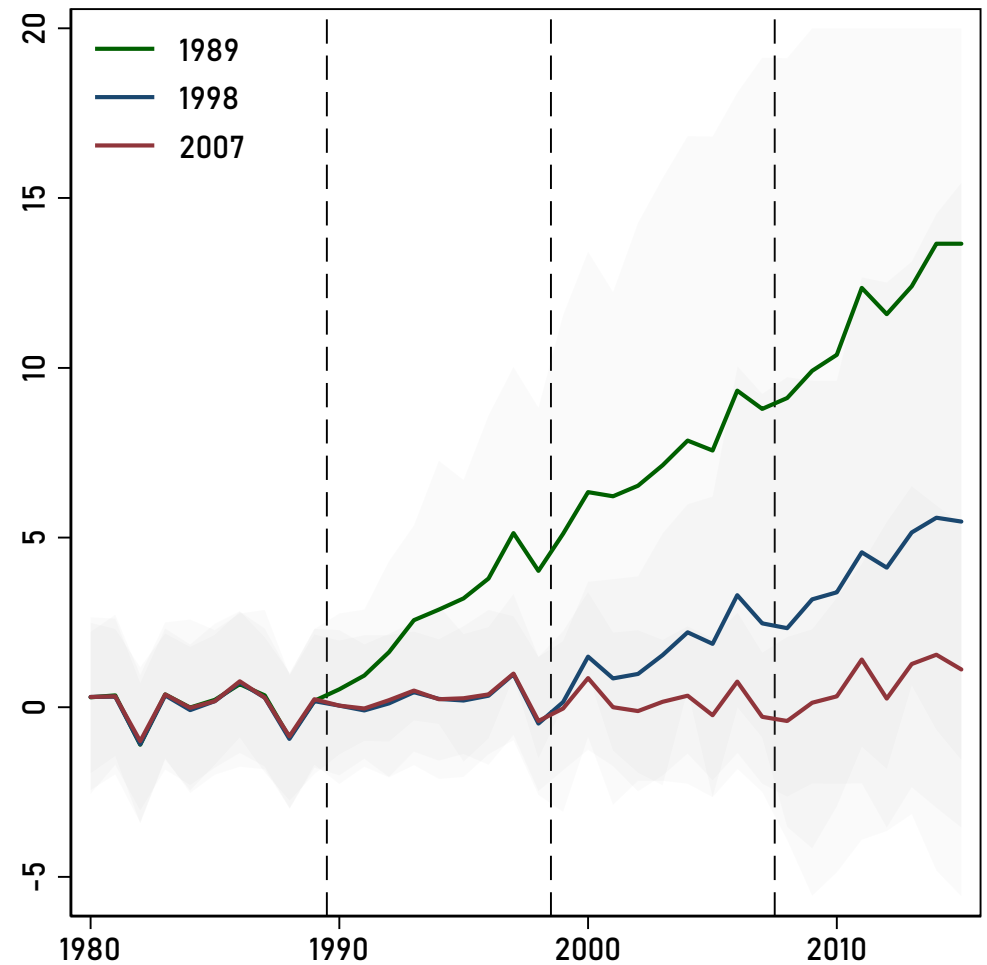
- Goodman-Bacon (JE, 2021) :
TWFE推定量は各 2×2 DID推定量の正の重み付け線形結合.
⇒ 2×2 DID推定値と重みのばらつきによる診断
(Stata: `bacondecomp`, R: `bacondecomp`)
- dCDH (AER, 2020) :
TWFE被推定量は真の処置効果の(一部負の)重み付け線形結合.
⇒ 重みの(処置グループ間の)分布による診断
(Stata: `twowayfeweights`, R: `TwoWayFEWeights`)
- Sun-Abraham (JE, 2021) : Event-study被推定量は全ての初期期間の真の処置効果の(一部負の)重み付け線形結合.
⇒ 重みの(処置時点間の)分布による診断
(Stata: `eventstudyweights`, Rは無い(2023.4.30時点))
- いずれも Negative Weightの問題に特化しており, データに内在する問題の特徴化はできない. ⇒ Jakiela (2021)

Goodman-Bacon Decomposition in Stata

Case 1. Homogeneous/No Dynamic Effect
(Simulation 3 in Baker *et al.*)

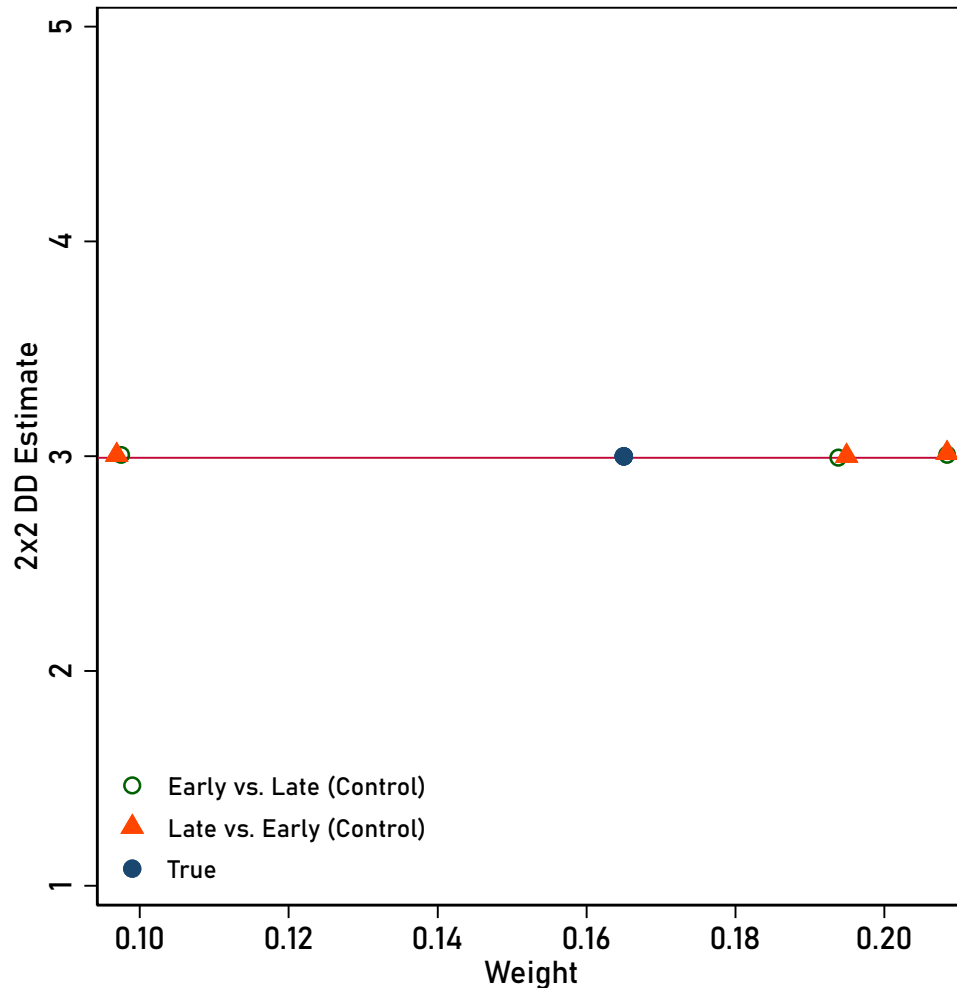


Case 2. Heterogeneous/Dynamic Effect
(Simulation 6 in Baker *et al.*)

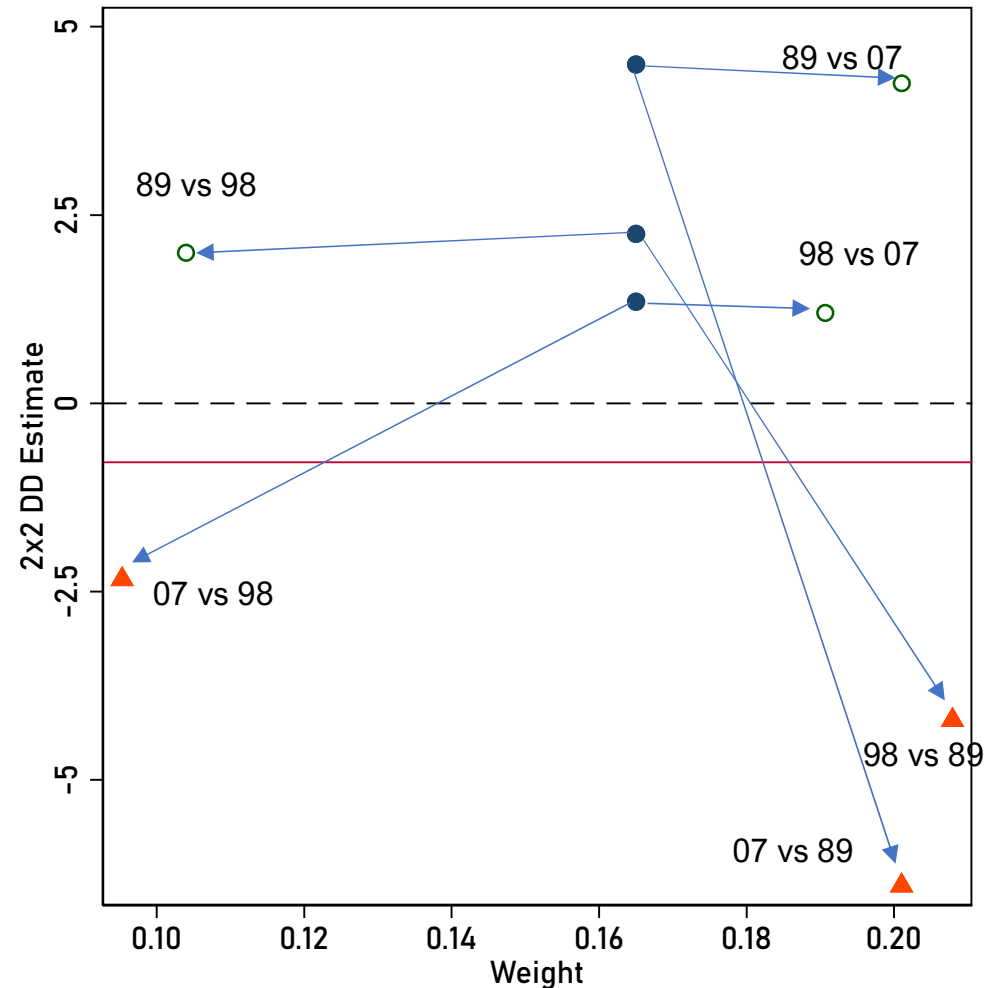


Goodman-Bacon Decomposition in Stata

Case 1. Homogeneous/No Dynamic Effect
(Simulation 3 in Baker *et al.*)



Case 2. Heterogeneous/Dynamic Effect
(Simulation 6 in Baker *et al.*)



Note 1: 分かり易くするため, “True effect”はBaker *et al.* (2022)とは異なり, 真のATTのみをプロットしている.

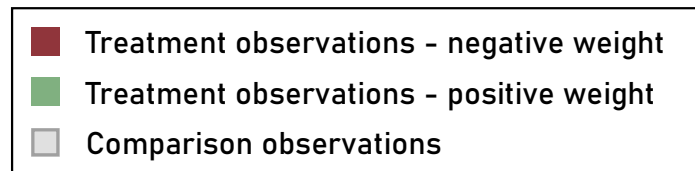
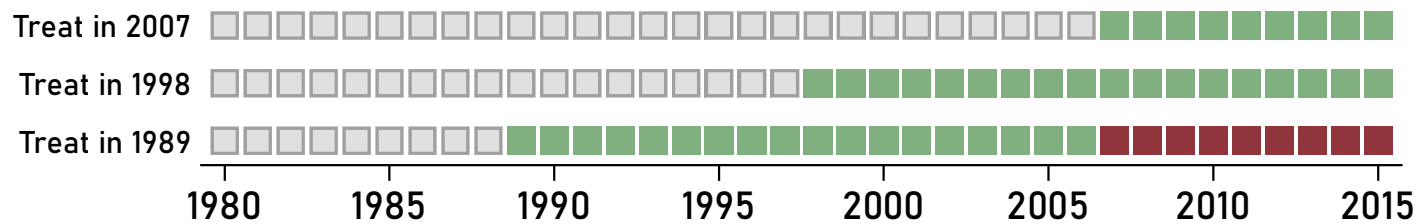
Note 2: 2007年の処置グループに関しては, 真の対照群として使用できるグループが存在しないため, 推定値が存在しない点に注意.

Jakiela's Diagnosis

Frisch–Waugh–Lovell (FWL) theoremから二つの性質が得られる：

$$(FWL1) \quad \hat{\delta}^{TWFE} = \sum_{it} Y_{it} \left(\frac{\tilde{D}_{it}}{\sum_{it} \tilde{D}_{it}^2} \right)$$

⇒ 観測値ごとの weights ($\tilde{D}_{it} / \sum_{it} \tilde{D}_{it}^2$)を計算し、どの観測値が negative weightsを持つかを評価.

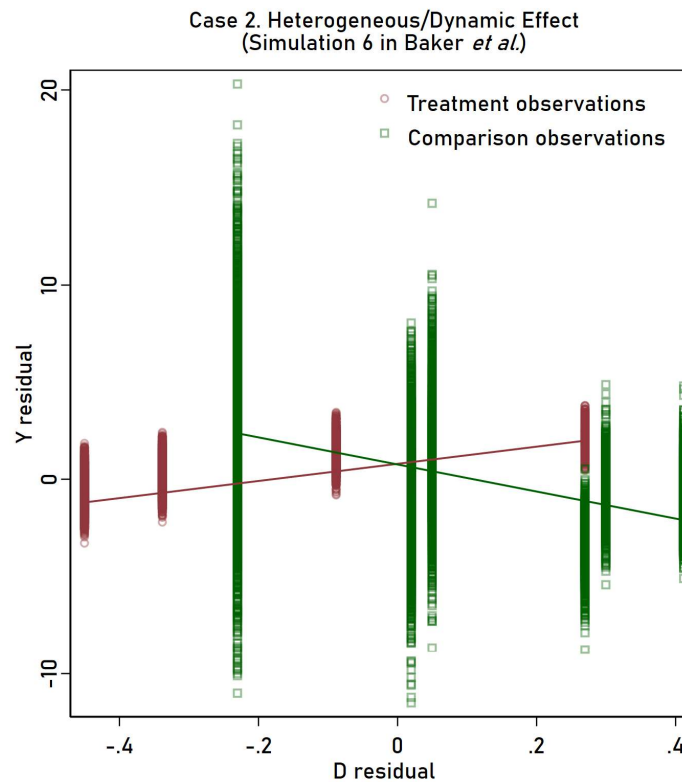
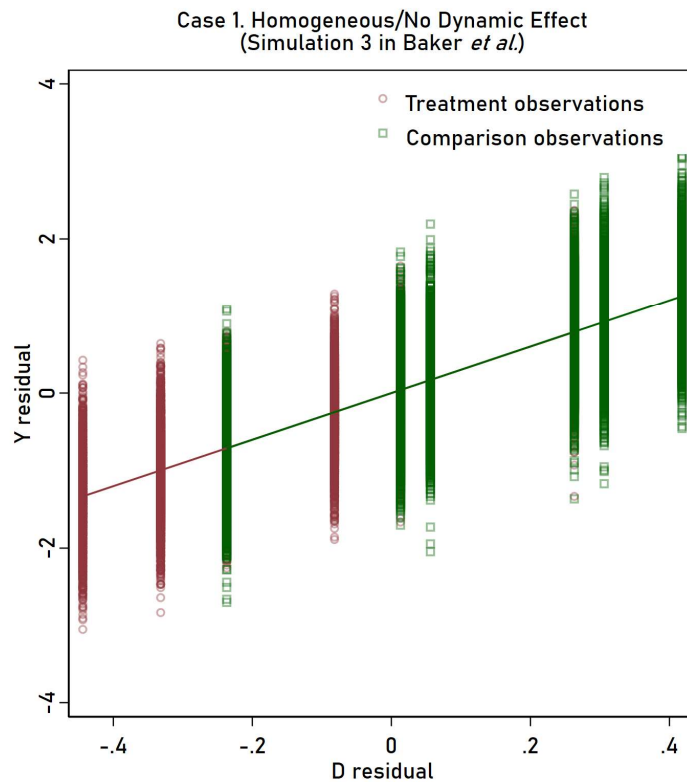


Jakiela's Diagnosis

Frisch–Waugh–Lovell (FWL) theoremから二つの性質が得られる：

$$(FWL2) \quad \hat{\delta}^{TWFE}: \tilde{Y}_{it} = \delta \tilde{D}_{it} + \tilde{\epsilon}_{it}$$

⇒ Residualize された被説明変数と処置変数の関係性が、線形か否か、大きな異質性を持つか否かを評価。



Summary So Far

1. 問題なのは「Staggered DIDの研究デザイン」ではなく「TWFE-DID estimator/specification」
 - 多くの場合（DIDの三つの仮定が成り立っていれば）DIDの研究デザインは因果効果の識別・推定戦略として妥当。
 - 問題は、処置が異なるタイミングで異なるグループに生じ、その効果に異質性がある場合、単一のDIDパラメーターを推定する方法では重大なバイアスが生じる。
 - 効果の異質性は、全ての実証研究で存在すると仮定するのが妥当(de Chaisemartin-D'Haultfoeuille, NBER, 2022)。
 - この問題は、処置が連続変数か否かに限らず生じる(de Chaisemartin-D'Haultfoeuille, NBER, 2022)。
- ⇒ どうすれば“Forbidden Comparison”を避けられるか？

Summary So Far

2. 処置効果の異質性が明らかでない場合，適切な“Aggregation”の方法を考える必要（バイアスはweightingからも生じる）。

- Group-specific 平均処置効果：

$$\hat{\theta}_k = \frac{1}{T - k + 1} \sum_{t \geq k} \widehat{ATT}_{t,k}$$

- Event time-specific 平均処置効果：

$$\hat{\theta}_l = \sum_k w_{l,k} \widehat{ATT}_{l,k}$$

- Overall 平均処置効果：

$$\hat{\theta}^* = \sum_k w_k \hat{\theta}_k$$

対処法は？

1. Fully-saturated TWFE Regression

- Cengiz *et al.* (QJE, 2019)
- Sun and Abraham (JE, 2021)
- Wooldridge (WP, 2021)

2. Rolling Methods

- Callaway-Sant'Anna (JE, 2021)
- de Chaisemartin-D'Haultfoeuille (AER, 2020)

3. Imputation Methods

- Borusyak, Jaravel, Spiess (R&R at RES, 2023)
- Gardner (WP, 2021)

Fully-saturated TWFE

1-A. Stacked Regression (Cengiz *et al.*, QJE, 2019)

- ① Event timing ごとに“クリーンな2×2 DIDデータセット”を作成 (各グループに対して not-yet treatedのみを対照群として使用).
- ② これらをデータセットをstackし, 推定用データセットを作成.
- ③ Stacked dataにfully saturated event-study DIDを適用する :

$$Y_{itk} = \alpha_{ik} + \lambda_{tk} + \sum_{l \neq -1, l \in L} \delta_l \mathbb{I}[t - k = l] + \epsilon_{itk} \quad (7)'$$

※. Fixed effectsをevent timingごとに入れる以外, 通常のevent-study DIDと同じ. 異なるのは, event timingごとにクリーンなデータセットを作成する点.

Fully-saturated TWFE

1-B. Sun and Abraham (JE, 2021)

- ① (Full data setに対して) Event timingグループを表すダミーとの交差項を入れたEvent-study DIDを推定：

$$Y_{it} = \alpha_i + \lambda_t + \sum_{k \notin C} \sum_{l \neq -1} \delta_{l,k} \mathbb{I}[G_k = k] \mathbb{I}[t - k = l] + \epsilon_{it} \quad (8)$$

- ② Dynamic treatment effects ATT_l を，次のように推定：

$$ATT_l = \sum_k w_{l,k} \hat{\delta}_{l,k}$$

※. (8)による推定は，実質的に never-treated のみを対照群として利用することになる．Never-treated が存在しない場合，last-treated を対照群として使用するため処置群に入れない．

Fully-saturated TWFE

1-C. Wooldridge (WP, 2021)

- ① (Full data setに対して) Event timingグループを表す固定効果と Event timingグループとの交差項を入れたTWFEを推定：

$$Y_{it} = \alpha_{k(i)} + \lambda_t + \sum_{k \notin C} \sum_{l > 0} \delta_{l,k} \mathbb{I}[G_k = k] \mathbb{I}[t - k = l] + \epsilon_{it} \quad (8)'$$

- ② Dynamic treatment effects ATT_l については、Stata/Rの制約付き検定パッケージを利用（例：Stata's `lincom/margins`）。

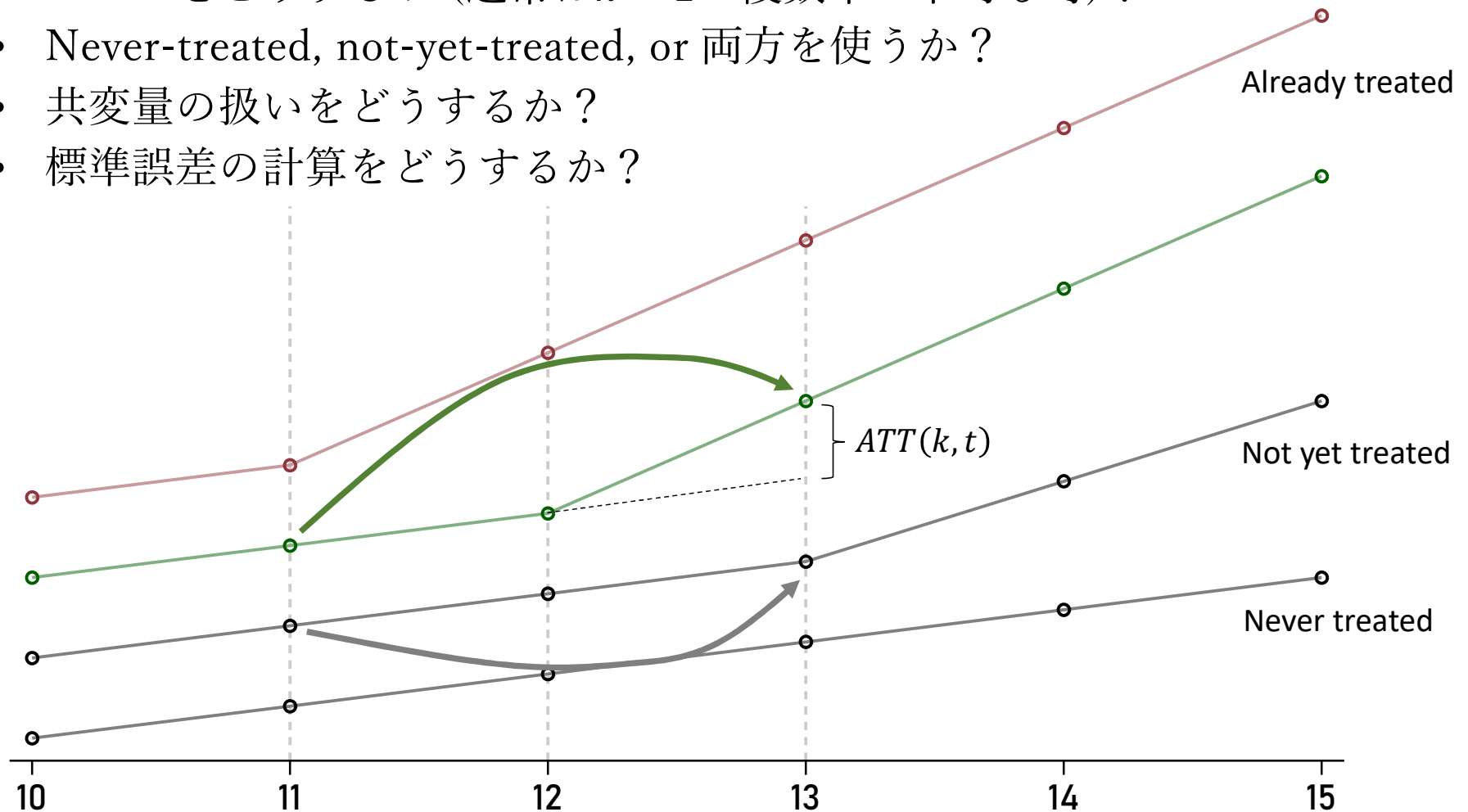
※1. Sun-Abrahamと同様に、実質的に never-treated のみを対照群として利用することになる。Never-treated が存在しない場合、last-treated を対照群として使用するため処置群に入れない。

※2. Time-varying covariatesがCPTに必要な場合の識別・推定についても明示的に議論している。

Rolling Methods

処置グループ k の t 年目の処置効果 $ATT(k, t)$:

- ベースをどうするか (通常は $k - 1$; 複数年の平均も可)?
- Never-treated, not-yet-treated, or 両方を使うか?
- 共変量の扱いをどうするか?
- 標準誤差の計算をどうするか?



Rolling Methods

2-A. Callaway-Sant'Anna Estimator (JE, 2021)

- (CPTが成り立つ時) ATTに関して、次の結果が成り立つ：

$$\begin{aligned} ATT(k, t) &\equiv E[Y_t^1 - Y_t^0 | X, G_k = 1] \\ &= E \left[\left(\frac{G_k}{E[G_k]} - \frac{\frac{p_k(X)C}{1-p_k(X)}}{E\left[\frac{p_k(X)C}{1-p_k(X)}\right]} \right) (Y_t - Y_{k-1} - m_{t,k}(X)) \right] \end{aligned} \quad (9)$$

$G_k = 1$ if treated in $t = k$; $C = 1$ if never-treated

$p_k(X) = \Pr(G_k = 1 | X, G_k, C)$ (generalized propensity score)

$m_{t,k}(X) = E[Y_t - Y_{k-1} | X, C = 1]$ (counterfactual outcome trend)

⇒ 式(9)を利用し、セミパラメトリックに推定； ATT_t に関しては、Sun and Abraham と同様。多重検定問題にも対応。

Rolling Methods

2-B. de Chaisemartin-D'Haultfoeuille Estimator (AER, 2020; WP, 2022)

- “joiners”と“leavers”に関して、次の推定量を定義：

$$ATT_+(t) \equiv \sum_{k:D_{k,t}=1, D_{k,t-1}=0} \frac{N_{k,t}}{N_{1,0,t}} (Y_{k,t} - Y_{k,t-1}) - \sum_{k:D_{k,t}=0, D_{k,t-1}=0} \frac{N_{k,t}}{N_{0,0,t}} (Y_{k,t} - Y_{k,t-1})$$

$$ATT_-(t) \equiv \sum_{k:D_{k,t}=1, D_{k,t-1}=1} \frac{N_{k,t}}{N_{1,1,t}} (Y_{k,t} - Y_{k,t-1}) - \sum_{k:D_{k,t}=0, D_{k,t-1}=1} \frac{N_{k,t}}{N_{0,1,t}} (Y_{k,t} - Y_{k,t-1})$$

$$\Rightarrow \delta^S = E \left(\sum_t \frac{N_{1,0,t}}{N_S} ATT_+(t) + \frac{N_{0,0,t}}{N_S} ATT_-(t) \right) \quad (10)$$

- ※1. $N_{a,b,t} = \sum_{k:D_{k,t}=a, D_{k,t-1}=b} N_{k,t}$; $N_{k,t} = \#$ of obs. for k in t .
- ※2. 処置にOn/Offがない場合, $ATT_+(t)$ のみを使用.
- ※3. Stataでは, CS estimator に比べ様々なオプションが付いている.
- ※4. 式(10)はStatic版. Dynamic版についてはWP(2022)を参照.

Imputation Methods

3-A. Borusyak *et al.* (R&R at RES, 2023)

- ① 処置を受けていない観測値($D_{it} = 0$)のみを使い, 次のような回帰式を推定する.

$$Y_{it} = \alpha_i + \lambda_t + u_{it}$$

- ② 処置を受けた観測値($D_{it} = 1$)を使い, 各観測体(i, t)の処置効果を次のように推定:

$$\hat{\delta}_{it} = Y_{it} - \hat{Y}_{it}(0) = Y_{it} - \hat{\alpha}_i - \hat{\lambda}_t$$

- ③ (SA/CS/deCDHと同様に) Dynamic treatment effects ATT_l を, 次のように推定:

$$\hat{\delta}_l = ATT_l = \sum_{\{i,t|t-k(i)=l\}} w_{it} \hat{\delta}_{it}$$

Imputation Methods

3-B. Gardner (WP, 2021)

- ① 処置を受けていない観測値($D_{it} = 0$)のみを使い, 次のような回帰式を推定する.

$$Y_{it} = \alpha_i + \lambda_t + u_{it}$$

- ② 全ての観測値に対して, 残差 (“Residualized outcome”)を計算 :

$$\tilde{Y}_{it} = Y_{it} - \hat{\alpha}_i - \hat{\lambda}_t$$

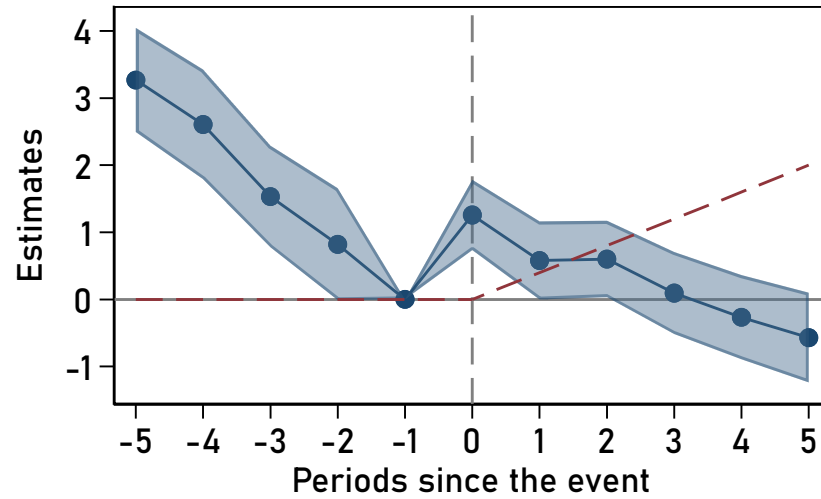
- ③ ②の残差を被説明変数として, 次の推定を行う :

(Static TWFE) $\tilde{Y}_{it} = \delta D_{it} + \epsilon_{it}$

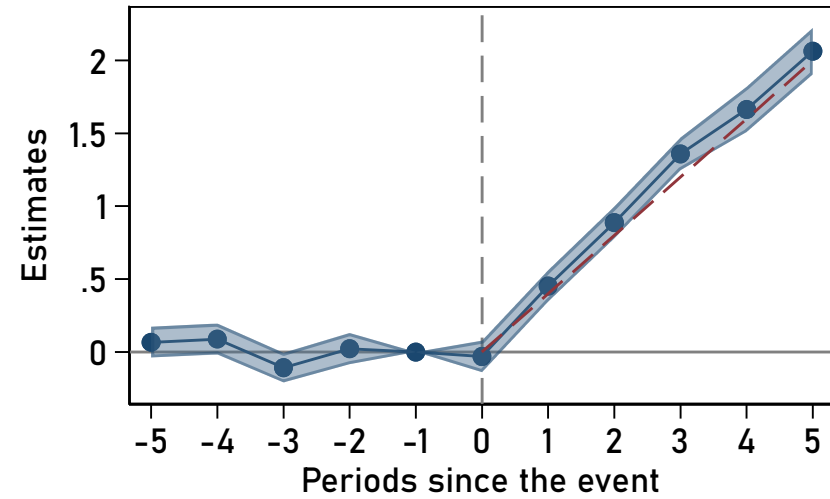
(Dynamic TWFE) $\tilde{Y}_{it} = \sum_{l \neq -1} \delta_l \mathbb{I}[t - k_i = l] + \epsilon_{it}$

Alternative Estimators in Stata

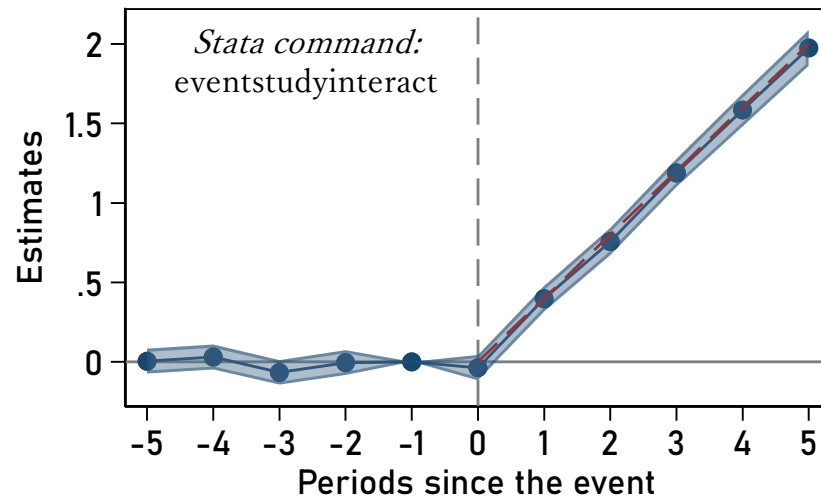
Naive TWFE Event Study



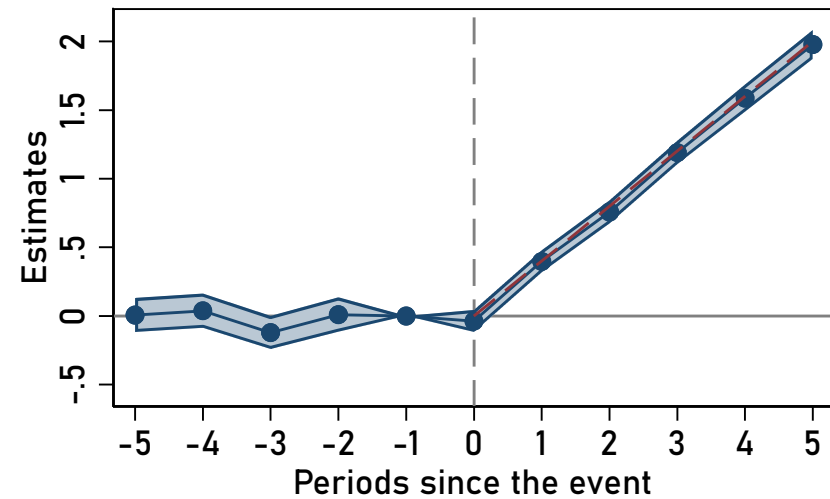
Stacked Regression



Sun and Abraham

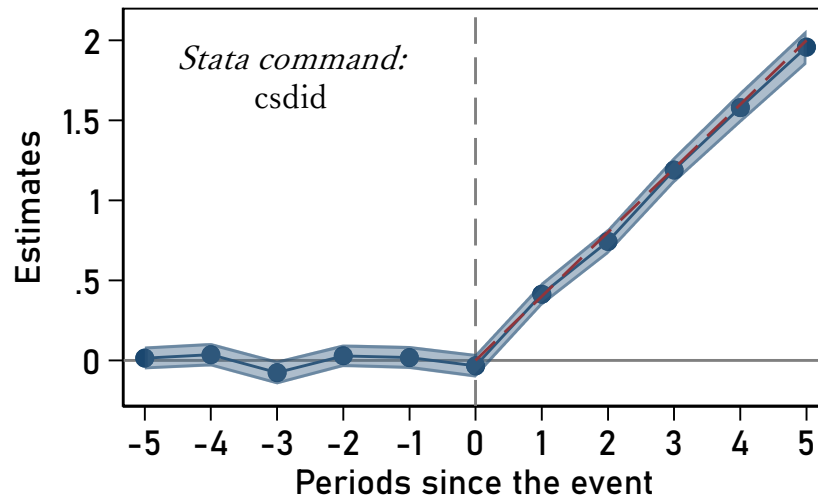


Wooldridge

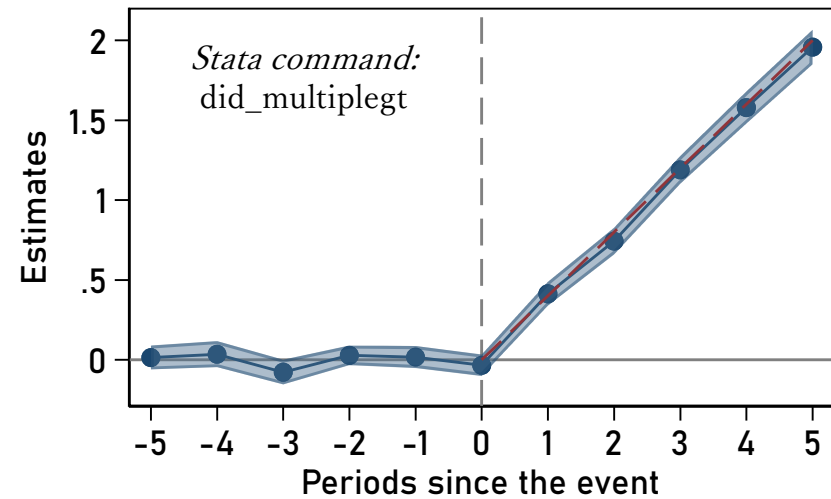


Alternative Estimators in Stata

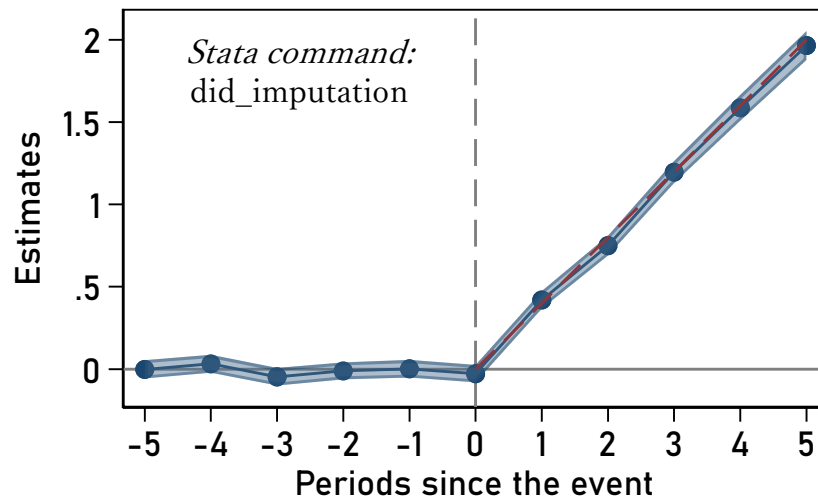
Callaway and Sant'Anna



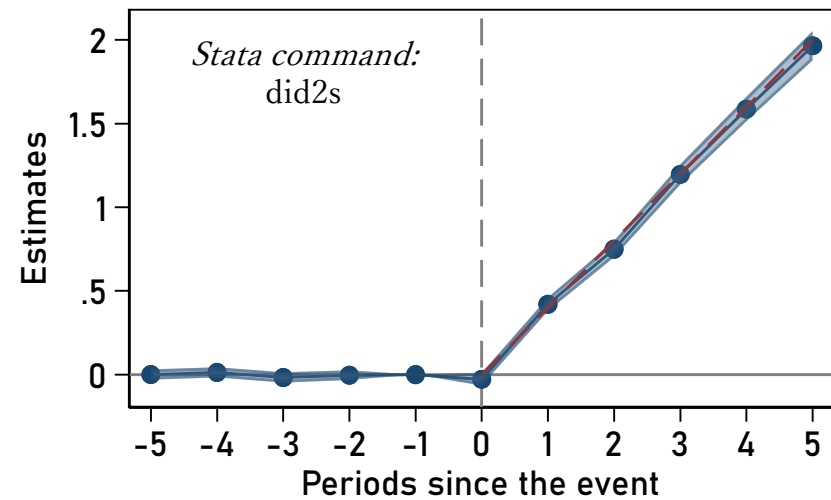
de Chaisemartin and D'Haultfoeulle



Borusyak, Jaravel, Spiess



Gardner



Stata/Rによる演習

(1) 基礎編

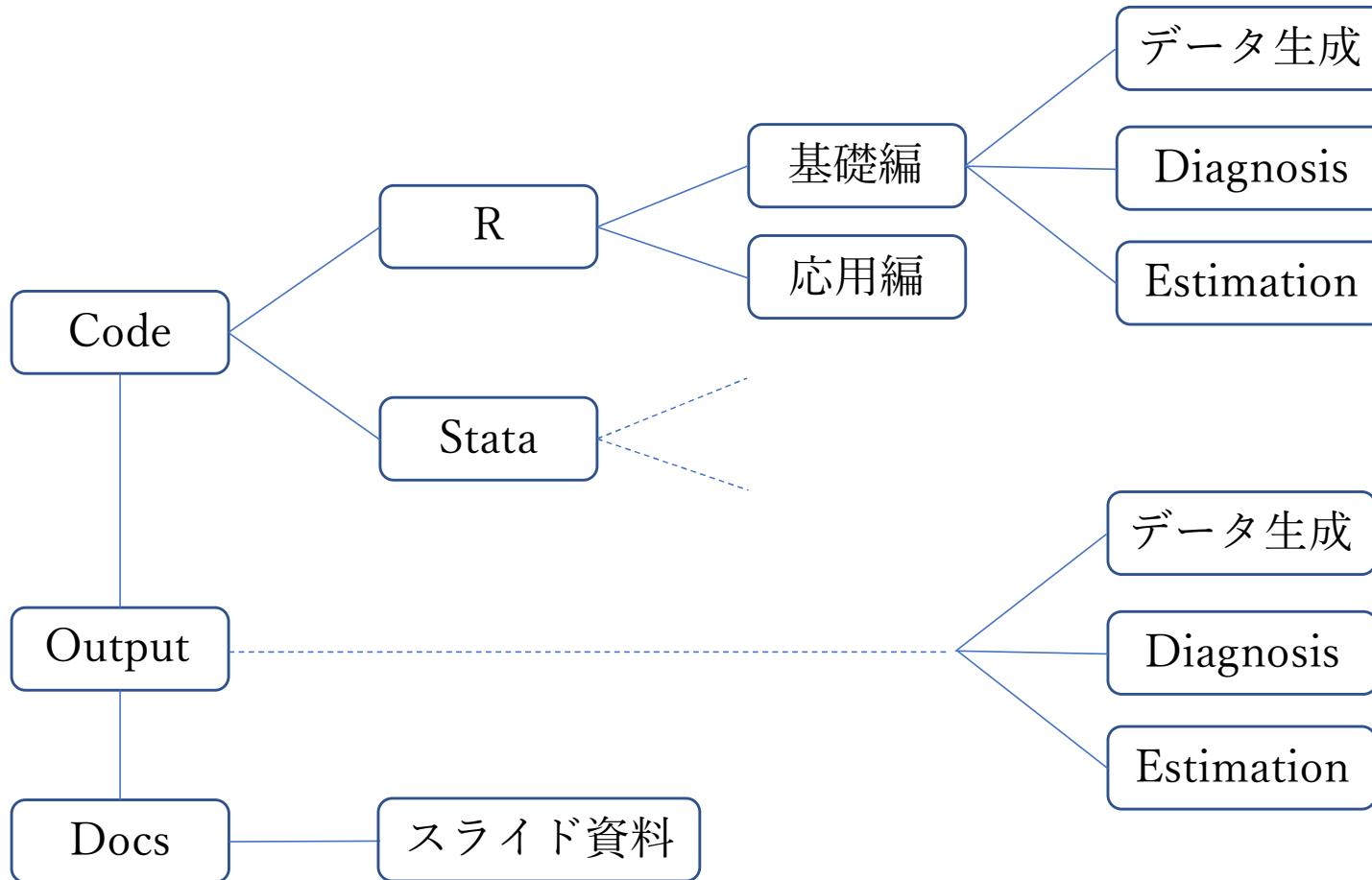
- Baker *et al.* の仮想例
(Heterogeneous, Dynamic Effects)
 - PackageのDL
 - Diagnosisの方法
 - Goodman-Bacon decomposition
 - Jakiela's diagnosis
 - 三つのアプローチによる推定
 - Fully saturated TWFE
 - Rolling methods
 - Imputation methods

Stata/Rによる演習

(2) 実践編

- Cicala (AER, 2022) のデータ
(電力自由化が経済厚生に与えた影響)
 - 実証研究の文脈と研究デザイン
 - 実証研究で生じうる問題は？
 - Diagnosis
 - Jakiela's diagnosis
 - 三つのアプローチによる推定
 - SA vs. CS vs. deCDH vs. BJS

サポートサイト (Github)



※1. サポートサイトはこちら：github.com/kazuyanagimoto/staggered_did_tutorial.

※2. Stataコードは私, Rコードは柳本和春さん(CEMFI)と梅谷隼人さん(神戸大学)作成.

※3. 教育・研究目的での利用は引用の上, 商用利用の場合は私の許諾を得てからご利用下さい.

Lessons from Stata/R Exercise

L 1. Negative Weightsの診断は簡単に行うことが出来る.

- Stata/R packageよりもJakielaのDiagnosisを推奨.

L 2. Sources of heterogeneityが既知の場合, 三つのアプローチいずれも真の処置効果およびPre-trendsを正しく推定する.

- Stacked regression, Wooldridge, Gardner 等による“手動”推定は, 正しくAggregationを行った場合とSEの推定値が異なる可能性.

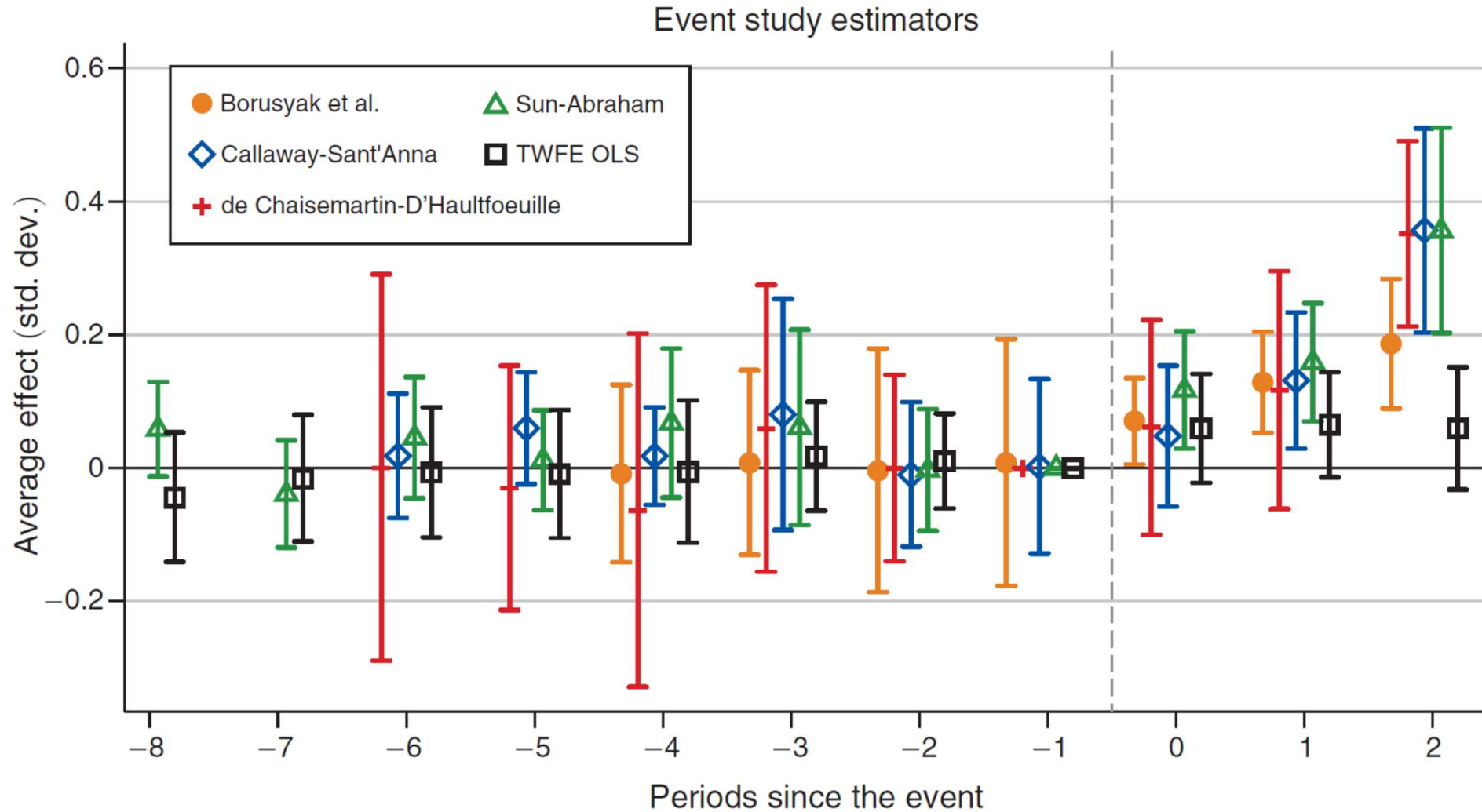
L 3. Sources of heterogeneityが既知でない現実の実証研究では, 三つのアプローチの推定値(およびSE)は大きく異なる可能性がある.

- それぞれのEstimatorは, 異なるデータの変分を利用しているため, 推定値の不偏性・効率性ともに実証上は不定.
- ⇒ 単一の推定量だけに頼らず複数のアプローチで推定.

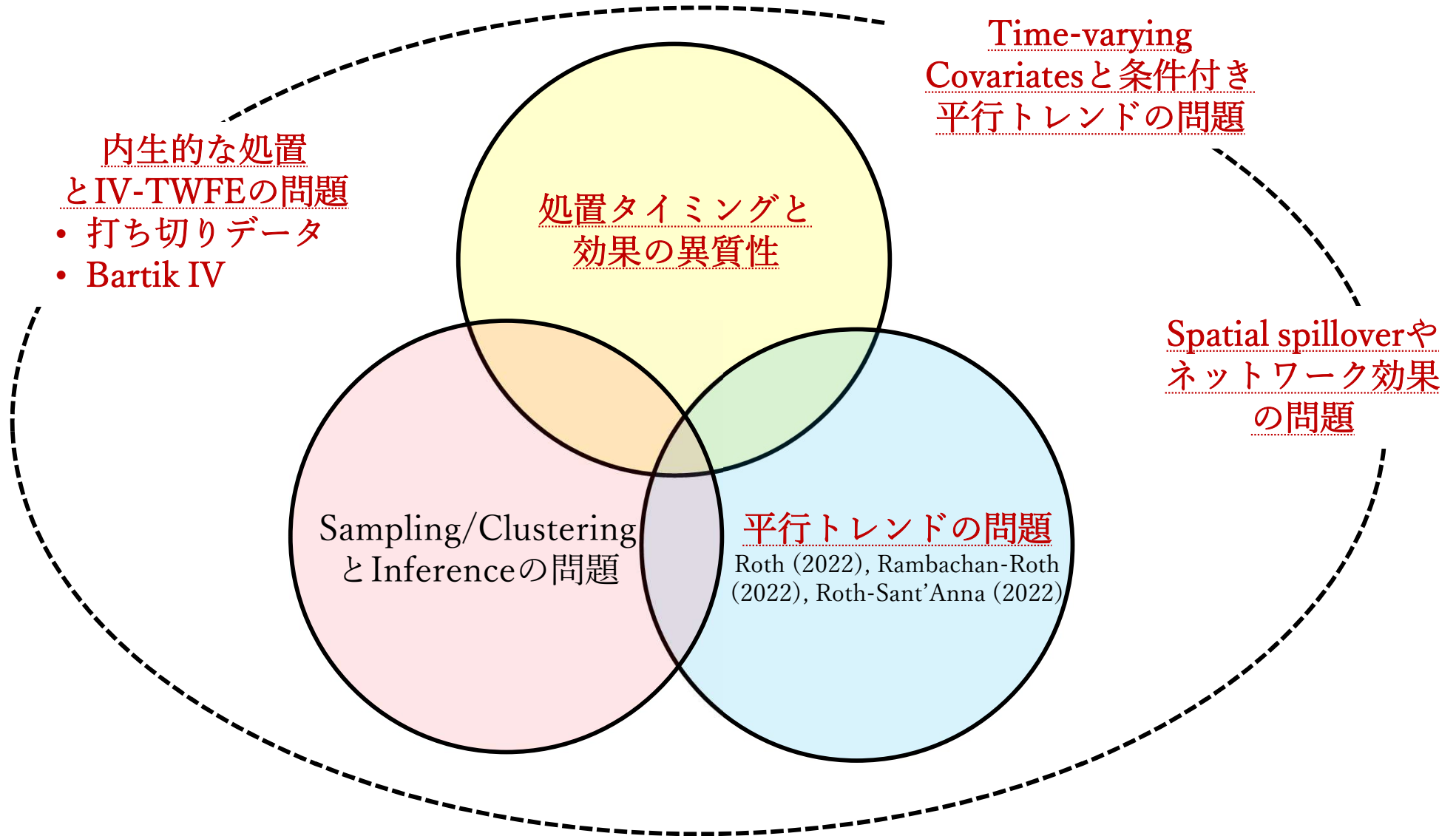
L 4. Time-varying Covariatesは, Imputation methodsとRolling methodsを組み合わせる事で対処可能.

- より複雑なケース($\Delta X(1) \neq \Delta X(0)$): Caetano *et al.* (WP, 2022).

Good Practice in Event Study



What's Next in DID?



参考文献(1)

- Athey, Susan, Guido W. Imbens (2022) Design-based analysis in Difference-In-Differences settings with staggered adoption. *Journal of Econometrics*
- Baker, Andrew, David F. Larcker, Charles Wang (2021) How Much Should We Trust Staggered Difference-In-Differences Estimates? *Journal of Financial Economics*
- Beck, Thorsten, Ross Levine, Alexey Levkov (2010) Big Bad Banks? The Winners and Losers from Bank Deregulation in the United States. *Journal of Finance*
- Borusyak, Kirill, Xavier Jaravel, Jann Spiess (2023) Revisiting Event Study Designs: Robust and Efficient Estimation. *R&R at Review of Economic Studies*
- Callaway, Brantly, Andrew Goodman-Bacon, Pedro Sant'Anna (2021) Difference-in-Differences with a Continuous Treatment. *Working paper*
- Callaway, Brantly, Pedro Sant'Anna (2021) Difference-in-Differences with Multiple Time Periods. *Journal of Econometrics*
- Caetano, Carolina, Brantly Callaway, Stroud Payne, Hugo Sant'Anna Rodrigues (2022) Difference in Differences with Time-Varying Covariates. *Working paper*
- Cengiz, Doruk, Arindrajit Dube, Attila Lindner, Ben Zipperer (2019) The Effect of Minimum Wages on Low-wage Jobs. *Quarterly Journal of Economics*

参考文献(2)

- Cicala, Steve (2022) Imperfect Markets versus Imperfect Regulation in US Electricity Generation. *American Economic Review*
- de Chaisemartin, Clément, Xavier D'Haultfoeuille (2020) Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. *American Economic Review*
- de Chaisemartin, Clément, Xavier D'Haultfoeuille (2022) Difference-in-differences Estimators of Intertemporal Treatment Effects. NBER *Working Paper*
- de Chaisemartin, Clément, Xavier D'Haultfoeuille (2022) Two-Way Fixed Effects and Difference-in-Differences with Heterogeneous Treatment Effects: A Survey. *NBER Working paper*
- Gardner, John (2021) Two-stage Differences in Differences. *Working paper*
- Goodman-Bacon, Andrew (2021) Difference-in-Differences with Variation in Treatment Timing. *Journal of Econometrics*
- Jakiela, Pamela (2021) Simple Diagnostics for Two-Way Fixed Effects. *Working paper*
- Powell, D. (2021) Labor Supply Consequences of the Opioid Crisis. *Working paper*
- Rambachan, Ashesh, Jonathan Roth (2023) A More Credible Approach to Parallel Trends. *The Review of Economic Studies*

参考文献(3)

Roth, Jonathan, Pedro Sant'Anna (2021) Efficient Estimation for Staggered Rollout Designs. *Working paper*

Roth, Jonathan, Pedro Sant'Anna (2022) When Is Parallel Trends Sensitive to Functional Form? *Forthcoming at Econometrica*

Roth, Jonathan, Pedro H. C. Sant'Anna, Alyssa Bilinski, John Poe (2022) What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature. *Forthcoming at Journal of Econometrics*

Sant'Anna, Pedro, Jun Zhao (2020) Doubly Robust Difference-in-Differences Estimators. *Journal of Econometrics*

Słoczyński, Tymon (2020) Interpreting OLS Estimands When Treatment Effects Are Heterogeneous: Smaller Groups Get Larger Weights. *Review of Economics & Statistics*

Sun, Liyang, Sarah Abraham (2020) Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects. *Journal of Econometrics*

Wooldridge, Jeffrey M. (2021) Two-Way Fixed Effects, the Two-Way Mundlak Regression, and Event Study Estimators. *Working paper*

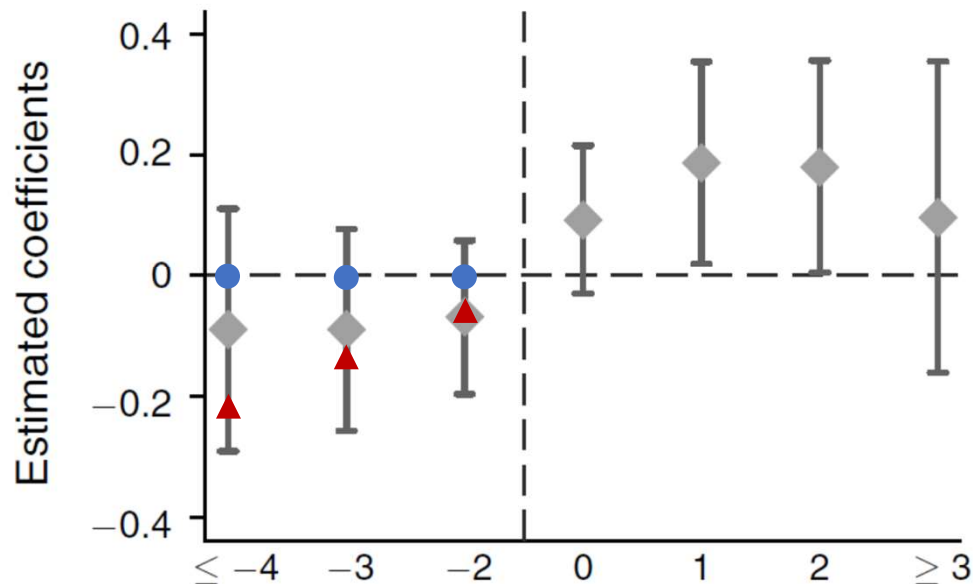
R and Stata packages

<https://asjadnaqvi.github.io/DiD/>

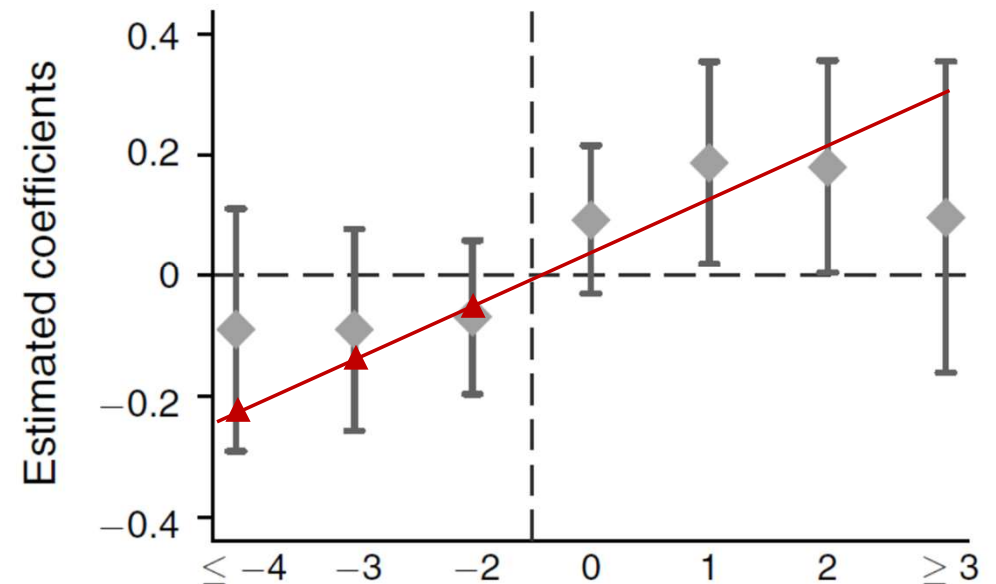
Pre-trend Check の問題点

- Roth (AER Insight, 2022) :
 - (1) Event studyのpre-trend testは検出力が弱い可能性.
 - (2) Pre-trendの結果に基づいてspecificationを選ぶとバイアスに繋がる可能性.
- 例えば, 左図において $H_0: \beta = 0$ (青丸)と $H_0: \beta =$ 赤丸という二つの帰無仮説のp値はほぼ同じ. もし仮に赤丸が真のトレンドであった場合, 右図のようなトレンド線を引くことが可能となり, ATTはほぼゼロになる.

Panel A. Effect on subsidized population



Panel A. Effect on subsidized population



Parallel Trend の感度分析

この問題の対処法は？

- Roth (AER Insight, 2022) : このような“線”が引けるかどうかで検出力判定.
Stata/R package: **pretrends** (github.com/mcaceresb/stata-pretrends)
- Rambachan and Roth (ECTA, 2022) : 次のような感度分析を行う.
 - Pre-trendからPost-treatment trendを予測
 - Post-treatment trendに $|M|$ のBoundsを入れることでRobust CIを作る
 - M を変化させることで, ATTの有意性の感度分析を行う

Stata/R package : **honestdid** (github.com/mcaceresb/stata-honestdid)

