

DiDの展開

Roth et al. (2023) What's trending in difference-in-differences?

資料作成： 田中 鮎夢
青山学院大学

2024-06-09

1. はじめに

- 本資料は、Roth et al. (**2023**) とそのワーキング・ペーパー・バージョンに従って、DiDの基本的な概念を説明し、最近の研究の展開について議論する。
 - Roth, Jonathan, Pedro HC Sant'Anna, Alyssa Bilinski, and John Poe. 2023. "What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature." *Journal of Econometrics* 235 (2): 2218–44. <https://doi.org/10.1016/j.jeconom.2023.03.008>.
- 論文の中でも、基礎的かつ重要な要点に絞る。（従って、原論文を適宜参照のこと）
 - 推定（estimation）に重点を置き、推論（inference）については解説を省く。証明はすべて省く。
 - 共変量なしの基礎ケースに重点を置き、共変量ありの応用ケースは簡潔な紹介にとどめる。
 - 数値例を追加。

2. 基本的なモデル

2.1. 処置の割り当てと時期

- 2つの期間 ($t = 1, 2$) が存在するモデルを考える。
- 個体は、 i でインデックスされ、次の2つの母集団のうちの1つから引き出される。
 - 処置群の母集団 ($D_i = 1$) の個体は、期間 $t = 1$ と $t = 2$ の間に処置を受ける。
 - 未処置群（比較群、対照群、対照群）の母集団 ($D_i = 0$) の個体は、両期間とも未処置のままである。

例) 最低賃金引き上げが雇用に与える影響

	Before (t=1)	After (t=2)
NJ, 処置群 (D=1)	No	処置
PA, 比較群 (D=0)	No	No

結果と処置状態

- 計量経済学者は、 $i = 1, \dots, N$ の個体、 $t = 1, 2$ の2期間のパネルについて、
- 結果(outcome): Y_{it} と
- 処置状態: D_i
を観察する。
- 単純なバランスト・パネル・データのセットアップに焦点を当てる。
 - (反復横断データや非バランスト・パネルにDiDを拡張することは可能)

例) 最低賃金引き上げが雇用に与える影響

	Before (t=1)	After (t=2)
NJ, 処置群 (D=1)	2	4
PA, 比較群 (D=0)	3	2

2.2. 潜在的な結果の枠組み

- 潜在的な結果 (potential outcomes) の枠組み (Rubin 1974, Robins 1986)
 - $Y_{it}(0, 0)$ は、 t 期において i が両期とも未処置であった場合の i の潜在的な結果。
 - $Y_{it}(0, 1)$ は、 i が最初の期間では未処置であるが、2番目の期間までに処置にさらされた場合の個体 i の期間 t の潜在的な結果。
- 表記を簡単にするために、以下のように書く。
 - $Y_{it}(0) = Y_{it}(0, 0)$
 - $Y_{it}(1) = Y_{it}(0, 1)$

最低賃金の例

例) 最低賃金引き上げが雇用に与える影響

	Before (t=1)	After (t=2)
NJ, 処置群 (D=1)	$Y_{i1}(1) = 2$	$Y_{i2}(1) = 4$
PA, 比較群 (D=0)	$Y_{i1}(0) = 3$	$Y_{i2}(0) = 2$

- 注: この例では簡単化のため、各群内の個体間でアウトカムに異質性がなく、期待値オペレータを省略できるとする。以下同様。

因果推論の基本的な問題

- 因果推論の基本的な問題（Holland, 1986）のために、我々は各個体*i*について2つの潜在的結果のうちの1つだけを観測。
- すなわち、観測された結果は、

$$Y_{it} = \underbrace{D_i Y_{it}(1)}_{\text{処置群}} + \underbrace{(1 - D_i) Y_{it}(0)}_{\text{比較群}}$$

によって与えられる。

- この潜在的結果の枠組みは、個体*i*の結果（outcome）が個体*j* ≠ *i*の処置状況に依存しないという**安定個体処置値仮定（SUTVA: stable unit treatment value assumption）**を暗黙的に内包しており、スピルオーバー効果を排除している。

ATT (処置群への平均処置効果)

- 正統的なDiDの設定において最も注目される因果推定対象 (causal estimand) は、期間 $t = 2$ における**処置群への平均処置効果** (ATT: average treatment effect on the treated) である。

$$\tau_2 = \mathbb{E}[Y_{i2}(1) - \underbrace{Y_{i2}(0)}_{\text{処置群}} | D_i = 1]$$

- これは、処置された個体に対する、処置された期間 ($t = 2$) における平均的な因果効果を測定したものである。
- 2番目の項、 $Y_{i2}(0)$ は、**反実仮想**。つまり、処置が行われなかった場合の処置群の平均結果を示す。

2.3. 平行傾向の仮定と識別

- τ_2 を識別する際の課題は、未処置の潜在的な結果（outcome）である $Y_{i2}(0)$ が、処置群（ $D_i = 1$ ）では観測されないことである。
- DiD法は、(1)未処置群の結果の変化と(2)処置群のベースラインの結果を用いて、処置群の平均的な反事実の未処置結果を推定する仮定によって、この識別問題を克服している。

仮定1「平行傾向の仮定」

- τ_2 を識別するための重要な仮定は、**平行傾向の仮定 (parallel trends assumption)** である。これは、処置が行われなかった場合、処置群と未処置群（比較群、対象群）の平均結果は平行に進展したであろうという直観的なものである。
- 仮定1（平行傾向 / Parallel Trends）

$$\underbrace{\mathbb{E} \left[\underbrace{Y_{i2}(0) - Y_{i1}(0)}_{\text{処置群}} \mid \underbrace{D_i = 1}_{\text{処置群}} \right]}_{\text{処置がなかった場合の変化 (反実)}} = \underbrace{\mathbb{E} \left[\underbrace{Y_{i2}(0) - Y_{i1}(0)}_{\text{比較群}} \mid \underbrace{D_i = 0}_{\text{比較群}} \right]}_{\text{処置がなかった場合の変化 (現実)}} \quad (1)$$

仮定 2 「予見効果なしの仮定」

- また、 τ_2 の識別に必要なもう一つの重要で隠れた仮定は、**予見効果なしの仮定 (no anticipatory effects assumption、無予測の仮定: no-anticipation assumption)** であり、これは、処置が実施される前には因果的効果がないとするものである。
- そうでなければ、第1期と第2期の処置群の結果 Y の変化は、 $t = 2$ 期の因果的効果だけでなく、 $t = 1$ 期の予見効果も反映している可能性があるからである (Abbring and van den Berg, 2003; Malani and Reif, 2015)。

- **仮定2 (予見効果なし / No Anticipatory Effects)**

$$Y_{i1}(0) = Y_{i1}(1) \text{ for all } i \text{ with } D_i = 1$$

- 意味: 処置群 ($D_i = 1$) にとって、処置があってもなくても、処置前の期間 ($t = 1$) の結果 (Y_{i1}) は同じはずである。

ATTの識別(1)

- 平行傾向と予見効果なしの仮定の下で、2期目のATT (τ_2) が識別される。
- 平行傾向の仮定 (再掲)

$$\mathbb{E} \left[\underbrace{Y_{i2}(0) - Y_{i1}(0)}_{\text{処置群}} \mid D_i = 1 \right] = \mathbb{E} \left[\underbrace{Y_{i2}(0) - Y_{i1}(0)}_{\text{比較群}} \mid D_i = 0 \right]$$

- 平行傾向の仮定 (上の式) の項を並べ替えることで、次のようになる。

$$\underbrace{\mathbb{E}[Y_{i2}(0) \mid D_i = 1]}_{\text{処置群の反実仮想の第2期の値}} = \underbrace{\mathbb{E}[Y_{i1}(0) \mid D_i = 1]}_{\text{処置群の反実仮想ベースライン}} + \underbrace{\mathbb{E}[Y_{i2}(0) - Y_{i1}(0) \mid D_i = 0]}_{\text{比較群の変化の期待値}}$$

ATTの識別(2)

- さらに、予見効果なしの仮定で、 $\mathbb{E}[Y_{i1}(0) | D_i = 1] = \mathbb{E}[Y_{i1}(1) | D_i = 1]$ のため、以下のように変形できる。

$$\begin{aligned}\mathbb{E}[Y_{i2}(0) | D_i = 1] &= \mathbb{E}[Y_{i1}(0) | D_i = 1] + \mathbb{E}[Y_{i2}(0) - Y_{i1}(0) | D_i = 0] \\ &= \mathbb{E}[Y_{i1}(1) | D_i = 1] + \mathbb{E}[Y_{i2}(0) - Y_{i1}(0) | D_i = 0]\end{aligned}$$

ATTの識別(3)

- 処置群について $Y(1)$ を、比較群について $Y(0)$ を観測していることを利用するとさらに変形できる:

$$\underbrace{\mathbb{E}[Y_{i2}(0) \mid D_i = 1]}_{\text{処置群の2期目の反実仮想的な平均結果}} = \mathbb{E}[Y_{i1}(1) \mid D_i = 1] + \mathbb{E}[Y_{i2}(0) - Y_{i1}(0) \mid D_i = 0]$$

$$= \underbrace{\mathbb{E}[Y_{i1} \mid D_i = 1]}_{\text{処置群の処置前の平均}} + \underbrace{\mathbb{E}[Y_{i2} - Y_{i1} \mid D_i = 0]}_{\text{比較群の結果の変化の平均}}$$

- よって、反実仮想値を、処置群の処置前の平均と比較群の結果の変化の平均の和によって代替できる。

ATTの識別(4)

- 処置群の $Y(1)$ を直接観測していることに注意して、以下のように式変形できる。

$$\begin{aligned}
 \tau_2 &= \mathbb{E}[Y_{i2}(1) - Y_{i2}(0) \mid D_i = 1] \\
 &= \mathbb{E}[Y_{i2}(1) \mid D_i = 1] - \mathbb{E}[Y_{i2}(0) \mid D_i = 1] \\
 &= \mathbb{E}[Y_{i2}(1) \mid D_i = 1] - \mathbb{E}[Y_{i1} \mid D_i = 1] - \mathbb{E}[Y_{i2} - Y_{i1} \mid D_i = 0] \\
 &= \underbrace{\mathbb{E}[Y_{i2} - Y_{i1} \mid D_i = 1]}_{\text{処置群の変化}} - \underbrace{\mathbb{E}[Y_{i2} - Y_{i1} \mid D_i = 0]}_{\text{比較群の変化}} \quad (2)
 \end{aligned}$$

- これにより、ATT (τ_2) が識別できる。

2.4. 推定と推論 (1) 手計算

- 手計算によるATT推定
- 期待値を標本平均値に置き換えることで、 τ_2 を推定できる。
$$\hat{\tau}_2 = (\bar{Y}_{t=2,D=1} - \bar{Y}_{t=1,D=1}) - (\bar{Y}_{t=2,D=0} - \bar{Y}_{t=1,D=0})$$
- ここで、 $\bar{Y}_{t=t',D=d}$ は、期間 t' の群 d の結果 Y の標本平均値。

最低賃金の例

- 観測された値からのATTの計算。

例) 最低賃金引き上げが雇用に与える影響

	Before (t=1)	After (t=2)	差
NJ, 処置群 (D=1)	$Y_{i1}(1) = 2$	$Y_{i2}(1) = 4$	$4-2=2$
PA, 比較群 (D=0)	$Y_{i1}(0) = 3$	$Y_{i2}(0) = 2$	$2-3=-1$
差	$2-3=-1$	$4-2=2$	$ATT=2-(-1)=3$

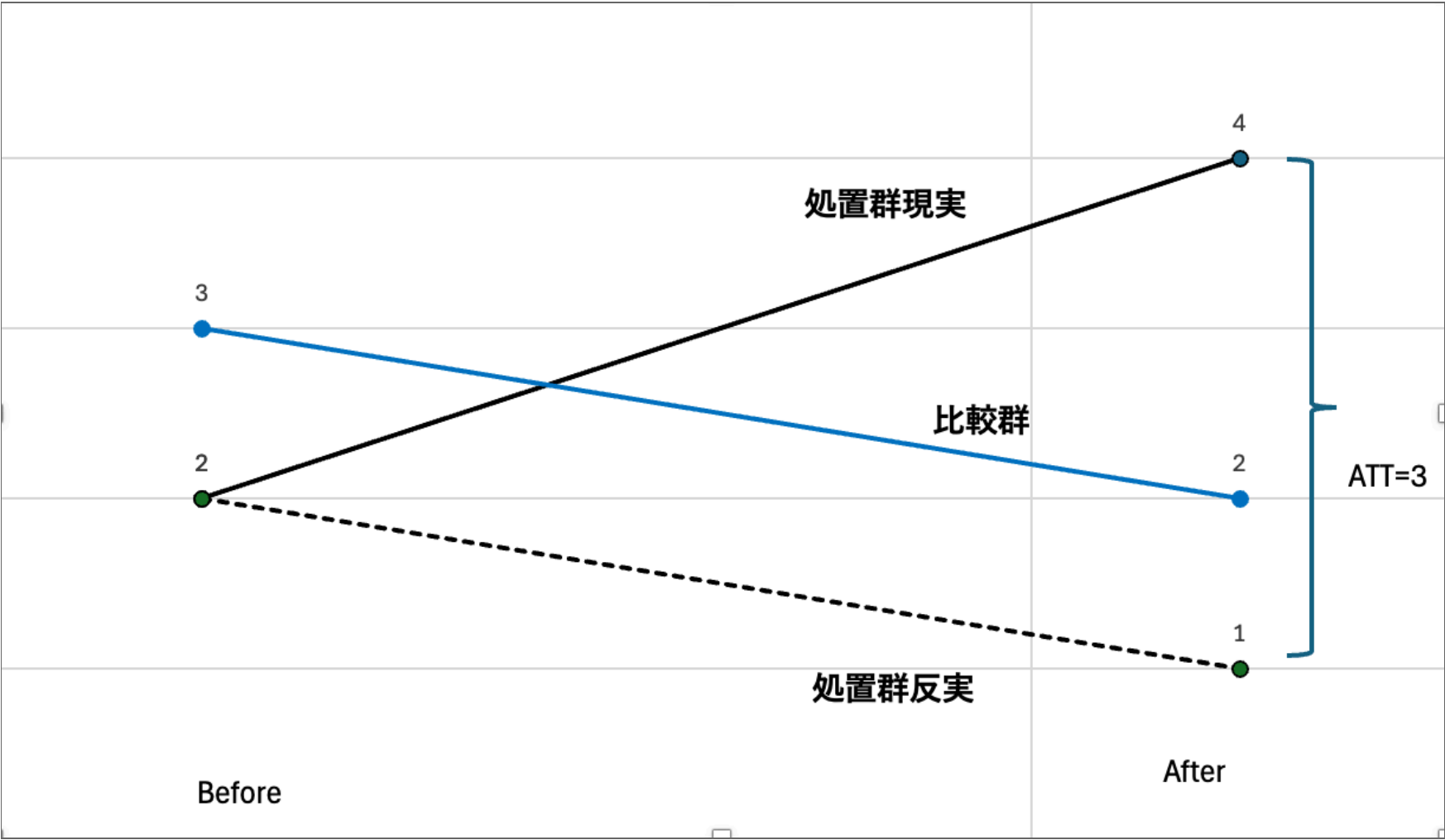
最低賃金の例

- 潜在的結果を明示したATTの計算。

例) 最低賃金引き上げが雇用に与える影響

	Before (t=1)	After (t=2)
現実, Y(1) NJ, 処置群 (D=1)	$Y_{i1}(1) = 2$	$Y_{i2}(1) = 4$
反実, Y(0) NJ, 処置群 (D=1)	予見効果なしより $Y_{i1}(0) = 2$	平行傾向より $Y_{i2}(0) = 1$
ATT		$= 4 - 1 = 3$

最低賃金の例



推定と推論 (2) TWFE

- Two-way fixed effects (TWFE) regression specification
- 二方向固定効果 (TWFE) 回帰式を使用して $\hat{\tau}_2$ を推定することで、標準誤差の計算が容易になる。

$$Y_{it} = \alpha_i + \phi_t + (1[t = 2] \cdot D_i) \beta + \epsilon_{it} \quad (3)$$

- 個体固定効果 α_i
- 時間固定効果 ϕ_t
- 処置後の指標 $1[t = 2]$ と処置群 D_i との相互作用
- この標準的なDiD設定では、通常 of 最小二乗法 (OLS) 係数 $\hat{\beta}$ が $\hat{\tau}_2$ と同等であることを簡単に示せる。

3. 処置の割り当てと時期に関する仮定の緩和

- 最近のDiDの論文では、処置の割り当てとタイミングに関する基本的な仮定の緩和に焦点が当てられている。
- 特に、2つ以上の期間があり、各個体が異なる時点で処置を受ける場合について大きな進展があった。
 - 例えば、異なる州で異なる時点で特定の種類の法律が制定される。

3.1. 時差のある処置を考慮した一般化モデル

仮定と表記法 (1) 処置のタイミング

- $t = 1, \dots, T$ で示される T 個の期間があり、各個体はどの期間でも関心のあるバイナリー処置を受けることができる。
- 処置は永続的に続く状態であり、一度処置を受けた個体はパネルの残りの期間、処置を受け続ける。
- ここで、個体 i が期間 t に処置を受けたかどうかの指標を D_{it} とし、 $G_i = \min \{t : D_{it} = 1\}$ は個体 i が処置を受けた最も早い期間である。
- もし i がサンプル期間中に一度も処置を受けなかった場合、 $G_i = \infty$ となる。
- 処置は永続的に続く状態であるため、すべての $t \geq G_i$ に対して $D_{it} = 1$ となる。処置に対する相対時間を $R_{it} = t - G_i + 1$ と定義し、個体 i の最初の処置期間では $R_{it} = 1$ となるようにする。

仮定と表記法（2）潜在的な結果

- 潜在的な結果の枠組みを多期間設定に拡張する。
- ここで、 $\mathbf{0}_s$ と $\mathbf{1}_s$ は、それぞれ0と1の s 次元ベクトルを表すとする。
- 個体 i が期間 g に初回処置を受けた場合の期間 t における潜在的な結果を $Y_{it}(\mathbf{0}_{g-1}, \mathbf{1}_{T-g+1})$ と表記する。
- “never-treated” の潜在的な結果を $Y_{it}(\mathbf{0}_T)$ と表記する。
- 我々は、処置がいったんオンになると「そのまま」であると仮定しているので、潜在的な結果の全経路は最初の処置時点 (g) によって要約される。したがって表記法を単純化するために、処置開始時期によって潜在的な結果をインデックス化することができる。
 - $Y_{it}(g) = Y_{it}(\mathbf{0}_{g-1}, \mathbf{1}_{T-g+1})$
 - $Y_{it}(\infty) = Y_{it}(\mathbf{0}_T)$

仮定4「時差あり設定の平行傾向」

- 平行傾向の仮定を時差あり設定に拡張する。
- 2群2期間版の平行傾向を、期間のすべての組み合わせと異なる時期に処置された「群」(groups)のすべての組み合わせで成立することを要求するように拡張。
- 仮定4 (時差あり設定の平行傾向)。全ての $t \neq t'$ と $g \neq g'$ について、
$$\mathbb{E}[Y_{it}(\infty) - Y_{it'}(\infty) \mid G_i = g] = \mathbb{E}[Y_{it}(\infty) - Y_{it'}(\infty) \mid G_i = g'] \quad (4)$$
- この仮定は、処置が行われなかった反実仮想では、すべての処置グループの結果が平行して進展していたことを意味している。
- (最終的にすべての個体が処置を受ける場合、処置される個体間でのみ平行傾向を課すと、ATTが識別できる期間の数は制限される。)

仮定5「時差ありの場合の予見効果なしの仮定」

- 基本モデルの予見効果なしの仮定は、時差のある設定にも簡単に適用できる。
- 直観的には、ある個体が期間 t に未処置であれば、その結果は、将来どの期間に処置されるかに依存しないことを課すものである。
- 仮定5 時差あり予見効果なしの仮定 (Staggered no anticipation assumption)。

$$Y_{it}(g) = Y_{it}(\infty) \text{ for all } t < g$$

- 意味：処置前の期間 ($t < g$) の結果は、将来期間 g に処置が行われることに依存しない。

3.2. TWFEモデルの推定対象の解釈

静学的TWFE

- まず、**静学的TWFEの定式化** (“static” TWFE specification) の議論から始める。これは、個体と期間の固定効果、および個体 i が期間 t で処置されたかどうかの指標 D_{it} に結果変数 Y_{it} を回帰させるものである。

$$Y_{it} = \alpha_i + \phi_t + D_{it}\beta_{\text{post}} + \epsilon_{it} \quad (5)$$

- 静学的定式化は、時間や個体によらず処置効果に不均質性がない場合に、賢明な推定対象 (sensible estimand) をもたらす。
- 処置効果の不均一性が、処置後の時間または個体間のいずれにも存在する場合**、静学的定式化に問題が発生する。
- 例えば、個々の処置効果 τ_s が全て正であっても、集計された処置効果の係数 β_{post} が負となる可能性 (negative weights) がある。 (**負の重みの可能性**)

Goodman-Bacon (2021) 分解

- Goodman-Bacon (2021) は、 $\hat{\beta}_{\text{post}}$ が、一方の個体が処置状態を変更し、他方が変更しなかった個体のペアと期間の間のDiD比較の凸加重平均 (convex weighted average) として分解できることを示した。
- 直感に反して、この分解により、より早い時期に処置を受けた個体を対照群("control" group)として用いるDiDを含んでいることが分かった。
- 初期に処置された個体 (early-treated unit) が、後に処置された多くの個体 (later-treated units) の「対照群」として現れると、負のウェイトを受ける。
- この分解により、処置効果が個体や時間によって異なる場合、 β_{post} は「禁じられた比較」(forbidden comparisons) を含むため、賢明な推定対象 (sensible estimand) ではない可能性があることがわかる。

動学的TWFE

- 動学的TWFEでは、個体と期間の固定効果、および処置時点からの相対時間（time relative to treatment）のダミーに結果を回帰させる。

$$Y_{it} = \alpha_i + \phi_t + \sum_{\substack{r \neq 0 \\ -\bar{T} \leq r \leq \bar{T}}} 1 [R_{it} = r] \beta_r + \epsilon_{it} \quad (7)$$

- ここで、 $R_{i,t} = t - G_i + 1$ であり、処置を受けてからの経過時間をあらわす。
- また、 $R_{i,t} = 1$ は処置を受けた最初の期間を示す。
- 処置後の時間のみに異質性がある場合**、賢明な因果推定対象（sensible causal estimand）を得ることができる。
- コホート間の異質性がある場合**、動学的定式化でも動的因果効果の賢明な推定値（sensible estimates of dynamic causal effects）を得ることができない。（この結果の導出は数学的に複雑）

3.3. 処置タイミングに時差がある場合のDiD推定量

CS推定量 (1) 概要

- Callaway and Sant'Anna (2021) のアプローチでは、時間 g に初めて処置されたコホートに対する時間 t における平均処置効果 (group-time average treatment effect on the treated) 、つまり、
$$ATT(g, t) = \mathbb{E} [Y_{it}(g) - Y_{it}(\infty) \mid G_i = g]$$
を構築要素 (building block) として考慮する。
- そして、仮定4やその変形のような平行傾向の仮定を一般化した下での識別と推定を検討する。
- 直感的には、仮定4 (平行傾向) と仮定5 (予見効果なし) の下では、 $ATT(g, t)$ は、コホート g の期間 $g - 1$ と t の結果の期待変化とまだ処置を受けていない対照群のそれを比較することにより、識別することができる。

CS推定量 (2) ATT

- Callaway and Sant'Anna (**2021**) は、平行傾向の仮定が共変量の条件付きでのみ成立する一般化についても考察しているが、ここでは共変量がない場合に焦点を当てる。
- 数学的には、任意の $g' > t$ (つまり、群 g' は t 時点で未処置) について、以下を式(2)の識別結果の多期間版と見なすことができる。

$$ATT(g, t) = \underbrace{\mathbb{E} [Y_{it} - Y_{i,g-1} \mid G_i = g]}_{\text{既に処置済みの群の期待変化}} - \underbrace{\mathbb{E} [Y_{it} - Y_{i,g-1} \mid G_i = g']}_{\text{未処置群の期待変化}}$$

- これは、 $g' > t$ となるどの比較群でも成り立つので、すべての $g' \in \mathcal{G}_{\text{comp}}$ について $g' > t$ となるような比較群の集合 $\mathcal{G}_{\text{comp}}$ を平均しても、成り立つ。

$$ATT(g, t) = \underbrace{\mathbb{E} [Y_{it} - Y_{i,g-1} \mid G_i = g]}_{\text{既に処置済みの群の期待変化}} - \underbrace{\mathbb{E} [Y_{it} - Y_{i,g-1} \mid G_i \in \mathcal{G}_{\text{comp}}]}_{\text{未処置群の期待変化}}$$

CS推定量 (3) 推定

- 期待値を標本平均に置き換えることで、 $ATT(g, t)$ を推定できる。

$$\widehat{ATT}(g, t) = \underbrace{\frac{1}{N_g} \sum_{i: G_i = g} [Y_{i,t} - Y_{i,g-1}]}_{\text{処置群の変化の平均}} - \underbrace{\frac{1}{N_{\mathcal{G}_{comp}}} \sum_{i: G_i \in \mathcal{G}_{comp}} [Y_{i,t} - Y_{i,g-1}]}_{\text{対照群の変化の平均}}$$

CS推定量 (4) 比較群

- Callaway and Sant'Anna (2021) は、比較群 $\mathcal{G}_{\text{comp}}$ について2つの選択肢を検討している。
 1. 一度も処置されることのない個体 (never-treated units) だけを使う。つまり、 $\mathcal{G}_{\text{comp}} = \{\infty\}$ 。
 2. その時点で未処置の個体 (not-yet-treated units) をすべて使用する。つまり、 $\mathcal{G}_{\text{comp}} = \{g' : g' > t\}$ 。

CS推定量 (5) 集計

- 期間と処置コホートの数が比較的少ない場合、関連するすべての (g, t) について $\widehat{ATT}(g, t)$ を報告することは合理的かもしれない。
- しかし、処置期間やコホートが多い場合、すべての $\widehat{ATT}(g, t)$ を報告することは煩雑である。
- 幸い、 $ATT(g, t)$ の任意の加重平均を推定することができる。

CS推定量 (6) 集計ATT

- 例えば、異なる採用（処置）コホート間の採用（処置）後 l 期間の処置効果の（加重）平均を与える「イベント・スタディ・パラメータ」（“event-study” parameter）を推定することができる。

$$ATT_l^w = \sum_g w_g ATT(g, g + l)$$

- ここで、 w_g は、コホート g の重みを表す。
- 重み w_g は、異なるコホートを等しく重み付けするように選ぶことも、処置された集団における相対頻度の観点から選ぶこともできる。

4. 平行傾向の仮定の緩和

仮定6 「条件付き平行傾向」

- 基本モデルの平行傾向の仮定は次のように共変量 (covariates) を組み込んだ形に拡張できる。

仮定6 (条件付き平行傾向/Conditional Parallel Trends)

$$\mathbb{E}[Y_{i2}(0) - Y_{i1}(0) \mid D_i = 1, X_i] = \mathbb{E}[Y_{i2}(0) - Y_{i1}(0) \mid D_i = 0, X_i] \text{ (almost surely)} \quad (10)$$

- ここで、 X_i は、観測された共変量の処置前のベクトルである。

仮定7「重複条件の仮定」

- 条件付き平行傾向の仮定に加えて、共変量 X_i を持つ各処置群個体に対して、 X_i の値が同じである未処置群個体が母集団中に少なくともいくつか存在することを保証する重複条件（overlap condition）、別名、正值条件（positivity condition）を課することになる。
- この重複条件の仮定は、標準的な推論手法を用いる場合に特に重要である(Khan and Tamer, 2010)。

仮定7（強い重複/Strong overlap）

- 観測された特性が与えられたとき、処置群に属する条件付き確率は、1から離れるように一様に束縛され、処置される個体の割合は0から離れるように束縛される。
- つまり、ある $\epsilon > 0$ について、
 $P(D_i = 1 \mid X_i) < 1 - \epsilon$, almost surely, and $\mathbb{E}[D] > 0$.

共変量で条件づけたATT

- 条件付き平行傾向の仮定、予見効果なしの仮定、重複条件の仮定のもとで、共変量 $X_i = x$ を条件としたATTは、次のように表される。

$$\tau_2(x) = \mathbb{E} [Y_{i2}(1) - Y_{i2}(0) \mid D_i = 1, X_i = x]$$

このATTは、 $P(D_i = 1 \mid X_i = x) > 0$ の任意の x について識別される。

- 特に、条件付き平行傾向の仮定が成立する場合、ATTは次のように表される。

$$\tau_2(x) = \underbrace{\mathbb{E} [Y_{i2} - Y_{i1} \mid D_i = 1, X_i = x]}_{\text{Change for } D_i=1, X_i=x} - \underbrace{\mathbb{E} [Y_{i2} - Y_{i1} \mid D_i = 0, X_i = x]}_{\text{Change for } D_i=0, X_i=x}. \quad (11)$$

二重頑健推定量 (1)

- 上記のアウトカム回帰とIPWのアプローチは、結果モデルと傾向スコアモデルのどちらかが正しく定式化されていれば有効な「二重頑健」法 (“doubly-robust” (DR) methods) を形成するために組み合わせることも可能である。
- 具体的には、Sant’Anna and Zhao (2020)は、仮定2、6、7の下で、二重頑健推定量 (Doubly-robust estimators) を用いて、ATTが以下のように識別されることを示している。

$$\tau_2 = \mathbb{E} \left[\underbrace{\left(\frac{D_i}{\mathbb{E}[D_i]} - \frac{\frac{(1-D_i)p(X_i)}{1-p(X_i)}}{\mathbb{E}\left[\frac{(1-D_i)p(X_i)}{1-p(X_i)}\right]} \right)}_{\substack{\text{逆確率重み付け} \\ \text{inverse probability weighting}}} \underbrace{\left(Y_{i,2} - Y_{i,1} - \underbrace{\mathbb{E}[Y_{i,2} - Y_{i,1} \mid D_i = 0, X_i]}_{\substack{\text{比較群の期待変化} \\ \text{population outcome regression}}} \right)} \right] \quad (15)$$

...(15)式

二重頑健推定量 (2)

- 前回と同様に、傾向スコアとCEF (conditional expectation function) の両方の推定値 (estimates) を代入することでATTを推定することができる。

$$\hat{\tau}_2 = \frac{1}{N} \sum_{i=1}^N \left(\frac{D_i}{\frac{1}{N} \sum_{j=1}^N D_j} - \frac{\frac{(1-D_i)\hat{p}(X_i)}{1-\hat{p}(X_i)}}{\frac{1}{N} \sum_{j=1}^N \frac{(1-D_j)\hat{p}(X_j)}{1-\hat{p}(X_j)}} \right) (Y_{i2} - Y_{i1} - \hat{\mathbb{E}}[Y_{i2} - Y_{i1} | D_i = 0, X_i])$$

- 結果式 (outcome equation) と傾向スコアは、パラメトリック法、セミ / ノンパラメトリック法のいずれでもモデル化でき、これらのモデルのいずれかが正しく定式化されていれば、DR法は一般に一致性 (consistent) が保たれる。

7. 結論

- 本論文の主要な結論は、以下の通りである
 1. 研究者は識別に用いる比較グループを明確にすること、
 2. 推定・推論方法を識別の仮定に合わせること、
 3. 仮定に違反する可能性に対して頑健性を探る。
- また、正しい識別仮定やそれに付随する方法を選択するためには、文脈に応じた知識（context-specific knowledge）が必要となる。

References

Callaway, Brantly, and Pedro H. C. Sant'Anna. 2021. "Difference-in-Differences with Multiple Time Periods." *Journal of Econometrics* 225 (2): 200–230.

<https://doi.org/10.1016/j.jeconom.2020.12.001>.

Goodman-Bacon, Andrew. 2021. "Difference-in-Differences with Variation in Treatment Timing." *Journal of Econometrics* 225 (2): 254–77.

<https://doi.org/10.1016/j.jeconom.2021.03.014>.

Roth, Jonathan, Pedro HC Sant'Anna, Alyssa Bilinski, and John Poe. 2023. "What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature." *Journal of Econometrics* 235 (2): 2218–44.

<https://doi.org/10.1016/j.jeconom.2023.03.008>.