Double-Robust Two-Way-Fixed-Effects Regression For Panel Data

Lihua Lei

Graduate School of Business, Stanford University

2023 ASSA/AEA Annual Meeting

Collaborators

Dmitry Arkhangelsky
(CEMFI)

Guido Imbens
(Stanford)

Xiaoman Luo
(Afiniti)

Two-way fixed effect (TWFE) regression model and estimator

$$\text{TWFE model}: \underbrace{Y_{it}}_{\text{outcome}} = \underbrace{\alpha_i}_{\text{unit FE}} + \underbrace{\lambda_t}_{\text{time FE}} + \underbrace{\tau}_{\text{effect}} \cdot \underbrace{W_{it}}_{\text{treatment}} + \beta \cdot \underbrace{X_{it}}_{\text{covariates}} + \epsilon_{it}$$

$$\text{TWFE estimator}: \hat{\tau}_{\text{TWFE}} \leftarrow \texttt{OLS}(Y_{it} \sim \text{unit dummy} + \text{time dummy} + W_{it} + X_{it})$$

Two-way fixed effect (TWFE) regression model and estimator

$$\text{TWFE model}: \underbrace{Y_{it}}_{\text{outcome}} = \underbrace{\alpha_i}_{\text{unit FE}} + \underbrace{\lambda_t}_{\text{time FE}} + \underbrace{\tau}_{\text{effect}} \cdot \underbrace{W_{it}}_{\text{treatment}} + \beta \cdot \underbrace{X_{it}}_{\text{covariates}} + \epsilon_{it}$$

$$\text{TWFE estimator}: \hat{\tau}_{\text{TWFE}} \leftarrow \texttt{OLS}(Y_{it} \sim \text{unit dummy} + \text{time dummy} + W_{it} + X_{it})$$

- DiD estimator $\Longleftrightarrow$ TWFE (with $T = 2$)
- $\hat{\tau}_{\text{TWFE}}$ is unbiased for τ under the TWFE model
- Biased with heterogeneous treatment effect or violation of parallel trend Borusyak et al '17, Goodman-Bacon '17, de Chaisemartin and d'Haultfoeuille '18, Athey and Imbens '18, Sun and Abraham '18
- Many alternative methods recently Imai and Kim '16, Athey et al. '17, Borusyak et al. '17, Callaway and Sant'Anna '18, de Chaisemartin and d'Haultfoeuille '18, Sun and Abraham '18, Arkhangelsky and Imbens '19, Arkhangelsky et al. '19, Ben-Michael et al. '19, Roth and Sant'Anna '20, ...

This paper

Has TWFE been fully understood?

This paper

Has TWFE been fully understood? NO!

This paper

Has TWFE been fully understood? NO!

- A class of estimands: **doubly average treatment effects (DATE)**
- A new estimator: **reshaped inverse probability weighting (RIPW)**-TWFE estimator
- Valid design-based inference:
  - **time- and unit-varying** effects (finite population framework)
  - many **dependent** designs: sampling without replacement, two-stage randomization, ...
- **Double robustness**: RIPW $\overset{p}{\to}$ DATE if
  - either the treatment assignment model is known/well estimated
  - or the TWFE model is correct
- Not limited to staggered adoption

Part I: DATE, RIPW, and design-based inference

Potential outcomes and doubly average treatment effect (DATE)

- Balanced panel: n units and T time periods; fixed T (harder than large T, discussed later)
- Binary treatment: $\mathbf{W}_i = (W_{i1}, \ldots, W_{iT})$; $\mathbf{W}_i \sim \boldsymbol{\pi}_i$ generalized propensity score Imben '00
- Potential outcomes: $(Y_{it}(1), Y_{it}(0))_{t=1}^T$; observed outcome $Y_{it} = Y_{it}(W_{it})$ (SUTVA)
- No covariates for this part (just for simplicity)
- Causal estimand: DATE with **user-specified** weights $\xi = (\xi_1, \ldots, \xi_t)$

$$\tau_{\text{DATE}}(\xi) = \sum_{t=1}^{T} \xi_t \left(\frac{1}{n} \sum_{i=1}^{n} (Y_{it}(1) - Y_{it}(0)) \right) \triangleq \sum_{t=1}^{T} \xi_t \tau_t, \quad \text{e.g., } \tau_{\text{eq}} = \frac{1}{T} \sum_{t=1}^{T} \tau_t$$

Potential outcomes and doubly average treatment effect (DATE)

- Balanced panel: n units and T time periods; fixed T (harder than large T, discussed later)
- Binary treatment: $\mathbf{W}_i = (W_{i1}, \ldots, W_{iT})$; $\mathbf{W}_i \sim \boldsymbol{\pi}_i$ generalized propensity score Imben '00
- Potential outcomes: $(Y_{it}(1), Y_{it}(0))_{t=1}^{T}$; observed outcome $Y_{it} = Y_{it}(W_{it})$ (SUTVA)
- No covariates for this part (just for simplicity)
- Causal estimand: DATE with **user-specified** weights $\xi = (\xi_1, \ldots, \xi_t)$

$$\tau_{\mathrm{DATE}}(\xi) = \sum_{t=1}^{T} \xi_t \left(\frac{1}{n} \sum_{i=1}^{n} (Y_{it}(1) - Y_{it}(0)) \right) \triangleq \sum_{t=1}^{T} \xi_t \tau_t, \quad \text{e.g., } \tau_{\mathrm{eq}} = \frac{1}{T} \sum_{t=1}^{T} \tau_t$$

How to leverage the treatment assignment mechanism to estimate DATE?

First thought: IPW estimator

For cross-sectional data, the Hájek-IPW estimator is given by

$$\hat{\tau} = \frac{\sum_{W_i=1} Y_i/\mathbb{P}(W_i = 1)}{\sum_{W_i=1} 1/\mathbb{P}(W_i = 1)} - \frac{\sum_{W_i=0} Y_i/\mathbb{P}(W_i = 0)}{\sum_{W_i=0} 1/\mathbb{P}(W_i = 0)} \xrightarrow{p} \mathrm{ATE}$$

First thought: IPW estimator

For cross-sectional data, the Hájek-IPW estimator is given by

$$\hat{\tau} = \frac{\sum_{W_i=1} Y_i/\mathbb{P}(W_i=1)}{\sum_{W_i=1} 1/\mathbb{P}(W_i=1)} - \frac{\sum_{W_i=0} Y_i/\mathbb{P}(W_i=0)}{\sum_{W_i=0} 1/\mathbb{P}(W_i=0)} \xrightarrow{p} \mathrm{ATE}$$

Numerically equivalent to an IP-weighted LS estimator:

$$\hat{\tau} \triangleq \arg\min_{\tau} \sum_{i=1}^{n} \underbrace{(Y_i - \mu - W_i\tau)^2}_{\text{least squares objective}} \quad \underbrace{\frac{1}{\pi_i(W_i)}}_{\text{propensity score}}$$

Key idea: reweighting the objective function via the treatment assignment mechanism

First thought: IPW estimator

Key idea: reweighting the objective function via the treatment assignment mechanism

Analogue in the panel data:

$$\hat{\tau}_{\text{IPW}} \triangleq \arg\min_{\tau} \sum_{i=1}^{n} \sum_{t=1}^{T} \underbrace{(Y_{it} - \alpha_i - \lambda_t - W_{it}\tau)^2}_{\text{TWFE objective}} \quad \frac{1}{\underbrace{\pi_i(\mathbf{W}_i)}_{\text{generalized propensity score}}}$$

First thought: IPW estimator

Key idea: reweighting the objective function via the treatment assignment mechanism

Analogue in the panel data:

$$\hat{\tau}_{\mathrm{IPW}} \triangleq \arg\min_{\tau} \sum_{i=1}^{n} \sum_{t=1}^{T} \underbrace{(Y_{it} - \alpha_i - \lambda_t - W_{it}\tau)^2}_{\text{TWFE objective}} \frac{1}{\underbrace{\pi_i(\mathbf{W}_i)}_{\text{generalized propensity score}}} \xrightarrow{p} \;\; \textbf{?}$$

Two examples with 3 time periods

Transient treatments

$$\mathbf{W}_i \in \{(0,0,0),(0,0,1),(0,1,0),(1,0,0)\}$$

Two examples with 3 time periods

Transient treatments

$$\mathbf{W}_i \in \{(0,0,0),(0,0,1),(0,1,0),(1,0,0)\}$$

$$\hat{\tau}_{\text{IPW}} \xrightarrow{p} \frac{1}{3}\tau_1 + \frac{1}{3}\tau_2 + \frac{1}{3}\tau_3 = \tau_{\text{eq}}$$

Two examples with 3 time periods

Transient treatments

$$\mathbf{W}_i \in \{(0,0,0),(0,0,1),(0,1,0),(1,0,0)\}$$

$$\hat{\tau}_{\text{IPW}} \xrightarrow{p} \frac{1}{3}\tau_1 + \frac{1}{3}\tau_2 + \frac{1}{3}\tau_3 = \tau_{\text{eq}}$$

Staggered rollouts

$$\mathbf{W}_i \in \{(0,0,0),(0,0,1),(0,1,1),(1,1,1)\}$$

Two examples with 3 time periods

Transient treatments

$$\mathbf{W}_i \in \{(0,0,0),(0,0,1),(0,1,0),(1,0,0)\}$$

$$\hat{\tau}_{\text{IPW}} \xrightarrow{p} \frac{1}{3}\tau_1 + \frac{1}{3}\tau_2 + \frac{1}{3}\tau_3 = \tau_{\text{eq}}$$

Staggered rollouts

$$\mathbf{W}_i \in \{(0,0,0),(0,0,1),(0,1,1),(1,1,1)\}$$

$$\hat{\tau}_{\text{IPW}} \xrightarrow{p} 0.3\tau_1 + 0.4\tau_2 + 0.3\tau_3$$

Two examples with T time periods

Transient treatments

$$W_{i1} + W_{i2} + \ldots + W_{iT} \leq 1$$

$$\hat{\tau}_{\text{IPW}} \xrightarrow{p} \frac{1}{T} \sum_{t=1}^{T} \tau_t = \tau_{\text{eq}}$$

Staggered rollouts

$$W_{i1} \leq W_{i2} \leq \ldots \leq W_{iT}$$

$$\hat{\tau}_{\text{IPW}} \xrightarrow{p} \sum_{t=1}^{T} \frac{(T+1-t)t}{\sum_{t=1}^{T}(T+1-t)t} \tau_t$$

Two examples with T time periods

Transient treatments

$$W_{i1} + W_{i2} + \ldots + W_{iT} \leq 1$$

Staggered rollouts

$$W_{i1} \leq W_{i2} \leq \ldots \leq W_{iT}$$

$$\hat{\tau}_{\text{IPW}} \xrightarrow{p} \frac{1}{T} \sum_{t=1}^{T} \tau_t = \tau_{\text{eq}}$$

$$\hat{\tau}_{\text{IPW}} \xrightarrow{p} \sum_{t=1}^{T} \frac{(T+1-t)t}{\sum_{t=1}^{T}(T+1-t)t} \tau_t$$

What if we want DATE with pre-specified weights (e.g., τ_{eq})?

Reshaped IPW estimator

Given a data-independent distribution $\mathbf{\Pi}$ on $\mathbb{S}$:

$$\text{RIPW estimator: } \hat{\tau}_{\text{RIPW}}(\mathbf{\Pi}) \triangleq \arg\min_{\tau} \sum_{i=1}^{n} \sum_{t=1}^{T} (Y_{it} - \alpha_i - \lambda_t - W_{it}\tau)^2 \frac{\mathbf{\Pi}(\mathbf{W}_i)}{\boldsymbol{\pi}_i(\mathbf{W}_i)}$$

Reshaped IPW estimator

Given a data-independent distribution $\mathbf{\Pi}$ on $\mathbb{S}$:

$$\text{RIPW estimator: } \hat{\tau}_{\text{RIPW}}(\mathbf{\Pi}) \triangleq \arg\min_{\tau} \sum_{i=1}^{n} \sum_{t=1}^{T} (Y_{it} - \alpha_i - \lambda_t - W_{it}\tau)^2 \frac{\mathbf{\Pi}(\mathbf{W}_i)}{\boldsymbol{\pi}_i(\mathbf{W}_i)}$$

- The IPW-TWFE estimator is the RIPW-TWFE estimator with $\mathbf{\Pi} \sim \mathrm{Unif}(\mathbb{S})$
- When $\boldsymbol{\pi}_i = \mathbf{\Pi}$, the RIPW-TWFE estimator reduces to the TWFE estimator

Reshaped IPW estimator

Given a data-independent distribution $\mathbf{\Pi}$ on $\mathbb{S}$:

$$\text{RIPW estimator: } \hat{\tau}_{\text{RIPW}}(\mathbf{\Pi}) \triangleq \arg\min_{\tau} \sum_{i=1}^{n} \sum_{t=1}^{T} (Y_{it} - \alpha_i - \lambda_t - W_{it}\tau)^2 \frac{\mathbf{\Pi}(\mathbf{W}_i)}{\boldsymbol{\pi}_i(\mathbf{W}_i)}$$

- The IPW-TWFE estimator is the RIPW-TWFE estimator with $\mathbf{\Pi} \sim \mathrm{Unif}(\mathbb{S})$
- When $\boldsymbol{\pi}_i = \mathbf{\Pi}$, the RIPW-TWFE estimator reduces to the TWFE estimator

For what $\mathbf{\Pi}$ does $\hat{\tau}_{\text{RIPW}}(\mathbf{\Pi}) \xrightarrow{p} \tau_{\text{DATE}}(\xi)$?

DATE equation

Theorem (Arkhangelsky, Imbens, L., and Luo '21)

Given $\mathbb{S}$ *and* $\mathbf{\Pi}$ *with* $\mathrm{Supp}(\mathbf{\Pi}) = \mathbb{S}$, $\hat{\tau}_{TWFE} \xrightarrow{p} \tau_{DATE}(\xi)$ *if and "only if"*

$$\mathbb{E}_{\mathbf{W}\sim\mathbf{\Pi}}\left[(\mathrm{diag}(\mathbf{W}) - \xi\mathbf{W}^\top)J(\mathbf{W} - \mathbb{E}_{\mathbf{W}\sim\mathbf{\Pi}}[\mathbf{W}])\right] = 0 \quad (\textit{DATE equation}),$$

where $J = I - \mathbf{1}_T\mathbf{1}_T^\top/T$.

- Only depends on $\mathbb{S}$
- Quadratic equations on $(\mathbf{\Pi}(w) : w \in \mathbb{S})$ with linear constraints (simplex, positivity)
- Closed-form solutions exist in many examples (DiD, cross-over, staggered rollouts, transient, ...)
- Generic solver based on nonlinear programming (BFGS algorithm)

Solutions of DATE equation: examples with estimand τ_{eq}

Transient treatments

$$\mathbf{W}_i \in \{(0,0,0),(0,0,1),(0,1,0),(1,0,0)\}$$

$$(\boldsymbol{\Pi}(0,0,0),\boldsymbol{\Pi}(0,0,1),\boldsymbol{\Pi}(0,1,0),\boldsymbol{\Pi}(1,0,0)) = \lambda\cdot(1,0,0,0)+(1-\lambda)\cdot\left(0,\frac{1}{3},\frac{1}{3},\frac{1}{3}\right)$$

$\lambda \in (0,1)$, Unif is a solution

Staggered rollouts

$$\mathbf{W}_i \in \{(0,0,0),(0,0,1),(0,1,1),(1,1,1)\}$$

$$(\boldsymbol{\Pi}(0,0,0),\boldsymbol{\Pi}(0,0,1),\boldsymbol{\Pi}(0,1,1),\boldsymbol{\Pi}(1,1,1)) = \lambda\cdot\left(\frac{2}{9},\frac{1}{3},0,\frac{4}{9}\right)+(1-\lambda)\cdot\left(\frac{4}{9},0,\frac{1}{3},\frac{2}{9}\right)$$

$\lambda \in (0,1)$, Unif is NOT a solution

An interpretation of DATE equation

- When $\boldsymbol{\pi}_i = \mathbf{\Pi}$, $\hat{\tau}_{\text{TWFE}} = \hat{\tau}_{\text{RIPW}}(\mathbf{\Pi}) \xrightarrow{p} \tau_{\text{DATE}}(\xi)$

- DATE equation gives all completely randomized experiments for which TWFE "works"!

An interpretation of DATE equation

- When $\boldsymbol{\pi}_i = \mathbf{\Pi}$, $\hat{\tau}_{\text{TWFE}} = \hat{\tau}_{\text{RIPW}}(\mathbf{\Pi}) \xrightarrow{p} \tau_{\text{DATE}}(\xi)$
- DATE equation gives all completely randomized experiments for which TWFE "works"!
- Conflict with the literature that TWFE has negative weights?

An interpretation of DATE equation

- When $\boldsymbol{\pi}_i = \boldsymbol{\Pi}$, $\hat{\tau}_{\mathsf{TWFE}} = \hat{\tau}_{\mathsf{RIPW}}(\boldsymbol{\Pi}) \xrightarrow{p} \tau_{\mathsf{DATE}}(\xi)$
- DATE equation gives all completely randomized experiments for which TWFE "works"!
- Conflict with the literature that TWFE has negative weights?
- Not really! $\mathbf{W}_i$'s are treated as fixed in the literature but as random in our work
- When talking about "weights", important to specify the sources of randomness

Design-based inference for RIPW estimator

$$\text{RIPW estimator: } \hat{\tau}(\mathbf{\Pi}) \triangleq \arg\min_{\tau} \sum_{i=1}^{n} \sum_{t=1}^{T} (Y_{it} - \alpha_i - \lambda_t - W_{it}\tau)^2 \frac{\mathbf{\Pi}(\mathbf{W}_i)}{\boldsymbol{\pi}_i(\mathbf{W}_i)}$$

- We propose an (asymptotically) conservative variance estimator and valid Wald CI under
  - Bernoulli design (independent $\mathbf{W}_i$'s)
  - Sampling without replacement
  - Cluster-wise randomization
  - Two-stage randomization
  - ...
- Roughly speaking, valid for **dependent designs that can be handled for cross-sectional data**

Part II: Double robustness of RIPW estimator

RIPW estimators with covariates

- Covariates: $\mathbf{X}_i = (X_{i1}, \ldots, X_{iT})$ (satisfying a **latent ignorability assumption**)
- Use $\mathbf{X}_i$ to fit an assignment model $\hat{\pi}_i(\cdot)$:
  - Staggered rollouts: duration models (e.g., Cox proportional hazard model)
  - General designs: discrete Markov model, conditional logit model ...

RIPW estimators with covariates

- Covariates: $\mathbf{X}_i = (X_{i1}, \ldots, X_{iT})$ (satisfying a **latent ignorability assumption**)
- Use $\mathbf{X}_i$ to fit an assignment model $\hat{\boldsymbol{\pi}}_i(\cdot)$:
  - Staggered rollouts: duration models (e.g., Cox proportional hazard model)
  - General designs: discrete Markov model, conditional logit model ...
- Use $\mathbf{X}_i$ to fit an outcome model $\hat{\mathbf{m}}_i = (\hat{m}_{i1}, \ldots, \hat{m}_{iT})$ for effects varying with units and time
  - Under TWFE $Y_{it} = \alpha_i + \lambda_t + m_{it} + \epsilon_{it}$ where $m_{it} = X_{it}^\top \beta$, then $\hat{m}_{it} = X_{it}^\top \hat{\beta}_{\text{TWFE}}$
  - No need to estimate FE; crucial since α_i **cannot be consistently estimated** for fixed T

RIPW estimators with covariates

- Covariates: $\mathbf{X}_i = (X_{i1}, \ldots, X_{iT})$ (satisfying a **latent ignorability assumption**)
- Use $\mathbf{X}_i$ to fit an assignment model $\hat{\pi}_i(\cdot)$:
  - Staggered rollouts: duration models (e.g., Cox proportional hazard model)
  - General designs: discrete Markov model, conditional logit model ...
- Use $\mathbf{X}_i$ to fit an outcome model $\hat{\mathbf{m}}_i = (\hat{m}_{i1}, \ldots, \hat{m}_{iT})$ for effects varying with units and time
  - Under TWFE $Y_{it} = \alpha_i + \lambda_t + m_{it} + \epsilon_{it}$ where $m_{it} = X_{it}^\top \beta$, then $\hat{m}_{it} = X_{it}^\top \hat{\beta}_{\text{TWFE}}$
  - No need to estimate FE; crucial since α_i **cannot be consistently estimated** for fixed T

$$\hat{\tau}(\boldsymbol{\Pi}) \triangleq \arg\min_{\tau} \sum_{i=1}^{n} \sum_{t=1}^{T} \Big(\underbrace{(Y_{it} - \hat{m}_{it})}_{\text{modified outcome}} - \alpha_i - \lambda_t - W_{it}\tau \Big)^2 \frac{\boldsymbol{\Pi}(\mathbf{W}_i)}{\hat{\pi}_i(\mathbf{W}_i)}$$

RIPW estimator is double robust for observational studies

$$\hat{\tau}(\boldsymbol{\Pi}) \triangleq \arg\min_{\tau} \sum_{i=1}^{n} \sum_{t=1}^{T} (\underbrace{(Y_{it} - \hat{m}_{it})}_{\text{regression adjustment}} - \alpha_i - \lambda_t - W_{it}\tau)^2 \frac{\boldsymbol{\Pi}(\mathbf{W}_i)}{\underbrace{\hat{\pi}_i(\mathbf{W}_i)}_{\text{assignment modeling}}}$$

- **Double robustness**: RIPW$\xrightarrow{p}$DATE if
  - either the assignment model is well estimated
  - or the TWFE model is correct

RIPW estimator is double robust for observational studies

$$\hat{\tau}(\boldsymbol{\Pi}) \triangleq \arg\min_{\tau} \sum_{i=1}^{n} \sum_{t=1}^{T} (\underbrace{(Y_{it} - \hat{m}_{it})}_{\text{regression adjustment}} - \alpha_i - \lambda_t - W_{it}\tau)^2 \frac{\boldsymbol{\Pi}(\mathbf{W}_i)}{\underbrace{\hat{\pi}_i(\mathbf{W}_i)}_{\text{assignment modeling}}}$$

- **Double robustness**: RIPW$\xrightarrow{p}$DATE if
  - either the assignment model is well estimated
  - or the TWFE model is correct
- Fundamentally different from the double robustness discussed in Sant'Anna and Zhao ('20)
  - Based on fundamentally different assumptions
  - Our double robustness holds for non-staggered adoption

Part III: Case study

State of emergency in the early COVID-19 pandemic

Inslee issues COVID-19 emergency proclamation

February 29, 2020

Story

Gov. Jay Inslee today declared a state of emergency in response to new cases of COVID-19, directing state agencies to use all resources necessary to prepare for and respond to the outbreak.

A **state of emergency** is a situation in which a government is empowered to perform actions or impose policies that it would normally not be permitted to undertake

OpenTable data in the early COVID-19 pandemic

California	9%	2%	-2%	2%	3%	6%	3%	3%	4%
Colorado	-4%	-9%	-13%	-2%	3%	6%	-3%	-5%	-2%
Connecticut	57%	9%	12%	-4%	49%	36%	30%	30%	20%
Delaware	-3%	-12%	-15%	-33%	-17%	0%	-3%	-11%	-5%
District of Columbia	-26%	-30%	-29%	-15%	-15%	-13%	-20%	-24%	-28%
Florida	26%	17%	19%	21%	29%	20%	16%	26%	32%
Georgia	0%	-4%	-3%	-2%	5%	10%	5%	2%	12%
Hawaii	-5%	1%	-5%	-7%	-12%	-6%	-10%	-10%	-12%

Daily data of year-over-year seated diners for a sample of restaurants in 36 states in the US

How the state of emergency affects economic activities in short term

- Interested in ATE of the state of emergency on dine-in rate during 02/29 – 03/13, 2020
  - State of emergency was less confounded; the first policy affecting the vast majority of the public
  - Restaurant industry is responding to the policy swiftly, thus immune to long-term confounders

How the state of emergency affects economic activities in short term

- Interested in ATE of the state of emergency on dine-in rate during 02/29 – 03/13, 2020

- Declaration time (assignment model) is easier to model than the dine-in rate (outcome model)
  - Dine-in rate is driven by many unmeasured behavioral variables
  - Declaration time is mainly driven by the progress of the pandemic and the authority's attitude

How the state of emergency affects economic activities in short term

- Interested in ATE of the state of emergency on dine-in rate during 02/29 – 03/13, 2020
- Declaration time (assignment model) is easier to model than the dine-in rate (outcome model)
- Covariates:
  - State-level accumulated confirmed cases
  - The vote share of Democrats based on the 2016 presidential election data
  - Number of hospital beds per-capita

RIPW estimate

- For assignments, fit a Cox proportional hazard model
- For regression adjustment, fit a standard (unweighted) TWFE model
- Estimate: -4.0% (95% CI $[-8.6\%, 0.6\%]$, 90% CI $[-7.9\%, -0.1\%]$)
- Unweighted TWFE: -1.1% (95% CI $[-4.3\%, 2.1\%]$, 90% CI $[-3.8\%, 1.6\%]$)

Summary

- IPW-TWFE converges to a DATE with potentially uninterpretable weights
- **RIPW-TWFE** solves the problem:
  - permits valid design-based **inference** for most practical designs
  - **double-robust** and work for **general designs** (not limited to staggered rollouts)
- Practically, they empower the users to leverage the information from the assignment model
- Easily computed from any existing software that supports weighted TWFE
- Objective-reweighting is a powerful general strategy (e.g., combined with an event-study model)

Thank you!

Paper link: https://arxiv.org/abs/2107.13737

Appendix: discussion

Negative weighting

Literature: "$\hat{\tau}_{\text{TWFE}}$ does not converge to a convex combination of τ_{it}s for most designs"

Our work: $\hat{\tau}_{\text{TWFE}}$ converges to DATE for any design for which the DATE equation has a solution

Conflict?

Negative weighting

Literature: "$\hat{\tau}_{\text{TWFE}}$ does not converge to a convex combination of τ_{it}s for most designs"

Our work: $\hat{\tau}_{\text{TWFE}}$ converges to DATE for any design for which the DATE equation has a solution

Conflict? **NO! Different sources of randomness**

Negative weighting

- The weights discussed in the literature:

$$\underbrace{\mathbb{E}[\hat{\tau}_{\text{TWFE}} \mid \mathbf{W}]}_{\text{conditional estimand}} = \sum_{i=1}^{n} \sum_{t=1}^{T} \underbrace{\zeta_{it}(\mathbf{W})}_{\text{conditional weight}} \tau_{it}$$

Negative weighting

- The weights discussed in the literature:

$$\underbrace{\mathbb{E}[\hat{\tau}_{\text{TWFE}} \mid \mathbf{W}]}_{\text{conditional estimand}} = \sum_{i=1}^{n} \sum_{t=1}^{T} \underbrace{\zeta_{it}(\mathbf{W})}_{\text{conditional weight}} \tau_{it}$$

The result proved in the literature: for most designs (e.g., staggered adoption with $T > 2$)

$$\exists (i,t) : \zeta_{it}(\mathbf{W}) < 0, \quad \text{almost surely}$$

Negative weighting

- The weights discussed in the literature:

$$\underbrace{\mathbb{E}[\hat{\tau}_{\text{TWFE}} \mid \mathbf{W}]}_{\text{conditional estimand}} = \sum_{i=1}^{n} \sum_{t=1}^{T} \underbrace{\zeta_{it}(\mathbf{W})}_{\text{conditional weight}} \tau_{it}$$

  The result proved in the literature: for most designs (e.g., staggered adoption with $T > 2$)

$$\exists (i, t) : \zeta_{it}(\mathbf{W}) < 0, \quad \text{almost surely}$$

- The weights discussed in our work:

$$\underbrace{\mathbb{E}[\hat{\tau}_{\text{TWFE}}]}_{\text{unconditional estimand}} = \sum_{i=1}^{n} \sum_{t=1}^{T} \underbrace{\mathbb{E}[\zeta_{it}(\mathbf{W})]}_{\text{unconditional weight}} \tau_{it}$$

Negative weighting

- The weights discussed in the literature:

$$\underbrace{\mathbb{E}[\hat{\tau}_{\text{TWFE}} \mid \mathbf{W}]}_{\text{conditional estimand}} = \sum_{i=1}^{n} \sum_{t=1}^{T} \underbrace{\zeta_{it}(\mathbf{W})}_{\text{conditional weight}} \tau_{it}$$

The result proved in the literature: for most designs (e.g., staggered adoption with $T > 2$)

$$\exists (i, t) : \zeta_{it}(\mathbf{W}) < 0, \quad \text{almost surely}$$

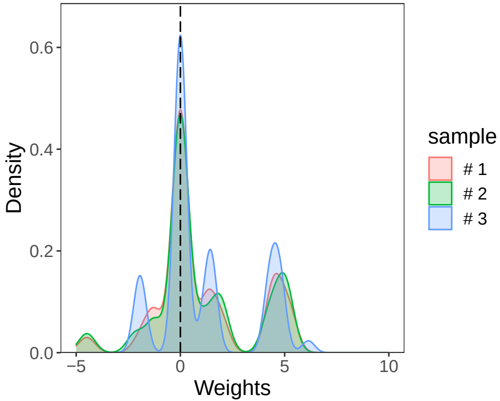
- The weights discussed in our work:

$$\underbrace{\mathbb{E}[\hat{\tau}_{\text{TWFE}}]}_{\text{unconditional estimand}} = \sum_{i=1}^{n} \sum_{t=1}^{T} \underbrace{\mathbb{E}[\zeta_{it}(\mathbf{W})]}_{\text{unconditional weight}} \tau_{it}$$

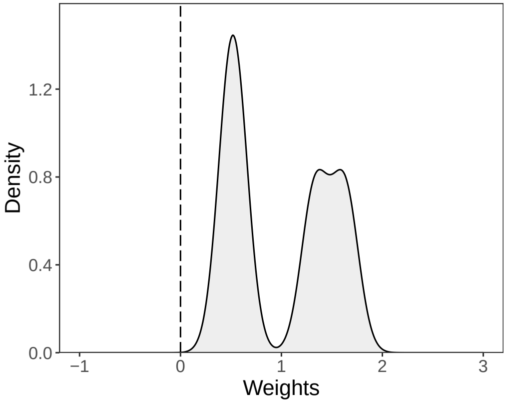
The result proved in our work: for any design for which the DATE equation has a solution:

$$\forall (i, t) : \mathbb{E}[\zeta_{it}(\mathbf{W})] > 0$$

Negative weighting: beyond RCT



Conditional weights $\mathbb{E}[\zeta_{it} \mid \mathbf{W}]$



Unconditional weights $\mathbb{E}[\zeta_{it}]$

Another view of DATE equation: effective estimand

$$\mathbb{E}_{\mathbf{W}\sim\mathbf{\Pi}}\left[(\mathrm{diag}(\mathbf{W}) - \zeta\mathbf{W}^\top)J(\mathbf{W} - \mathbb{E}_{\mathbf{W}\sim\mathbf{\Pi}}[\mathbf{W}])\right] = 0$$

$$\Longrightarrow \zeta = \frac{\mathbb{E}_{\mathbf{W}\sim\mathbf{\Pi}}\left[\mathrm{diag}(\mathbf{W})J(\mathbf{W} - \mathbb{E}_{\mathbf{W}\sim\mathbf{\Pi}}[\mathbf{W}])\right]}{\mathbb{E}_{\mathbf{W}\sim\mathbf{\Pi}}\left[\mathbf{W}^\top J(\mathbf{W} - \mathbb{E}_{\mathbf{W}\sim\mathbf{\Pi}}[\mathbf{W}])\right]} \quad (*)$$

$\Longrightarrow \tau(\zeta)$ is the effective estimand of the RIPW estimator

Theorem (Arkhangelsky, Imbens, L., and Luo '22+)

Let ζ be defined in $(*)$. For any $\mathbf{\Pi}$ and t, $\zeta_t \geq 0$.