Causal Inference Workshop

Week 6 - TWFE
Application and Implementation

Causal Inference Workshop

February 23, 2024

Workshop outline

A. Causal inference fundamentals
   - Modeling assumptions matter too
   - Conceptual framework (potential outcomes framework)

B. Design stage: common identification strategies
   - IV + RDD [coding]
   - Event Studies, DiD, DiDiD, New TWFE Lit [coding]
   - Synthetic Control / Synthetic DiD [coding]

C. Analysis stage: strengthening inferences
   - Limitations of identification strategies, pre-estimation steps
   - Estimation [controls] and post-estimation steps [supporting assumptions]

D. Other topics in causal inference and sustainable development
   - Inference (randomization inference, bootstrapping)
   - Weather data regressions, other common/fun SDev topics [coding]
   - Remote sensing data, other common/fun SDev topics

Causal inference roadmap

- *Potential outcomes* [framework]
  - Causal effect is the difference between two potential outcomes
  - We can't observe this difference, but can see differences in average observed outcomes
  - If **(conditional) independence assumption** holds, can estimate unbiased ATT
- *Identification* [application/implementation] [last week, and today, ... and next week!]
  - In most empirical settings, IA and CIA do not hold, which is why we need an **identification strategy**
  - Want to eliminate selection bias (identification problem)
- *Estimation* [application/implementation]
  - (Usually) use linear regression model
  - $\hat{\beta}_{OLS}$ unbiased estimator for ATT if e is uncorrelated with treatment (regression problem)

Outline

Workshop outline

Two-Way Fixed Effects (TWFE)

Outline

Workshop outline

Two-Way Fixed Effects (TWFE)
- Review of TWFE problem
- Suggestions for implementation
- Coding TWFE estimators

TWFE problem overview

- Big picture takeaway: TWFE regressions estimate *weighted sums* of the ATE in each group and period, with weights that sum to one but may make TWFE estimator biased (and *may* even be negative)

$$y_{it} = \lambda_{g[i]} + \delta_t + \beta_{TWFE} D_{g[i]t} + e_{g[i]t}$$

$\rightarrow$ $\hat{\beta}_{TWFE}$ is a specific weighted sum of the ATE in each treated (g,t) cell (Chaisemartin and D'Haultfœuille 2020)

$$\beta_{ATT} = \mathbb{E}\left[\sum_{(gt):D_{gt}=1} \frac{N_{gt}}{N_1} ATE_{gt} \right], \qquad \mathbb{E}[\hat{\beta}_{TWFE}] = \mathbb{E}\left[\sum_{(gt):D_{gt}=1} \frac{N_{gt}}{N_1} w_{gt} ATE_{gt} \right]$$

$\rightarrow$ weights may vary and can even be **negative!** (Goodman-Bacon 2021)
negative weights huge potential issue, as it may be $\hat{\beta}_{TWFE} < 0$ even if all $ATE_{gt} > 0$

More on TWFE weights

- TWFE regressions estimate *weighted sums* of the ATE in each group and period

$$y_{it} = \lambda_{g[i]} + \delta_t + \beta_{TWFE} D_{g[i]t} + e_{g[i]t}$$

- $\hat{\beta}_{TWFE}$ is a specific weighted sum of the ATE in each treated (g,t) cell (Chaisemartin and D'Haultfœuille 2020)

$$\beta_{ATT} = \mathbb{E}\left[\sum_{(gt):D_{gt}=1} \frac{N_{gt}}{N_1} ATE_{gt}\right], \qquad \mathbb{E}[\hat{\beta}_{TWFE}] = \mathbb{E}\left[\sum_{(gt):D_{gt}=1} \frac{N_{gt}}{N_1} w_{gt} ATE_{gt}\right]$$

- Weights are proportional to and have the same sign as

$$w_{g,t} \propto N_1 (D_{g,t} - D_{g,.} - D_{.,t} + D_{.,.})$$

More on TWFE weights

- Weights are proportional to and have the same sign as

$$w_{g,t} \propto N_1 (D_{g,t} - D_{g,.} - D_{.,t} + D_{.,.})$$

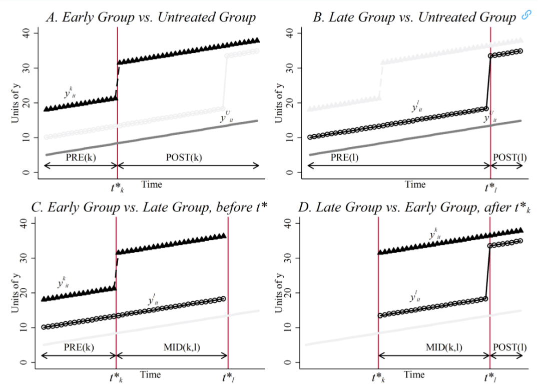
  - $D_{g,t}$: treatment in group g at period t
  - $D_{g,.}$: average treatment of group g across periods
  - $D_{.,t}$: average treatment at period t across groups
  - $D_{.,.}$: average treatment across groups and periods

$\rightarrow$ if $D_{g,t} - D_{g,.} - D_{.,t} - D_{.,.}$ is constant or uncorrelated with treatment effects, then $\hat{\beta}_{FE}$ is unbiased for ATT

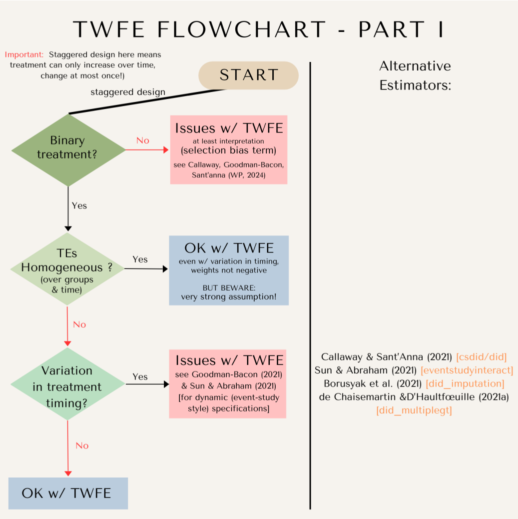
  - Constant in the case of binary, heterogeneous, no variation, staggered
  - But in most cases, no-correlation condition is implausible
    - Note specifically that $\hat{\beta}_{FE}$ downweights TEs of groups with highest average treatment and time periods with highest average treatment

TWFE and when it's a problem

- **Negative** weights may originate from comparing newly treated to previously treated groups(Goodman-Bacon 2021)



TWFE and when it's a problem



TWFE and when it's a problem

From most recent de Chaisemartin and D'Haultfoeuille review (Table 1):

Table 1: A summary of available heterogeneity-robust DID estimators

Panel A: Estimators ruling out dynamic effects (outcome unaffected by past treatments)		
Treatment	*Estimators available*	*Stata*[15] *commands*
Binary	de Chaisemartin and D'Haultfœuille (2020)	`did_multiplegt`
	Imai and Kim (2018)	
	Borusyak et al. (2021)	`did_imputation`
Discrete	de Chaisemartin and D'Haultfœuille (2020)	`did_multiplegt`
Continuous, with stayers	de Chaisemartin et al. (2022)	See paper
Continuous, w/o stayers	de Chaisemartin et al. (2022)	See paper
	Graham and Powell (2012)	`gmm`
	Chamberlain (1992)	`gmm`
Several treatments	de Chaisemartin and D'Haultfœuille (2021*b*)	`did_multiplegt`

TWFE and when it's a problem

From most recent de Chaisemartin and D'Haultfoeuille review (Table 1):

Panel B: Estimators allowing dynamic effects (outcome affected by past treatments)

Treatment	*Estimators available*	*Stata*[15] *commands*
Binary and staggered	Callaway and Sant'Anna (2021)	`csdid`
	Sun and Abraham (2021)	`eventstudyinteract`
	Borusyak et al. (2021)	`did_imputation`
	de Chaisemartin and D'Haultfœuille (2021*a*)	`did_multiplegt`
Binary or discrete, non-staggered	de Chaisemartin and D'Haultfœuille (2021*a*)	`did_multiplegt`
Continuous and staggered	de Chaisemartin and D'Haultfœuille (2021*a*)	`did_multiplegt`
	Callaway et al. (2021)	
Continuous and non-stagg., with stayers	de Chaisemartin et al. (2022)	See paper
Continuous and non-stagg., w/o stayers	No estimator available yet	
Several treatments	de Chaisemartin and D'Haultfœuille (2021*b*)	`did_multiplegt`

Overview of alternative estimators

For binary and staggered designs (with heterogeneous treatment effects):

- Methods that calculate individual treatment effects and then aggregate to single ATT
    - Callaway and Sant'Anna (2021) [csdid]
    - Sun and Abraham (2021) [eventstudyinteract]: event study context
    - de Chaisemartin and D'Haultfoeuille (2020) [did_multiplegt]

- Borusyak et al. (2021) [didimputation]: imputation based, estimate missing counterfactuals in multi-step process (extrapolate potential outcomes for the treated units, by using data for those observations that were not treated yet)

Outline

Workshop outline

Two-Way Fixed Effects (TWFE)

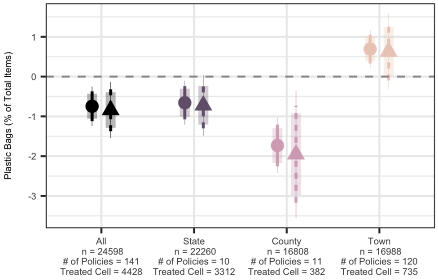
- Review of TWFE problem
- Suggestions for implementation
- Coding TWFE estimators

Suggestions for implementing new methods

- Start with and show TWFE estimator
- Then implement appropriate new estimator (maybe show robustness to others in appendix)

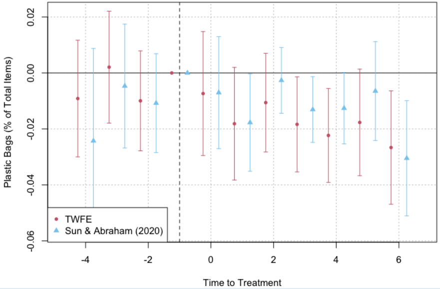
Suggestions for implementing new methods

- Start with and show TWFE estimator
- Then implement appropriate new estimator (maybe show robustness to others in appendix)
- Example (work in progress):
  - Effect of plastic bag bans and taxes on plastic in shoreline cleanups across the US
  - binary treatment, heterogeneous treatment effects, variation in treatment timing
  - (unbalanced panel, so implement Borusyak et al. as the main estimator, but show robustness to de Chaisemartin and d'Haultfoeuille in appendix)



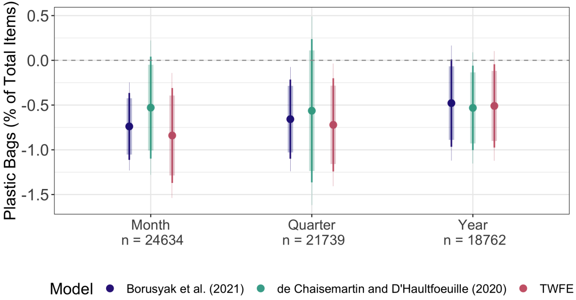
Suggestions for implementing new methods

- Start with and show TWFE estimator
- Then implement appropriate new estimator (maybe show robustness to others in appendix)
- Example (work in progress):
  - Effect of plastic bag bans and taxes on plastic in shoreline cleanups across the US
  - binary treatment, heterogeneous treatment effects, variation in treatment timing
  - (unbalanced panel, so implement Borusyak et al. as the main estimator, but show robustness to de Chaisemartin and d'Haultfoeuille in appendix)



Suggestions for implementing new methods

- Start with and show TWFE estimator
- Then implement appropriate new estimator (maybe show robustness to others in appendix)
- Example (work in progress):
  - Effect of plastic bag bans and taxes on plastic in shoreline cleanups across the US
  - binary treatment, heterogeneous treatment effects, variation in treatment timing
  - (unbalanced panel, so implement Borusyak et al. as the main estimator, but show robustness to de Chaisemartin and d'Haultfoeuille in appendix)



Outline

Workshop outline

Two-Way Fixed Effects (TWFE)

- Review of TWFE problem
- Suggestions for implementation
- Coding TWFE estimators

Coding new TWFE estimators, data

- Stevenson and Wolfers (2006), *QJE*
  - Explores the impacts of divorce law reform in the United States
  - Exploits the variation in the timing of unilateral divorce laws across states
  - Impact on female suicide rates, domestic violence, and murders of women by partners
    $\rightarrow$ decrease in all of these
- Cheng and Hoekstra (2013), *Journal of Human Resources*
  - Explores the impacts of "Castle Doctrine" ("stand your ground" laws, which expand legal justification for using lethal force in self-defense)
  - Exploits variation in the timing of these laws in 20+ states
  - Impact on homicide and violent crime
    $\rightarrow$ increase in reported murders
- $\rightarrow$ binary treatment, heterogeneous treatment effects, variation in treatment timing

Coding new TWFE estimators, summary

Use: `01_twfe.R`

- Load data (Stevenson and Wolfers by default)
- Old school TWFE and dynamic event-study version
- Bacon decomposition (earlier vs. later treated, etc.)
- Implement and compare new TWFE estimators
  - Sun and Abraham (2021)
  - Callaway and Sant'anna (2021)
  - Borusyak et al. (2021)
  - de Chaisemartin and D'Haultfoeuille (2020)

Questions? Comments?

Thank you!

References

Heavily based on Claire Palandri's 2022 version of the Causal Inference Workshop.

Chaisemartin, Clément de, and Xavier D'Haultfœuille. 2020. "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects." *The American Economic Review* 110 (9): pp. 2964–2996. ISSN: 00028282, 19447981, accessed February 15, 2024. https://www.jstor.org/stable/26966322.

Cheng, Cheng, and Mark Hoekstra. 2013. "Does Strengthening Self-Defense Law Deter Crime or Escalate Violence?" *Journal of Human Resources* 48 (3): 821–854. ISSN: 0022-166X. https://doi.org/10.3368/jhr.48.3.821. eprint: https://jhr.uwpress.org/content/48/3/821.full.pdf. https://jhr.uwpress.org/content/48/3/821.

Goodman-Bacon, Andrew. 2021. "Difference-in-differences with variation in treatment timing." Themed Issue: Treatment Effect 1, *Journal of Econometrics* 225 (2): 254–277. ISSN: 0304-4076. https://doi.org/https://doi.org/10.1016/j.jeconom.2021.03.014. https://www.sciencedirect.com/science/article/pii/S0304407621001445.

Stevenson, Betsey, and Justin Wolfers. 2006. "Bargaining in the Shadow of the Law: Divorce Laws and Family Distress*." *The Quarterly Journal of Economics* 121 (1): 267–288. ISSN: 0033-5533. https://doi.org/10.1093/qje/121.1.267. eprint: https://academic.oup.com/qje/article-pdf/121/1/267/5230654/121-1-267.pdf. https://doi.org/10.1093/qje/121.1.267.