MSI
A TETRA TECH COMPANY

ASSESSING THE NEW CLASS OF D-I-D ESTIMATORS

A replication

Dan Killian

DALL-E INTERPRETATION OF MY FINAL PREPARATION

OUTLINE OF PRESENTATION

- Background
- Problem
- Solutions
- Case study - MISTI
- Final thoughts

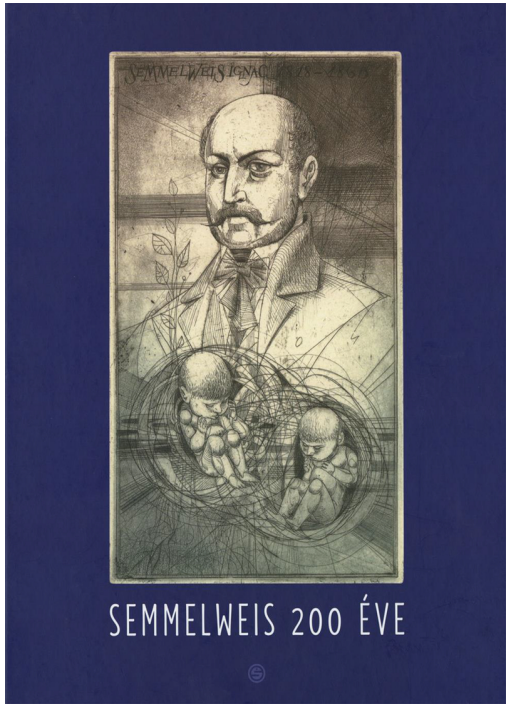
BOTTOM LINE UP FRONT

- In certain settings, beware the Two-Way Fixed Effects Estimator!
- Don’t conflate your modeling approach (TWFE) with your estimation strategy
- Examine the different groups created by differential timing
- Use event study designs
- Specify a fully flexible model (Two-way Mundlak)

- **Background**
- Problem
- Solutions
- Case study - MISTI
- Final thoughts

D-I-D HAS A LONG AND STORIED HISTORY

- Ignaz Semmelweis - Let’s track mortality rates in two maternity wards, one staffed with midwives and the other with medical students (who were busy with cadavers)



D-I-D HAS A LONG AND STORIED HISTORY

- John Snow: Let’s track cholera infection in two London neighborhoods, one with treated water and one without

D-I-D HAS A LONG AND STORIED HISTORY

- Economists - Let's steal repeated measures ANOVA from the statisticians!

- "This estimator has been labeled the difference-in-differences estimator in the recent program evaluation literature, although it has a long history in analysis of variance." [Wooldridge 2010]
- Cross Validated: Difference in Difference vs repeated measures

SO WHAT IS THE CANONICAL D-I-D SETUP?

$$y_{it} = \beta_0 + \delta_{0,t} Post_t + \beta_{1,i} Treat_i + \delta_{1,it} Post_t * Treat_i + \epsilon_{it}$$

where..

β_0 is the comparison group at baseline

δ_0 is the secular change from baseline to endline, unrelated to treatment

β_1 is the difference between the treatment and comparison groups at baseline, and

δ_1 is the treatment effect, the interaction of treatment and time

Algebraically, δ_0 can be expressed as the difference between the pre/post difference in each of the treatment and comparison groups

$$\delta_1 =$$

$$(\bar{y}_{POST,TREAT} - \bar{y}_{PRE,TREAT})$$

$$-$$

$$(\bar{y}_{POST,COMPARISON} - \bar{y}_{PRE,COMPARISON})$$

hence, difference-in-differences (d-i-d or DiD or DD)

CANONICAL D-I-D, 2X2

$$y_{it} = \beta_0 + \delta_{0,t} Post_t + \beta_{1,i} Treat_i + \delta_{1,it} Post_t * Treat_i + \epsilon_{it}$$

Canonical d-i-d 2x2 setup

	Pre	**Post**	**Post - Pre**
Comparison	β_0	$\beta_0 + \delta_0$	δ_0
Treatment	$\beta_0 + \beta_1$	$\beta_0 + \delta_0 + \beta_1 + \delta_1$	$\delta_0 + \delta_1$
Treatment - Comparison	β_1	$\beta_1 + \delta_1$	δ_1

HOW DOES THE CANONICAL D-I-D GENERALIZE TO MULTIPLE TIME PERIODS AND/OR GROUPS?

When we generalize the two-period setup to multiple time periods and/or groups, we have the two-way fixed effect (TWFE) estimator

$$y_{it} = \alpha_i + \alpha_t + \beta^{DD} it + \epsilon it$$

where..

α_i are group fixed effects

α_t are time fixed effects

B_{it}^{DD} indicates whether group *i* in period *t* is treated

TWFE IS A WORKHORSE IN PROGRAM EVALUATION

- 744 d-i-d studies across ten journals in finance and accounting, 2000-2019 [Baker 2022]
- 19 percent of all empirical articles published by the American Economic Review (AER) between 2010 and 2012 used TWFE [de Chaisemartin and D'Haultfoeuille 2020]

- Background
- **Problem**
- Solutions
- Case study - MISTI
- Final thoughts

BUT WHAT IS β_{it}^{DD} ACTUALLY TELLING US?

- For the canonical 2x2, we know exactly what we are estimating
- For *i* groups and *t* time periods, we are getting some average of multiple 2x2s
- But how does this work, exactly?
- Goodman-Bacon (2021) decided to work it out

Let's take a single step from two time periods to three, where treatment can be adopted at either t = 2 or t = 3

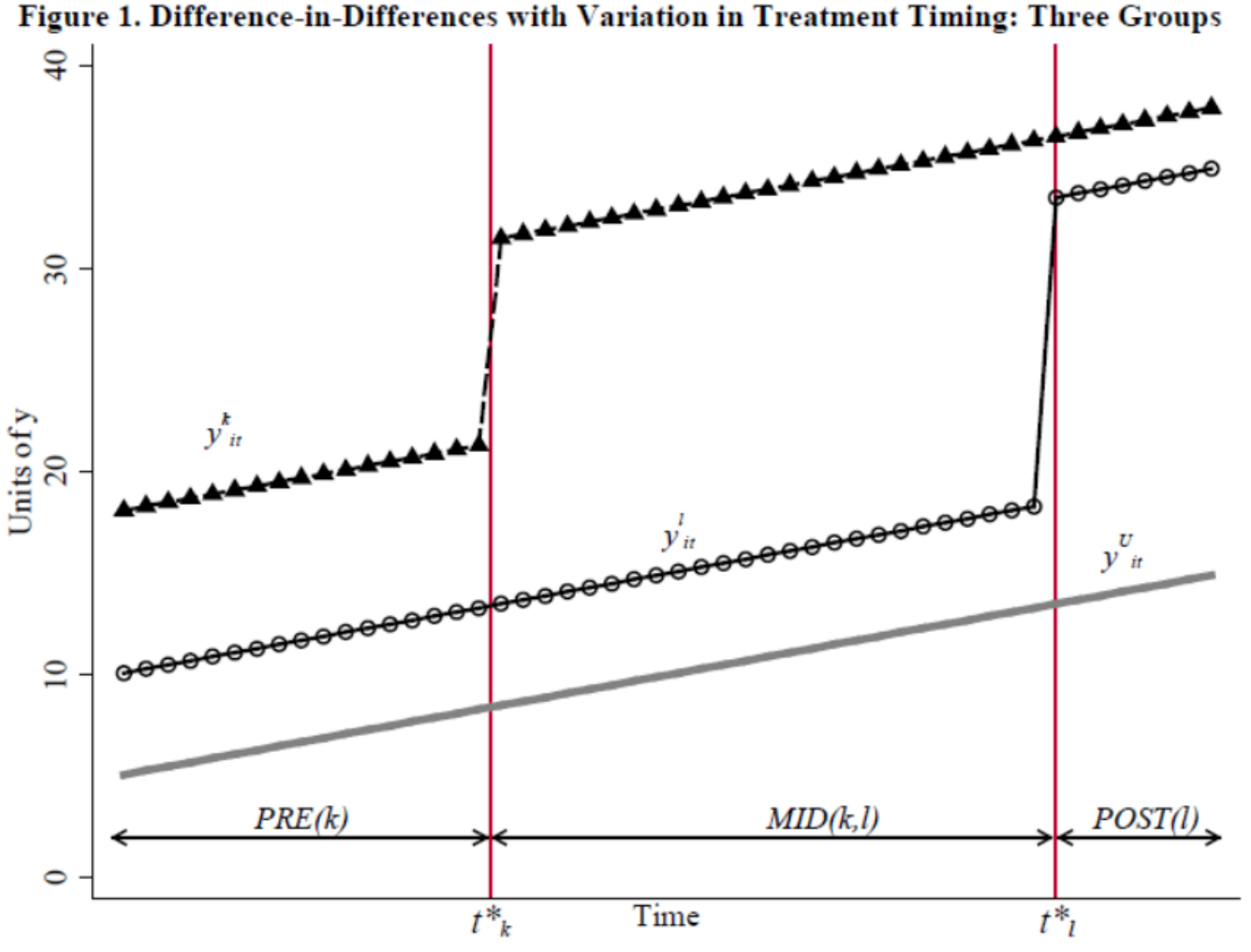


Figure 1 from Goodman-Bacon (2021)

- Baker [2022] showed that any design with multiple treatment timings will have k^2 groups, where *k* is the number of timings.

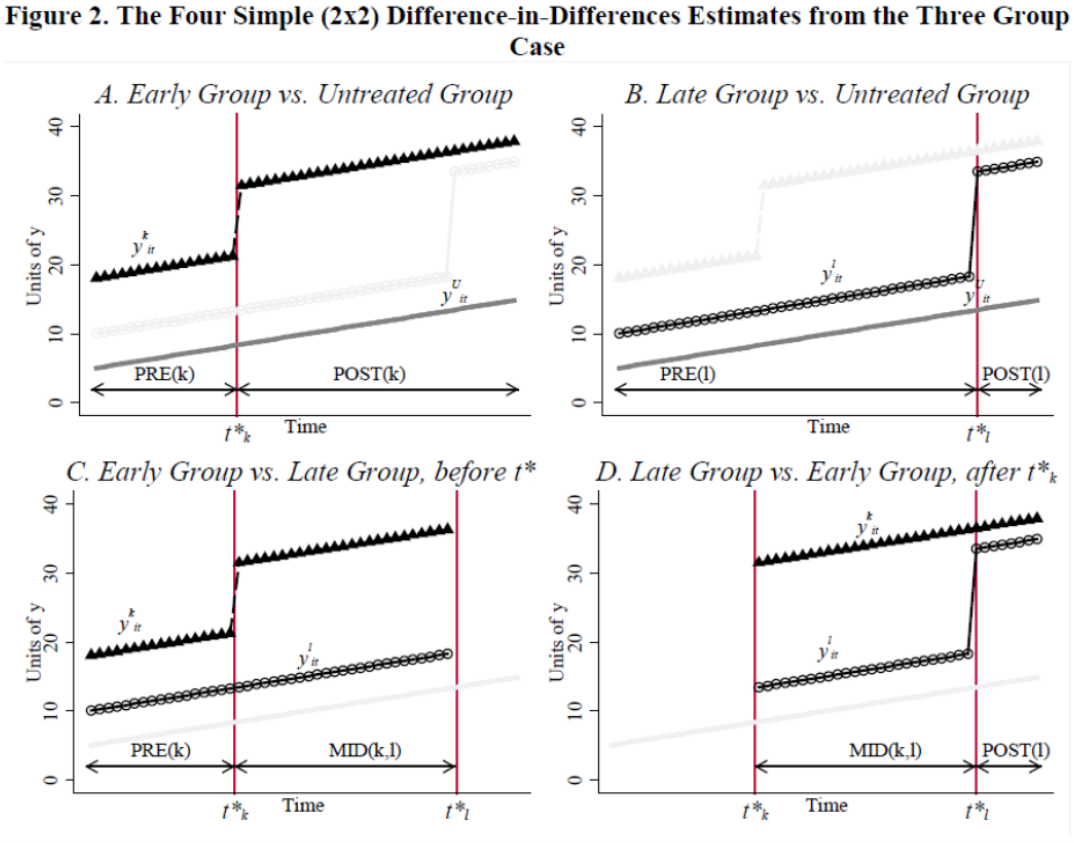
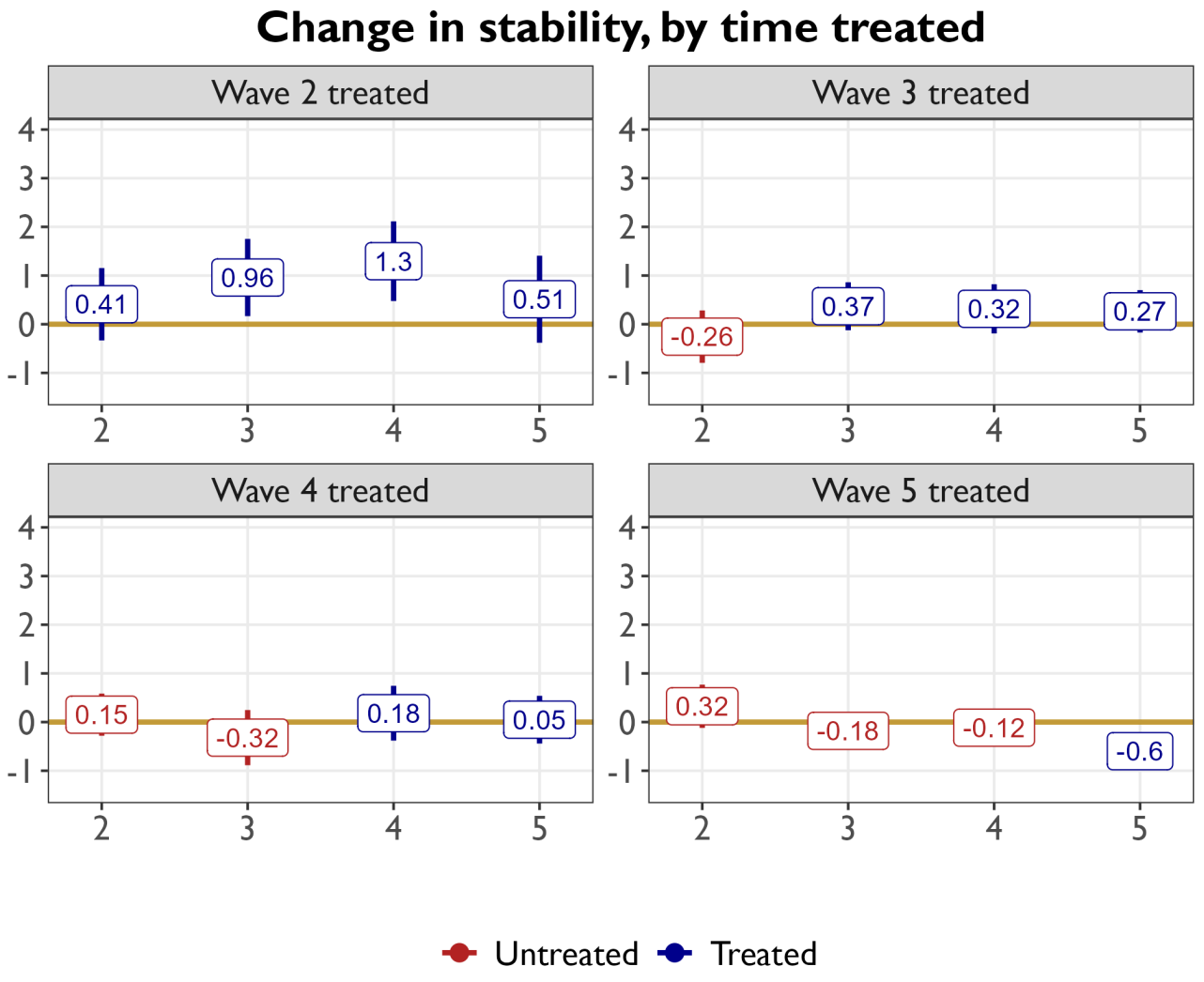


Figure 2 from Goodman-Bacon (2021)

- Background
- Problem
- **Solutions**
- Case study - MISTI
- Final thoughts

- Background
- Problem
- Solutions
- **Case study - MISTI**
- Final thoughts

CALLAWAY AND SANT'ANNA (2020)



Callaway Sant'Anna did
Outcomes in standard deviation units

- Background
- Problem
- Solutions
- Case study - MISTI
- **Final thoughts**