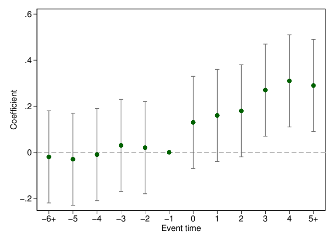
Plotting

Linear Panel Event Studies

Liyang (Sophie) Sun (CEMFI)
Jesse M. Shapiro (Harvard and NBER)

Motivation

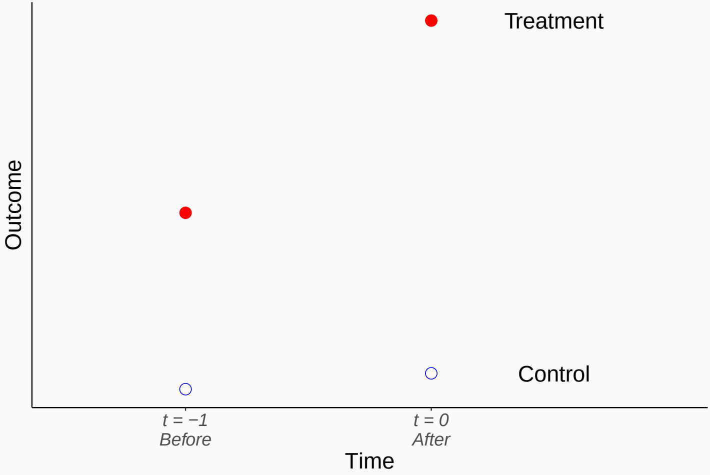


Source: Freyaldenhoven et al. (2021)

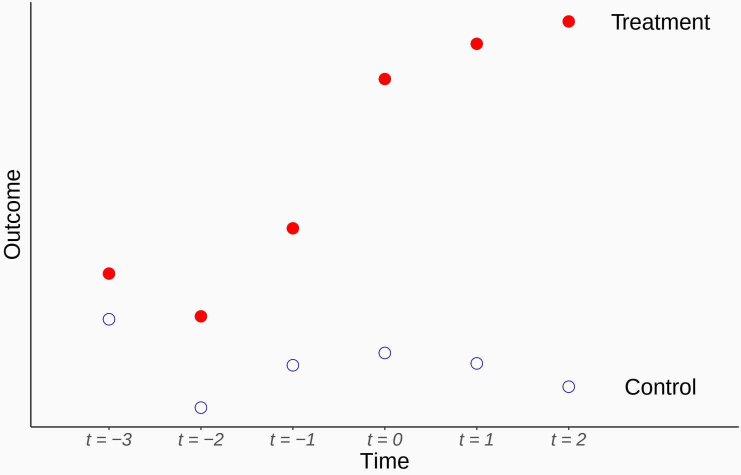
- Plotting an essential, not incidental, part of methodology
  - Of 16 papers in the 2022 AER mentioning DID or event studies, 12 do some form of pre-event testing and 10 include some form of plot of dynamic treatment effects and pre-event trends

Basics

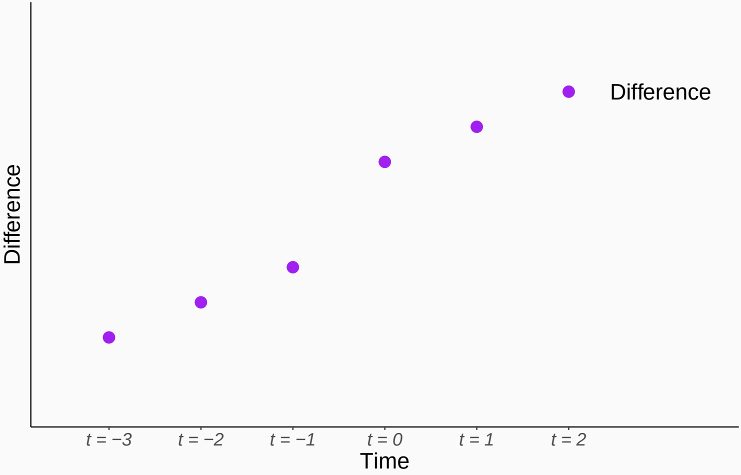
Two Groups, Two Periods



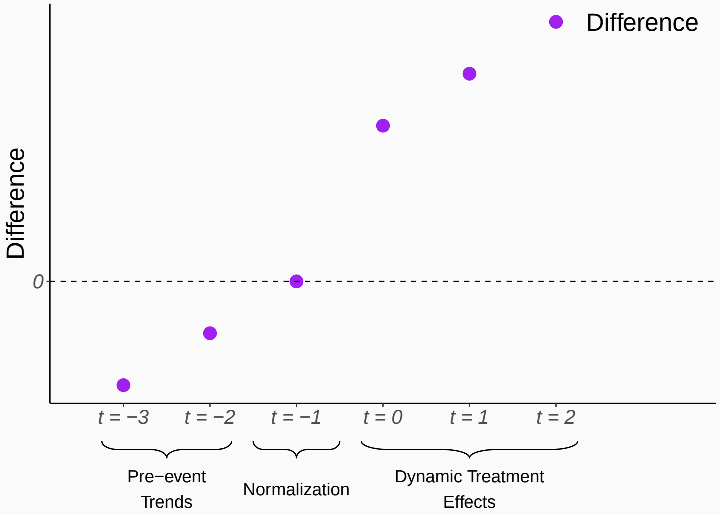
Two Groups, Many Periods



Differences, Many Periods



Normalized Differences, Many Periods



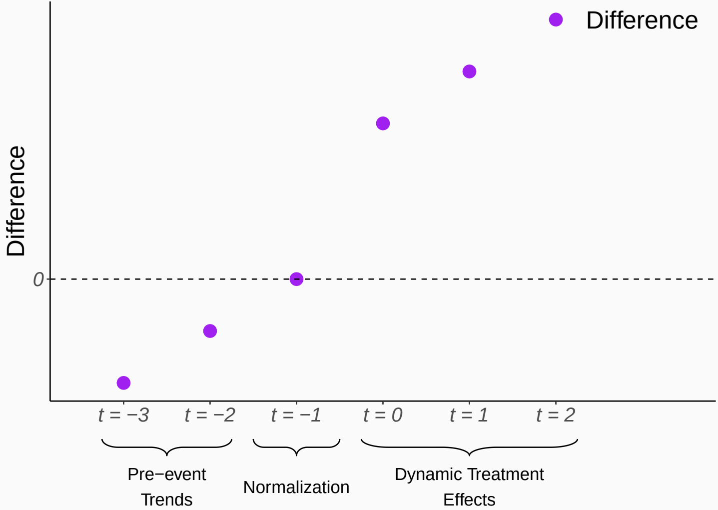
Regression Trick

- Let z_{it} be
  - 1 if i is in treatment group and t is after treatment date
  - 0 otherwise
- Estimate

$$y_{it} = \alpha_i + \gamma_t + \sum_{k=-\infty}^{\infty} \delta_k \Delta z_{i,t-k} + \varepsilon_{it}$$

  - Unit fixed effect α_i
  - Time fixed effect γ_t
  - Normalize $\delta_{-1} = 0$ so δ_k is in normalized differences
- Then plot $\left\{ \left(k, \hat{\delta}_k \right) \right\}_{k=-\infty}^{\infty}$

Normalized Differences, Many Periods



What If?

- Different units treated at different dates
  - e.g., staggered adoption of state law
- Policy $z_{it} \in \{0, 1\}$ is not binary
  - e.g., minimum wage
- Time series is not infinite
  - e.g., all real situations

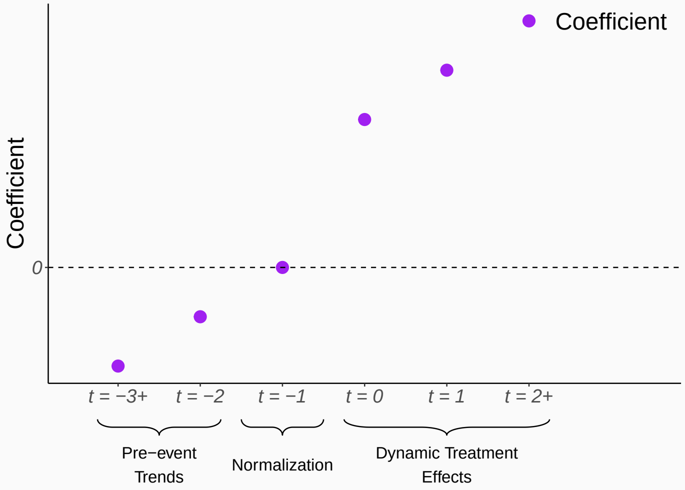
Regression Trick

- Estimate

$$y_{it} = \alpha_i + \gamma_t + \sum_{k=-(B-1)}^{A-1} \delta_k \Delta z_{i,t-k} + \delta_A z_{i,t-A} + \delta_B \left(1 - z_{i,t+B-1}\right) + \varepsilon_{it}$$

  - Normalize $\delta_{-1} = 0$
- Then plot $\left\{\left(k, \hat{\delta}_k\right)\right\}_{k=-B}^{A}$
  - $A =$ number of periods **A**fter to plot
  - $B =$ number of periods **B**efore to plot
- NB: For algebra, see Freyaldenhoven et al. (2021) or Schmidheiny and Siegloch (2023)

Event-study Plot



Substantive Decisions

$$y_{it} = \alpha_i + \gamma_t + \sum_{k=-(B-1)}^{A-1} \delta_k \Delta z_{i,t-k} + \delta_A z_{i,t-A} + \delta_B \left(1 - z_{i,t+B-1}\right) + \varepsilon_{it}$$

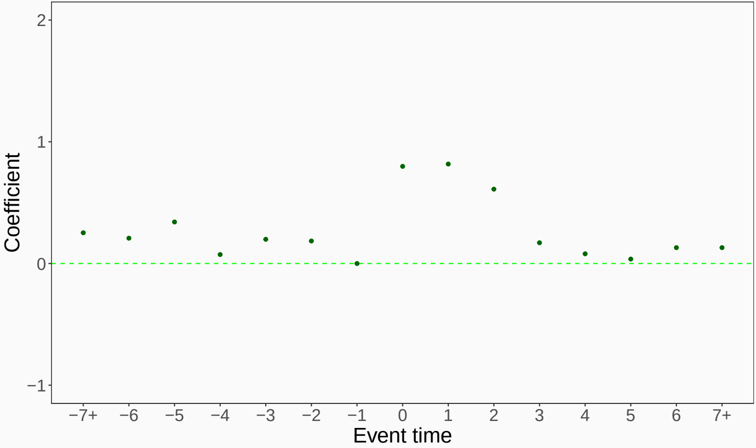
- Treating dynamics as stable more than B periods before event, A periods after
  - Can't allow for infinite dynamics due to finite data
- Estimating dynamics relative to a fixed normalization, e.g., $\delta_{-1} = 0$
  - Can't identify causal effects without a base period

Warning

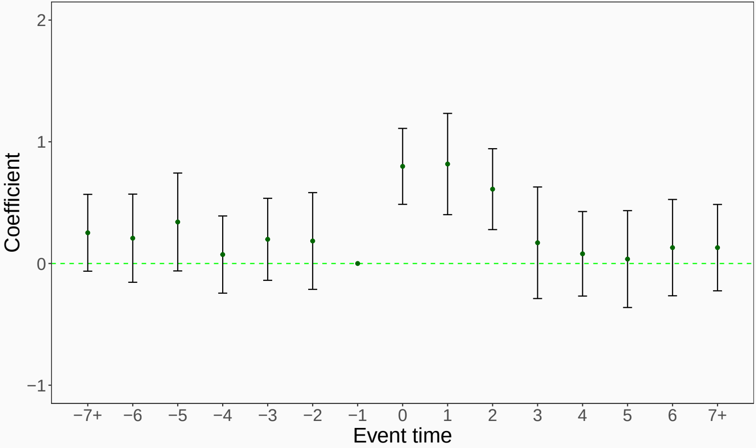
- This “trick” is one possible regression generalization of DID
- It has the virtue of being flexible
- Think of it as a starting point
- Other approaches may be more suited to your setting
- Will come back to this!

Making More Informative Plots

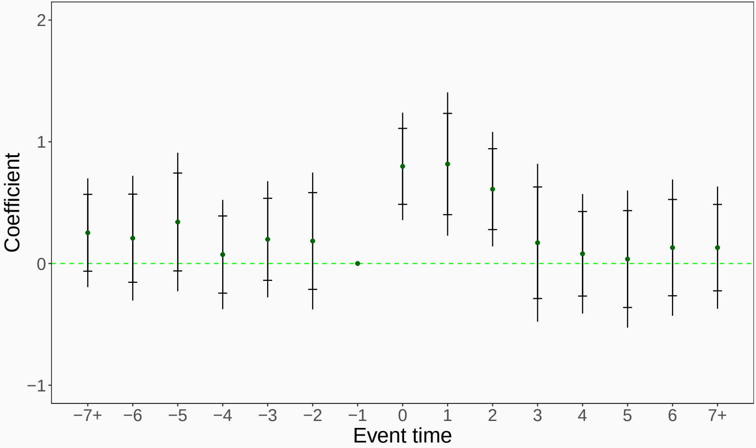
Point Estimates



Confidence Intervals

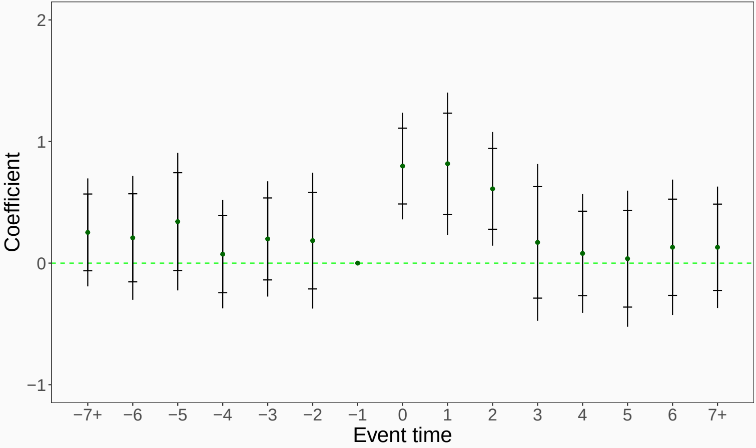


Confidence Bands



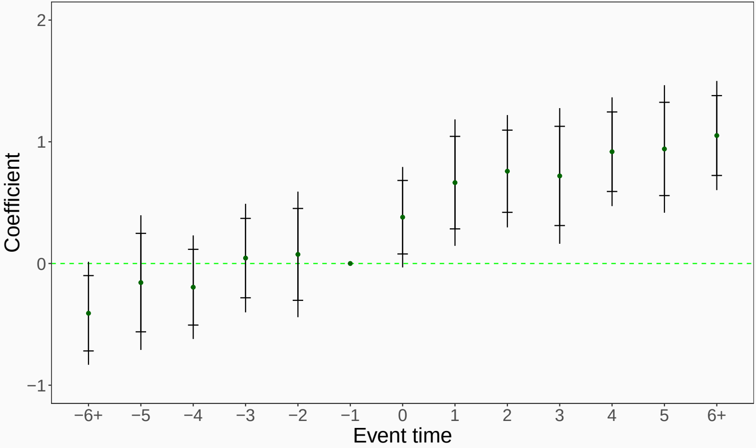
- Sup-t bands a la Montiel Olea and Plagborg-Møller (2018)

Testing



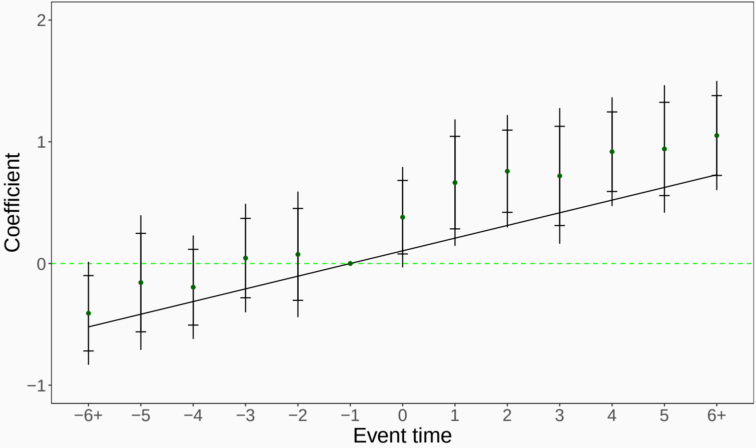
Pretrends p−value = 0.6 −− Leveling off p−value = 0.8

Confounding



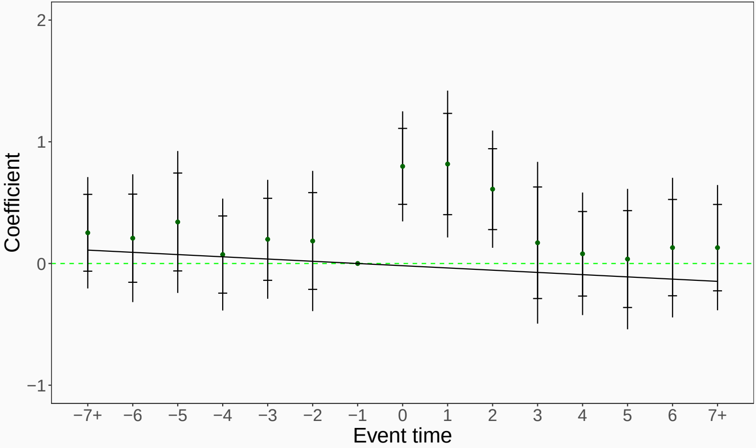
- Could confounding explain this pattern?

Confounding



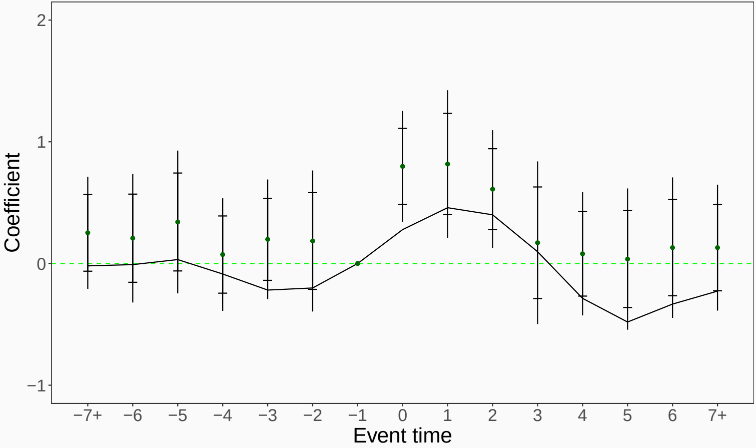
- A linear path in event-time, not statistically rejected.

Confounding



- A linear path in event-time, statistically rejected.

Least Wiggly Confound



- Defined in Freyaldenhoven et al. (2021)

Implementation

Software

- Stata: xtevent
- R: EventStudyR

Today

- ~~Overview (Jesse)~~
- ~~Basics of identification and estimation (Liyang)~~
- ~~Basics of plotting (Jesse)~~
- Pitfalls and some solutions
    - Confounds and pre-trend testing (Liyang)
    - Heterogeneous effects (Jesse)
- Conclusions (Liyang)