Confounds and Pre-trend Testing

Linear Panel Event Studies

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

Motivation

- Difference-in-differences and related methods rely on a “no anticipation” assumption and a “parallel trends” assumption
- In practice, we’re often not sure if these assumptions hold!
- Discuss common practice of testing for pre-trends
  - Role of anticipatory effects
  - Power of tests
- Discuss alternative ways to address confounding
  - Extrapolation of pre-period trends
  - Proxy IV methods

Basis of the pre-trend test

The Classical Example is Just Identified

- In the classical two-period two-group example, the model is just identified
  - Under the “no anticipation” and “parallel trends” assumptions, only one way to identify the ATT based on observed data

$$\beta = E[y_{i,0} - y_{i,-1} \mid D_i = 1] - E[y_{i,0} - y_{i,-1} \mid D_i = 0]$$

- No additional restriction is left from these assumptions

Reminder: Multiple Periods

- One treatment group and one control group
- Estimate a "dynamic" specification with normalization $\delta_{-1} = 0$:

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

- "no anticipation": $y_{it}(0) = y_{it}(1)$ for all i with $D_i = 1$ for all $t < t^*$
- "parallel trends": for all $t \neq t'$

$$E[y_{it'}(0) - y_{it}(0) \mid D_i = 1] = E[y_{it'}(0) - y_{it}(0) \mid D_i = 0]$$

Pre-trend test

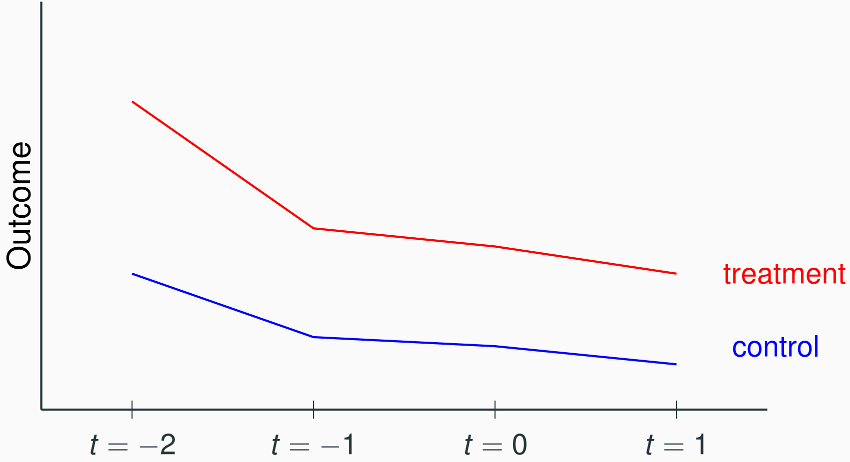
- Under “no anticipation” and “parallel trends”, we have

$$\delta_k = E[y_{i,t^*+k}(1) - y_{i,t^*+k}(0) \mid D_i = 1] \text{ for } k \geq 0$$
$$\delta_k = 0 \text{ for } k < -1$$

- Now we have the additional restrictions from the “no anticipation” and “parallel trends” assumptions to test:

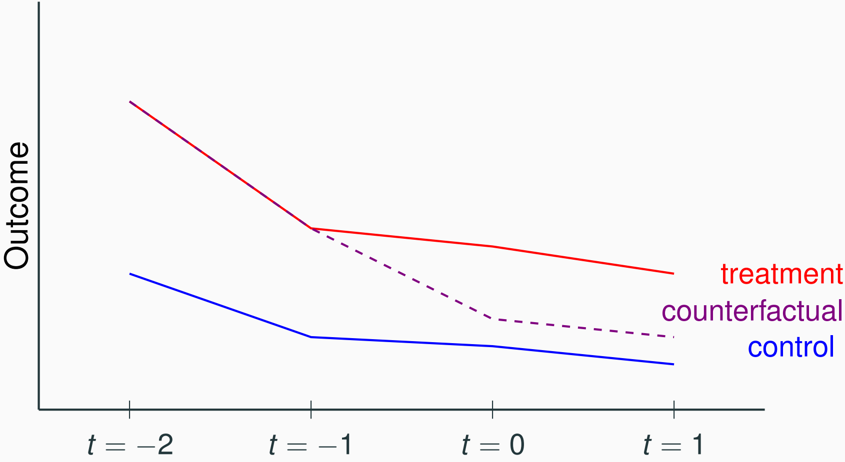
$$\text{pre-trend test } H_0 : \{\delta_k = 0\}_{k<-1}$$

Can We Test Both Assumptions?

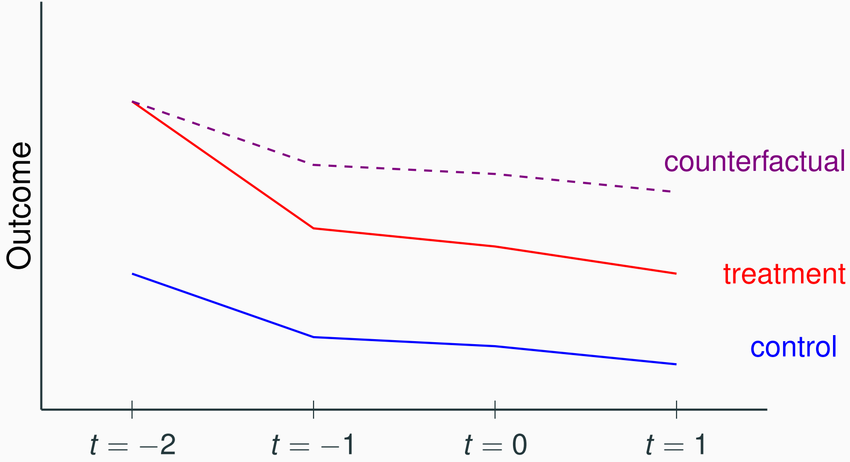


- Graphical (hypothetical) illustration for one treatment group and one control group
- Suppose we observe diverging trends between the two groups

No Anticipation, Only Selection on Trends



Only Anticipatory Effect, Parallel Trends



Summary

- Conceptually, violations of “no anticipation” and “parallel trends” are distinct
    - Anticipatory effect: treatment has causal effect prior to its implementation
    - Non-parallel trends: comparing the treatment and control group, treatment group experiences a confounding trend around the time of treatment implementation
- Observationally, violations of "no anticipation" and "parallel trends" are not distinct
- Rejection of the pre-trend test needs careful interpretation

Pitfalls with Pre-trend Tests

Issue 1 - Low Power

- Estimate a "dynamic" specification

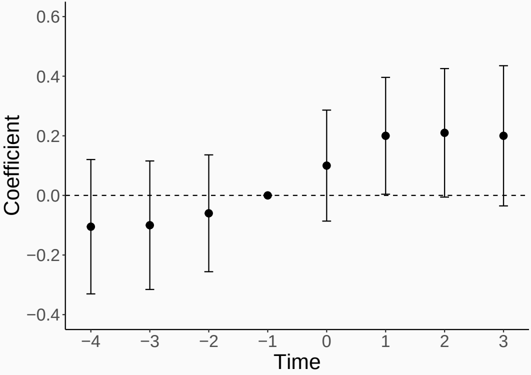
$$y_{it} = \alpha_i + \gamma_t + \sum_{-\infty}^{\infty} \delta_k \Delta z_{i,t-k} + \varepsilon_{it}$$

  and test

$$H_0 : \delta^{pre} = 0 \text{ where } \delta^{pre} = \{\delta_k\}_{k<-1}$$

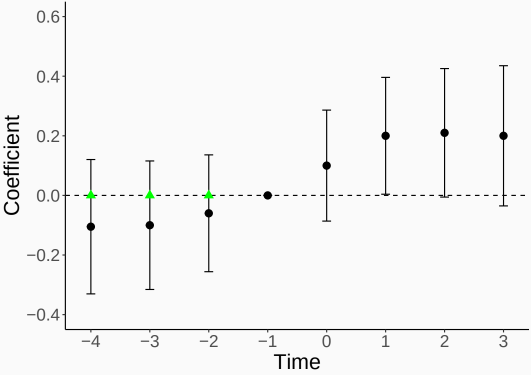
- Recent work pointed out the pre-trend test may fail to detect violations of "parallel trends" (Freyaldenhoven, Hansen, and Shapiro 2019, Kahn-Lang and Lang 2020, Bilinski and Hatfield 2020, Roth 2022)
- Graphical (hypothetical) illustration based on Roth (2022)

Issue 1 - Low Power



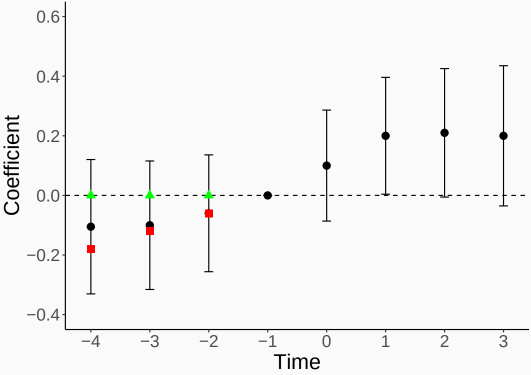
- Can we reject parallel trends in this event study?

Issue 1 - Low Power



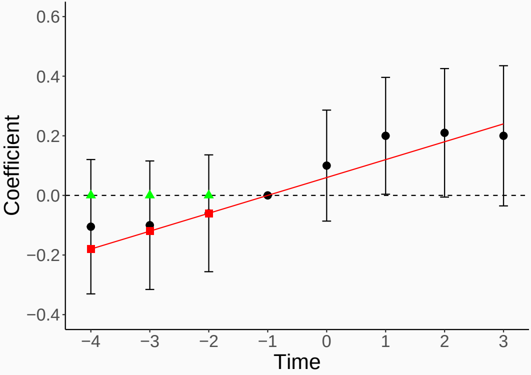
- P-value for $H_0 : \delta^{pre} =$ green triangles (no pre-trend): 0.7

Issue 1 - Low Power



- P-value for $H_0 : \delta^{pre} =$ green triangles (no pre-trend): 0.7
- P-value for $H_0 : \delta^{pre} =$ red squares: 0.7

Issue 1 - Low Power



- P-value for $H_0 : \delta^{pre} =$ green triangles (no pre-trend): 0.7
- P-value for $H_0 : \delta^{pre} =$ red squares: 0.7
- We can't reject zero pre-trends, but also can't reject pre-trends that under linear extrapolations would produce substantial bias

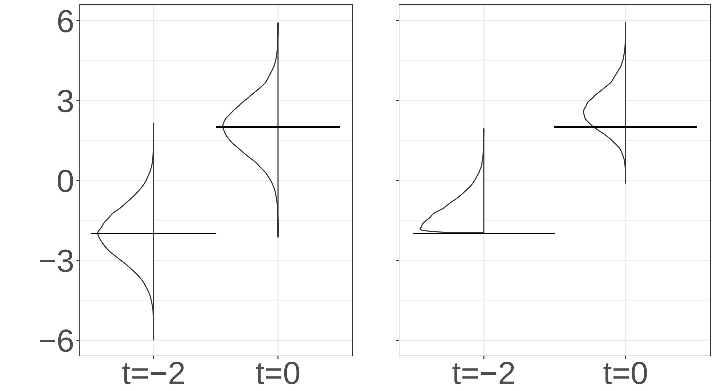
More Systematic Evidence

- Roth (2022): simulations calibrated to papers published in AEA journals
  - Many tests have limited power against reasonable alternatives, for example, linear confounding trends
- Roth (2022) provides package that evaluates power for any given application
  - `pretrends` package / Shiny app
- If power for reasonable alternatives is too low, then we might feel skeptical whether parallel trend holds even though $H_0 : \delta^{pre} = 0$ cannot be rejected

Issue 2 - Screen based on the pre-trend test

- Report estimates only if the pre-trend test passes. Does that yield an improved estimator?
- Estimates for δ_k for $k < -1$ are correlated with estimates for δ_k for $k \geq 0$
- When there is indeed confounding trend,
  - Condition on passing the pre-trend test $\leftrightarrow$ screen on whether $\hat{\delta}_k$ for $k < -1$ are small enough
  - Affects the original asymptotic normal approximation for $\hat{\delta}_k$ for $k \geq 0$
- Roth (2022): simulations calibrated to papers published in AEA journals
  - Screening induces a large bias that can be similar in magnitude to estimated effect
- Solution: emphasize tests for pre-trends only when these are powerful

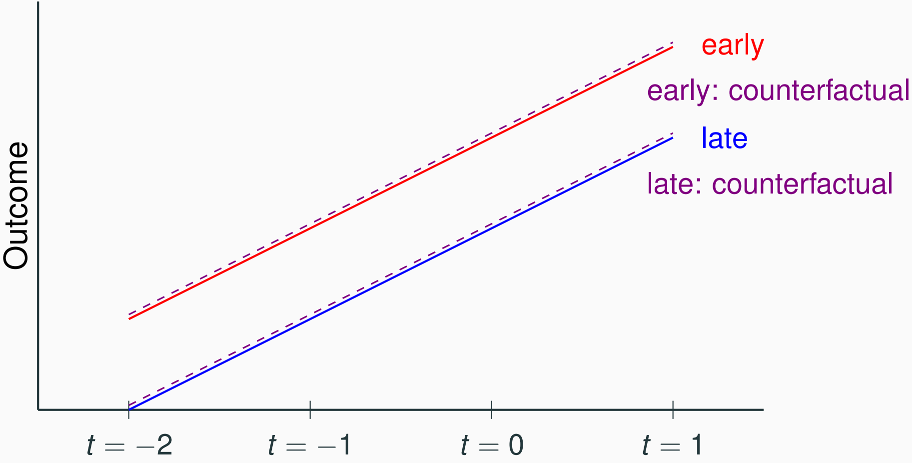
Issue 2 - Screen based on the pre-trend test: Illustration



- Upward confounding trend and positively correlated $(\hat{\delta}_{-2}, \hat{\delta}_0)$
- Upward biased estimate without screening (left)
- Screening exacerbates the bias (right) $\rightarrow$ pre-test bias

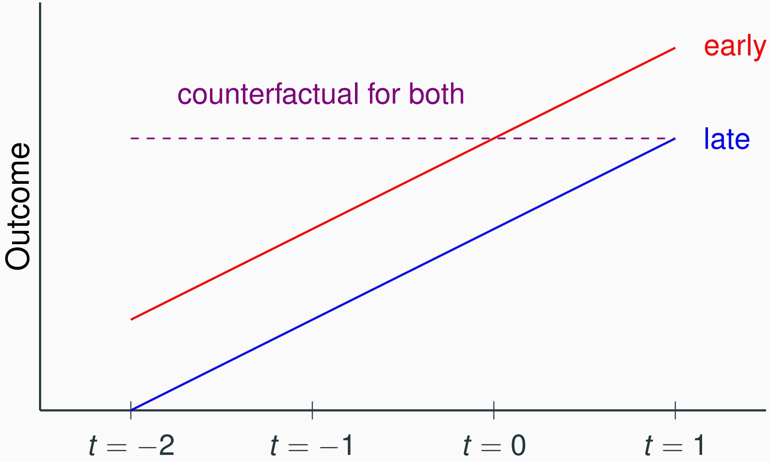
Issue 3 - Cannot Detect a Linear Violation

Only observe an early ($g(i) = 0$) and a late ($g(i) = 1$) treatment group. The data is consistent with no violation.



Issue 3 - Cannot Detect a Linear Violation

Borusyak, Jaravel and Spiess (2023): the data is also consistent with linear violations.



Issue 3 - Cannot Detect a Linear Violation

- The issue is that for "dynamic" specification,

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

  - when estimated without a control group,
  - includes all possible relative time indicators $\Delta z_{i,t-k}$
- The relative time indicators are multicollinear with the calendar time indicators
  - Note that $t - g(i) = k$

Issue 3 - Cannot Detect a Linear Violation

- Need to introduce some restriction about the DGP first and then test the remaining restrictions
- Since common software packages directly omit the collinear regressors, it would be good to check which ones are omitted

Issue 3 - Cannot Detect a Linear Violation

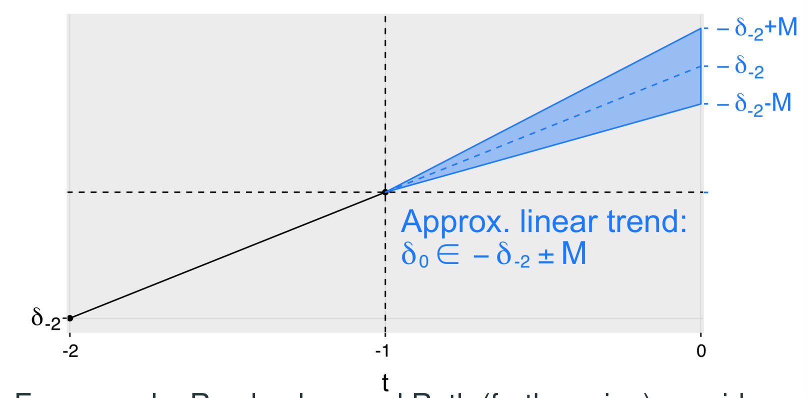
- Solution: make a conscious decision of normalization (in addition to $\delta_{-1} = 0$)
- For example,
  - Normalize at least another distant lead: assumes "no anticipation" and "parallel trends" assumptions hold between $g(i) - 1$ and $g(i) - B$ for each group
- In the "plotting" module, we suggest
  - Treat dynamics as stable more than B periods before event, A periods after

Solutions Under Potential Violations to Parallel Trends

Sensitivity Analysis

- Non-zero pre-trends can be informative about the violations to the parallel trends assumption
  - Provides information on the amount of bias in $\hat{\delta}_k$ for $k \geq 0$ (sensitivity analysis)
  - Empirical papers informally extrapolate the pre-trends to remove the bias, e.g., Dobkin et al. (2018)
- Manski and Pepper (2018) and Rambachan and Roth (forthcoming) relax the exact extrapolation

Sensitivity Analysis: Illustration



- For example, Rambachan and Roth (forthcoming) consider bounds on how far δ_0 can deviate from a linear extrapolation of the pre-trend: $\delta_0 \in [-\delta_{-2} - M, -\delta_{-2} + M]$
- Construct confidence sets with correct coverage under the assumed bounds: `HonestDiD package / Shiny app`

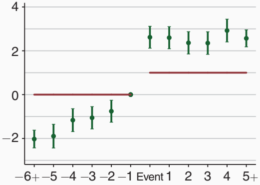
Proxy IV Estimation

- Sometimes we know the cause of confounding trend, e.g., labor demand is the confounder in the example of minimum wage increase on youth employment
- But we only observe a noisy measure for labor demand
  - For example, prime-age employment
- Freyaldenhoven, Hansen and Shapiro (2019) argue that under some conditions, leads of the treatment can be used as instruments for the noisy proxy
  - Stata: xtevent
  - R: EventStudyR
- Including the noisy proxy as a control variable does not fully remove bias

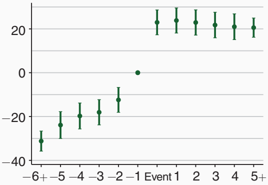
Proxy IV Estimation: Illustration

- Intuition: remove bias by subtracting off rescaled noisy proxy

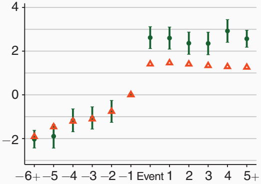
Panel A. Outcome of interest y_{it} around event time



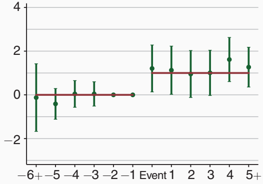
Panel B. Unaffected covariate x_{it} around event time



Panel C. Overlaying outcome of interest y_{it} (with confidence intervals) and rescaled unaffected covariate x_{it} (triangles) around event time



Panel D. Outcome of interest y_{it} around event time, using the behavior of the covariate to net out the effect of the confound



Further Reading

Borusyak, Kirill, Xavier Jaravel and Jann Spiess. 2023. Revisiting Event Study Designs: Robust and Efficient Estimation. In *arxiv [econ].*

Bilinski, Alyssa and Laura A. Hatfield. 2020. Nothing to See Here? Non-Inferiority Approaches to Parallel Trends and Other Model Assumptions. In *arxiv [stat].*

Freyaldenhoven, Simon, Christian Hansen, and Jesse M. Shapiro. 2019. Pre-event Trends in the Panel Event-Study Design. In *American Economic Review.*

Kahn-Lang, Ariella and Kevin Lang. 2020. The Promise and Pitfalls of Differences-in-Differences: Reflections on 16 and Pregnant and Other Applications. In *Journal of Business & Economic Statistics.*

Manski, Charles F. and John V. Pepper. 2020. How Do Right-to-Carry Laws Affect Crime Rates? Coping with Ambiguity Using Bounded-Variation Assumptions. In *The Review of Economics and Statistics.*

Rambachan, Ashesh and Jonathan Roth. Forthcoming. A More Credible Approach to Parallel Trends. In *The Review of Economic Studies.*

Roth, Jonathan. 2022. Pretest with Caution: Event-Study Estimates after Testing for Parallel Trends. In *American Economic Review: Insights.*

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)