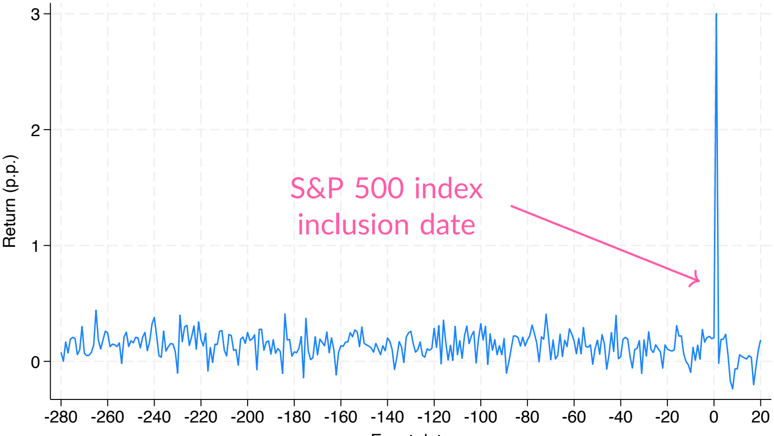
Causal inference in Financial Event Studies

Paul Goldsmith-Pinkham[1]   Tianshu Lyu[2]

[1]Yale SOM & NBER [2]Yale SOM

Tying together two literatures, and extending an old debate

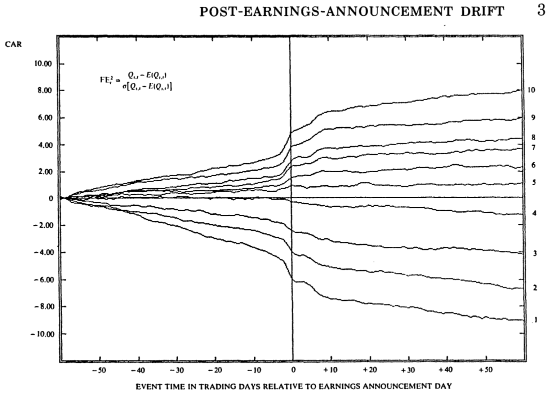
- Finance literature studying the impact of events on asset prices
- Econometrics literature estimating the average treatment effect on the treated using model and design-based inference



Historically, event studies are an important tool

What types of financial events? Examples...

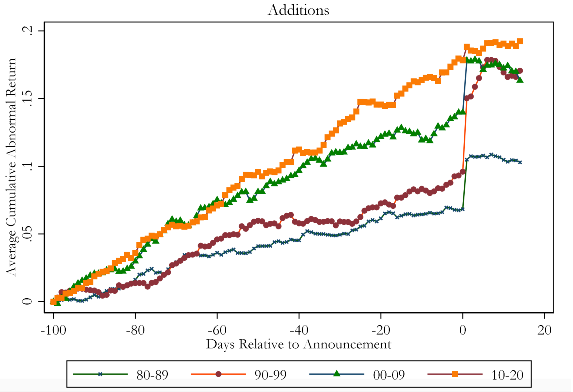
- Earnings Announcements
- Index Inclusion
- Mergers and acquisitions
- IPO, SEO, Shares repurchased
- CEO/CFO Changes
- Patent Issuance
- FOMC Announcements
- Labor Issues
- Political events



Historically, event studies are an important tool

What types of financial events? Examples...

- Earnings Announcements
- Index Inclusion
- Mergers and acquisitions
- IPO, SEO, Shares repurchased
- CEO/CFO Changes
- Patent Issuance
- FOMC Announcements
- Labor Issues
- Political events



Historically, event studies are an important tool

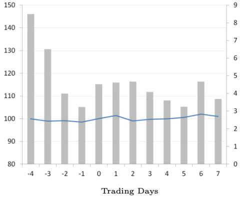
What types of financial events? Examples...

- Earnings Announcements
- Index Inclusion
- Mergers and acquisitions
- IPO, SEO, Shares repurchased
- CEO/CFO Changes
- Patent Issuance
- FOMC Announcements
- Labor Issues
- Political events

(a) Patent 4,946,778 granted to Genex on Aug, 7 1990, "Single Polypeptide Chain Binding Molecules."

(b) Patent 5,585,089 granted to Protein Design on Dec 17, 1996, "Humanized Immunoglobulins."

(c) Patent 6,317,722 granted to Amazon.com on Nov 13, 2001, "Use Of Electronic Shopping Carts To Generate Personal Recommendations."



(d) Patent 6,329,919 granted to IBM on Dec 11, 2001, "System and Method For Providing Reservations For Restroom Use."

Contribution of this paper (1/2)

- Reframe event studies in the view of causal inference literature
  - What is the counterfactual return?
- Characterize when standard abnormal return estimates are biased
  - Short-run – it depends
  - Long-run – almost always
- Almost all existing approaches use *model*-based counterfactuals
  - Counterfactual return is based on expected return
  - Requires *model stability* of factor structure
- Connection between historical approaches (CAR, BHAR, Calendar Time)
  - *buy-and-hold (geometric)* measures need to match on both counterfactual means *and variance*
  - BHAR may bias treatment effects downward due to volatility drag
- Propose alternative estimators
  - Synthetic control
  - PCA regression (GSynth)
  - Potentially many others!

Contribution of this paper (2/2)

- Highlight results in three applications
  1. Political Connections (Acemlogu et al. (2016))
  2. S&P 500 Index Inclusion (Greenwood and Sammon (2025))
  3. Effects of mergers on acquirer value (Malmandier (2018))

- Key takeaways:
  1. significant potential bias in short-run events when only one event
  2. no bias in short-run when many events with random timing
  3. significant potential bias in long-run events, even with many random events

- Key things still in progress:
  - Robust results on inference
  - Framework for partial information incorporation

What is the effect of an event on stock returns? Potential outcomes

- Unit of analysis: a path of stock returns $R_i = \{R_{i1}, \ldots, R_{iT}\}$
  - Set of n securities (firms) observed over T time periods

- $D_i \in \{0, 1\}$: an event happens at t_0 to $n_0 < n$ firms

- For each stock and time period, there are two potential versions of R_{it}:
  - $R_{it}(1)$: the firm experienced the event
  - $R_{it}(0)$: the firm without the event
  - Researchers are interested in the *causal effect* of the event:

$$\tau_{it} = R_{it}(1) - R_{it}(0)$$

- Fundamental problem of causal inference:

$$R_{it} = R_{it}(1)D_i + R_{it}(0)(1 - D_i)$$

Placing a model on the structure of counterfactual returns

Textbook approach approximates with *abnormal returns* (Campbell, Lo, Mackinlay (1997))

$$AR_{it} = R_{it} - \underbrace{\mathbb{E}(R_{it}|X_t)}_{\text{Normal Returns given } X_t}$$

- $\mathbb{E}(R_{it}|X_t)$ can reflect many models of expected returns (MacKinley (1997))
  - Market Model, CAPM, Fama-French

Placing a model on the structure of counterfactual returns

Textbook approach approximates with *abnormal returns* (Campbell, Lo, Mackinlay (1997))

$$AR_{it} = R_{it} - \underbrace{\mathbb{E}(R_{it}|X_t)}_{\text{Normal Returns given } X_t}$$

- $\mathbb{E}(R_{it}|X_t)$ can reflect many models of expected returns (MacKinley (1997))
    - Market Model, CAPM, Fama-French

$$R_{i,t} = \alpha_i + \beta_{i,1} \underbrace{F_{1,t}}_{\text{Risk Factor}} + \underbrace{\beta_{i,2}}_{\text{Factor Loading}} F_{2,t} + \tau_t D_i + \varepsilon_{i,t}$$

- Use pre-event data to estimate factor loadings (hence $\tau_s = 0$ for $s \leq t_0$)
- If model is exactly correctly specified, no issues

Misspecification in the abnormal return estimator

- What happens when a factor is omitted?

$$\widehat{AR}_{it} = R_{it} - \hat{\alpha}_i - \hat{\beta}_{i,1} F_{1,t}$$

$$\overline{AR}_t = n_s^{-1} \sum_{i \in n_1} \widehat{AR}_{it} \approx \tau_t + \underbrace{\overline{\beta}_2 \left(F_{2,t} - \overbrace{\frac{Cov(F_{2,t}, F_{1,t})}{Var(F_{1,t})}}^{\text{OVB}} F_{1,t} \right)}_{\text{misspecification error}} + \underbrace{n_1^{-1} \sum_{i \in n_1} \varepsilon_{i,t}}_{\text{noise}}$$

where $\overline{\beta}_2 = n_1^{-1} \sum_{i:T_i=s} \beta_{i,2}$

The short-and-long consequences of misspecification

$$\overline{AR}_t - \tau_t \approx \underbrace{\overline{\beta}_2 \left(F_{2,t} - \frac{Cov(F_{2,t}, F_{1,t})}{Var(F_{1,t})} F_{1,t} \right)}_{\text{misspecification error}} + \underbrace{n_1^{-1} \sum_{i \in n_1} \varepsilon_{i,t}}_{\text{noise}}$$

- Average noise is mean zero, and disappears with large n_1
- Misspecification error does not disappear with large n_1
  - Single event: $\overline{AR}_t$ is stochastic
- Trade-off between magnitudes of $F_{2,t}$ and τ_t
- If τ_t is large relative to $F_{2,t}$, then bias will be second order
  - However, $F_{2,t}$ is stochastic, and may coincide with event
  - Size of factor loading matters as well ($\overline{\beta}_2$)

More general framework: setup and notation

- $i = 1, \ldots, N$ securities ; $t = 1, \ldots, T$ time.
- Binary treatment path $D_{i,t}$ is irreversible : $D_{i,1} = 0, \; D_{i,t} = 1 \Rightarrow D_{i,t+1} = 1$
- Event timing $T_i = \begin{cases} t & \text{if event hits } i \text{ at } t \\ \infty & \text{if never} \end{cases}$
- Let $C = \{i : T_i = \infty\}$ and S the set of possible event dates.
- Potential returns $R_{i,t}(s)$ if event happens at s, and $R_{i,t}(\infty)$ if never.

Counterfactual Returns: Linear Factor Model

Assumption: Factor structure

$$\mathbb{E}\left[R_{i,t}(\infty) \mid T_i = s, F_t\right] = \alpha_s + \beta_s F_t,$$

with K common factors F_t and group means (α_s, β_s)

- Explicitly delivers $E[R_{i,t}(0) \mid T_i = s]$ used by most event-study models.
- Motivated by finance theory papers but strong
  - e.g. Chamberlain and Rothschild (1983)
  - Key question: is α_s non-zero? Should it be incorporated into the counterfactual estimate?

Limited Anticipation + Limited Effects

Assumption: Limited Anticipation

$$R_{i,t}(T_i) = R_{i,t}(\infty) \quad \text{for all } t < T_i - \delta_1$$

- Rules out pre-event price effects within the estimation window
- Justifies using pre-event data to learn counterfactuals

Assumption: Limited Effects

- If concerned that treatment *changes* risk loadings, can also consider a "post-treatment" stability assumption:

$$\mathbb{E}\left[R_{i,t}(s) \mid T_i = s, F_t\right] = \alpha_s^* + \beta_s^* F_t, \quad \text{for all } t > s + \delta_2$$

- Could use this to decompose treatment effects

Event-Assignment Mechanics

Timing propensity score

$p_t(X_i, F) = \Pr(T_i = t \mid X_i, F), \; X_i = (\alpha_i, \beta_i)$

- **Random assignment:** $p_t(X_i, F) = p_t(F)$
- **Random timing:** $p_t(X_i, F) = p_t(X_i)$

Random assignment controls who is treated; random timing controls *when*

Average Treatment Effect on the Treated as a building block

$$\tau_i(s,t) = R_{i,t}(s) - R_{i,t}(\infty), \qquad \tau_{\text{ATT}}(s,t) = \mathbb{E}[\tau_i(s,t) \mid T_i = s]$$

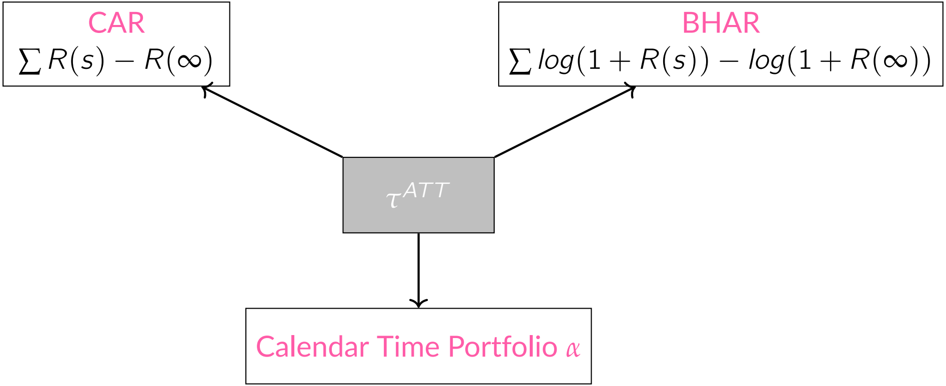
$$\begin{aligned}\tau_{\text{ATT}}(s,t) &= \mathbb{E}[\tau_i(s,t) \mid T_i = s] = \mathbb{E}[R_{i,t}(s) - R_{i,t}(\infty) \mid T_i = s] \\ &= \underbrace{\mathbb{E}[R_{i,t} \mid T_i = s]}_{\text{Observed}} - \underbrace{\mathbb{E}[R_{i,t}(\infty) \mid T_i = s]}_{\text{Counterfactual Return}}\end{aligned}$$

- Just a question of how we generate the average counterfactual return

$$\theta_\kappa = \sum_{s \in S} w_s \, \tau_{\text{ATT}}(s, s+\kappa), \quad w_s = \frac{N_s}{\sum_{s'} N_{s'}} \text{ (under random timing)}$$

- Common special case: cumulative effect $\Theta_H^{\text{CATT}} = \sum_{\kappa=0}^{H} \theta_\kappa$.

Connection between different estimators/estimands



- Control group for BHAR needs to match both *levels* and *variance*

$$\tau^{geo,ATT}(s,t) \approx \tau^{ATT}(s,t) - E(R_{it}(\infty)\tau_i(s,t) + \frac{1}{2}(\tau_i(s,t))^2 \mid T_i = s).$$

Estimator 1: Abnormal Returns

$$\widehat{R}_{i,t} = \hat{\alpha}_i + \hat{\beta}_i F_t^o \qquad (t < T_i - \delta), \qquad AR_{i,t} = R_{i,t} - \widehat{R}_{i,t}$$

$$\widehat{\tau}^{AR}(s,t) = \mathbb{E}[AR_{i,t} \mid T_i = s]$$

- Standard CAPM / Fama-French approach
- Subject to issues in simple exmple above unless F_t^o spans the true factors
- Counterfactual return generated by $\hat{R}_{it}$ model

Estimator 2: Difference-in-Means

$$\hat{\tau}^{\text{cont}}(s,t) = \mathbb{E}[R_{i,t} \mid T_i = s] - \mathbb{E}[R_{i,t} \mid i \in C].$$

- If C is the full market, $\approx$ equal-weighted market-adjusted return model
- Counterfactual return generated by average of other stocks
  - Consistent under random assignment

Estimator 3: Synthetic Control

$$\widehat{\tau}^{\text{SC}}(s,t) = R_{s,t} - \sum_{j \in C} \hat{\omega}_j R_{j,t},$$

with weights $\hat{\omega}$ chosen to *exactly* fit pre-event paths.

- Requires that a weighted portfolio of controls can replicate treated pre-trend
  - Ben-Michael and Feller (2021) show that even with imperfect fit this can be used
- No need for the factor model to be specified by researcher
  - With linear factor model, will exactly recover model
- Counterfactual return generated constructing replicating pre-period portfolio

Estimator 4: PCA regression / Gsynth Xu (2017)

$$\widehat{\tau}^{SC}(s,t) = R_{s,t} - \hat{\alpha}_s - \hat{\lambda}_s \hat{f}_t,\ t \geq \delta_1$$

- factors $\hat{f}_t$ are constructed using control firms
- loadings $\hat{\lambda}_s$ are constructed in the pre-period
- This approach is effectively PCA regression in pre-period
  - Factors constructed using PCA, and then dimensionality chosen via cross-validation
- No need for the factor model to be specified by researcher
  - With linear factor model, will exactly recover model
- Counterfactual return generated using control stocks' factor structure, and pre-event treated firms' loadings

Finite-sample bias approximation, ignoring idiosyncratic error

Under the assumptions of a limited anticipation and linear factor model:

$$\tau^{AR}(s,t) - \tau_{\text{ATT}}(s,t) = (\alpha_s - \hat{\alpha}_s) + (\beta_s F_t - \hat{\beta}_s F_t^o)$$

$$\hat{\tau}^{\text{cont}}(s,t) - \tau_{\text{ATT}}(s,t) = (\alpha_s - \alpha_\infty) + (\beta_s - \beta_\infty) F_t$$

$$\hat{\tau}^{\text{sc}}(s,t) - \tau_{\text{ATT}}(s,t) = (\alpha_s - \hat{\alpha}_s^{\text{sc}}) + (\beta_s - \hat{\beta}_s^{\text{sc}}) F_t$$

$$\hat{\tau}^{\text{gsynth}}(s,t) - \tau_{\text{ATT}}(s,t) = (\alpha_s - \hat{\alpha}_s^{\text{gsynth}}) + (\beta_s F_t - \hat{\beta}_s^{\text{gsynth}} \hat{f}_t)$$

- Estimator's error is proportional to difference with α_s, β_s in the pre-event window.
- Misspecifying factors shows up through $\beta_s F_t$ terms

Large-sample limits with both n and T to ∞

$$\tau^{AR}(s,t) - \tau_{\text{ATT}}(s,t) \xrightarrow{p} (\alpha_s - \tilde{\alpha}_s) + (\beta_s F_t - \tilde{\beta}_s F_t^o)$$

$$\hat{\tau}^{\text{cont}}(s,t) - \tau_{\text{ATT}}(s,t) \xrightarrow{p} (\alpha_s - \alpha_\infty) + (\beta_s - \beta_\infty) F_t$$

$$\hat{\tau}^{\text{gsynth}}(s,t) - \tau_{\text{ATT}}(s,t) \xrightarrow{p} 0$$

$$\hat{\tau}^{\text{SC}}(s,t) - \tau_{\text{ATT}}(s,t) \xrightarrow{p} 0$$

- Even with $n_s, n_c, T \to \infty$, AR and DiD are biased if the factor model is wrong
- Synthetic control is unbiased under exact pre-event fit
- PCA regression able to recover underlying factor structure as well

When do simple estimators work?

- **Random assignment** $\Rightarrow$ Difference-in-mean is unbiased even with a fixed T:

$$\hat{\tau}^{\text{cont}}(s,t) - \tau_{\text{ATT}}(s,t) \xrightarrow{p} 0$$

- **Correct factors** $(F_t^o = F_t)$ $\Rightarrow$ Abnormal-returns estimator is consistent:

$$\tau^{AR}(s,t) - \tau_{\text{ATT}}(s,t) \xrightarrow{p} 0$$

- **Synthetic control** + **Gsynth** unbiased under linear factor model
  - Synthetic control constructs *replicating* portfolio - tradable

Large-sample limits with staggered events

Assume $n_s, n_c, T \to \infty$ and each date in $\mathcal{S}$ has non-trivial treatment probability. Then

$$\widehat{\theta}_\kappa^{AR} - \theta_\kappa^{ATT} \xrightarrow{p} \mathbb{E}\left[(\alpha_s - \tilde{\alpha}_s) + (\beta_s F_{s+\kappa} - \tilde{\beta}_s F^o_{s+\kappa}) \mid T_i \in S\right]$$

$$\widehat{\theta}_\kappa^{\text{cont}} - \theta_\kappa^{ATT} \xrightarrow{p} \mathbb{E}\left[(\alpha_s - \alpha_\infty) + (\beta_s - \beta_\infty) F_{s+\kappa} \mid T_i \in S\right]$$

$$\widehat{\theta}_\kappa^{\text{gsynth}} - \theta_\kappa^{ATT} \xrightarrow{p} 0 \qquad \widehat{\theta}_\kappa^{\text{SC}} - \theta_\kappa^{ATT} \xrightarrow{p} 0$$

- AR and diff-in-mean inherit factor-model/timing bias
- Synthetic control and gsynth remain unbiased

What additional assumptions help?

1. **Random assignment** $\widehat{\theta}_\kappa^{\text{cont}} \xrightarrow{p} \theta_\kappa^{ATT}$ even with fixed T.
2. **Random timing** gives closed-form bias expressions:

$$\widehat{\theta}_\kappa^{AR} - \theta_\kappa^{ATT} = E[(\alpha_s - \tilde{\alpha}_s)|T_i \in S] + \beta_s E[F_t]\tilde{\beta}_s E[F_{s+\kappa}^o]$$

3. If the reported factors are correct $(F_t^o = F_t)$, AR is consistent

Practical implications of bias

- Bias that is " negligible" day-by-day *compounds* over long horizons
  - If daily avg. factor premium is 0.02 percent $\rightarrow$ 250 day period, avg of 5 percent
- Random timing averages out factor *realizations*, not factor *premia*; misspecification still matters for horizons where $E[F_t] \neq 0$.
- Synthetic control or gsynth is the safest route for long-run event studies
  - Stable factor model is a strong assumption over long horizon (Kelly, Pruitt, Su (2019))
  - Clear evidence of shifting loadings in empirical examples

Empirical example 1: Acemoglu, Johnson, Kermani, Kwak, Mitton

- Acemoglu et al. (2016) study how the leak of Timothy Geithner's nomination as U.S. Treasury Secretary on Nov 21, 2008 affected firms connected to him
  - Focus on pooled average treatment effect (ATT) for five methods: abnormal returns, synthetic control, gsynth and synth did
- Paper compares *within* banks connected vs. not, we expand control group
- Key features:
  - Single event
  - Unusual timing
  - Short horizon

Contents lists available at ScienceDirect

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/jfec

The value of connections in turbulent times: Evidence from the United States☆

Daron Acemoglu[a], Simon Johnson[b], Amir Kermani[c], James Kwak[d], Todd Mitton[e,*]

[a] MIT Department of Economics, 77 Massachusetts Avenue Building E52, Room 446, Cambridge, MA 02142-1347, United States
[b] MIT Sloan School of Management, 50 Memorial Drive E62-420, Cambridge, MA 02142, United States
[c] Haas School of Business, University of California, Berkeley 545 Student Services Building #1900, Berkeley, CA 94720-1900, United States
[d] University of Connecticut School of Law, 55 Elizabeth Street Hosmer 118, Hartford, CT 06105-2290, United States
[e] Brigham Young University, Marriott School of Management, 640 TNRB, Provo, UT 84602, United States

ARTICLE INFO

Article history:
Received 28 May 2015
Revised 14 September 2015
Accepted 12 October 2015
Available online 27 April 2016

JEL Classification:
G01
G14
G21
G28

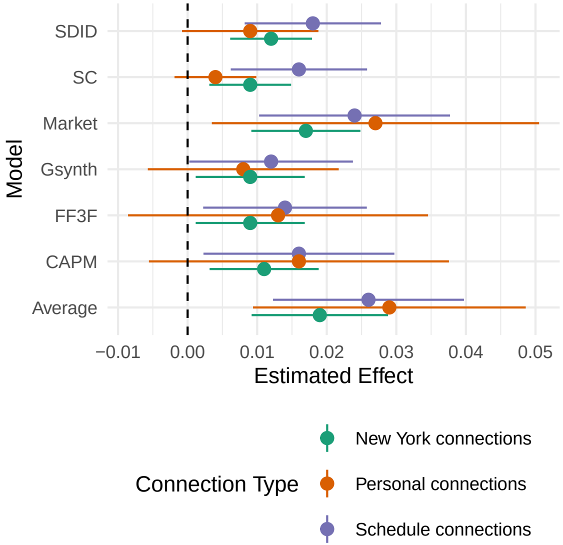
Keywords:
Political connections
Economic crises
Institutions

ABSTRACT

The announcement of Timothy Geithner as nominee for Treasury Secretary in November 2008 produced a cumulative abnormal return for financial firms with which he had a prior connection. This return was about 6% after the first full day of trading and about 12% after ten trading days. There were subsequently abnormal negative returns for connected firms when news broke that Geithner's confirmation might be derailed by tax issues. Personal connections to top executive branch officials can matter greatly even in a country with strong overall institutions, at least during a time of acute financial crisis and heightened policy discretion.

Results are much closer to zero using synthetic methods

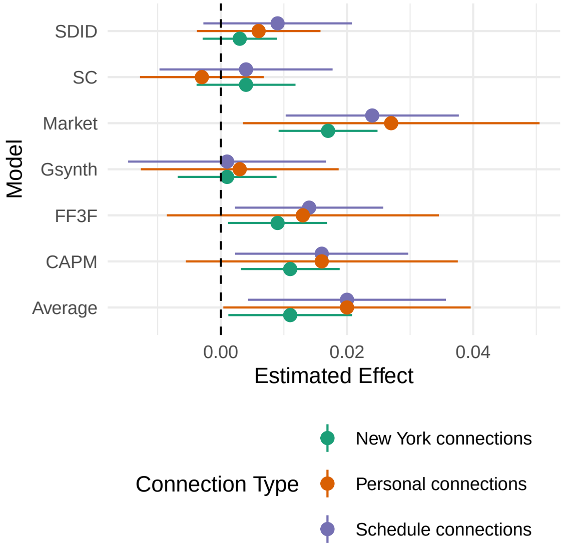
- Effect sizes are large in the paper, shrink significantly with synthetic approach, even when focused on just banks as controls



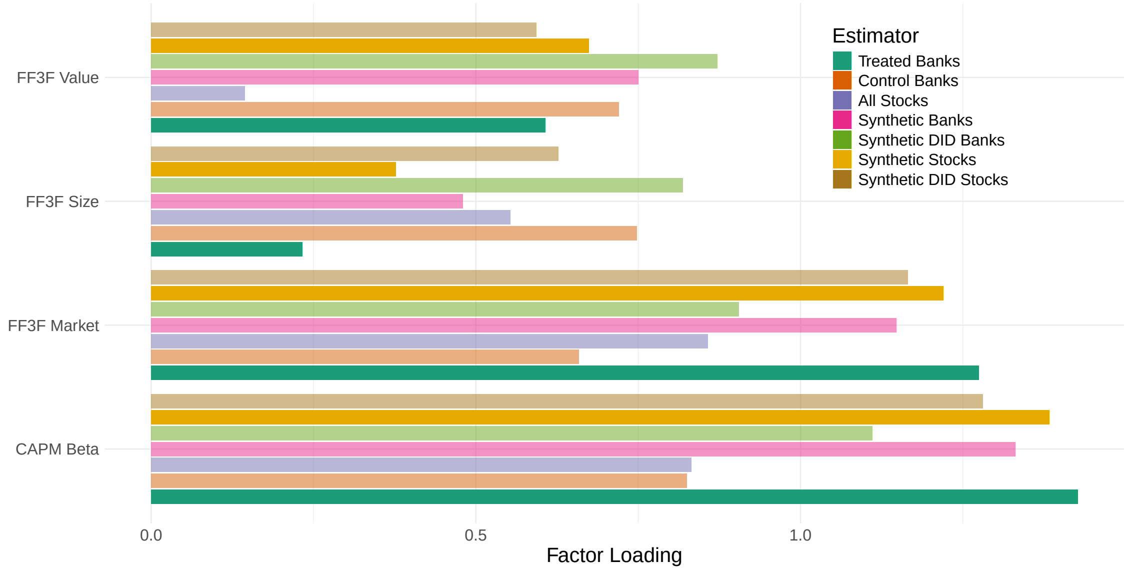
Results are much closer to zero using synthetic methods

- Effect sizes are large in the paper, shrink significantly with synthetic approach, even when focused on just banks as controls

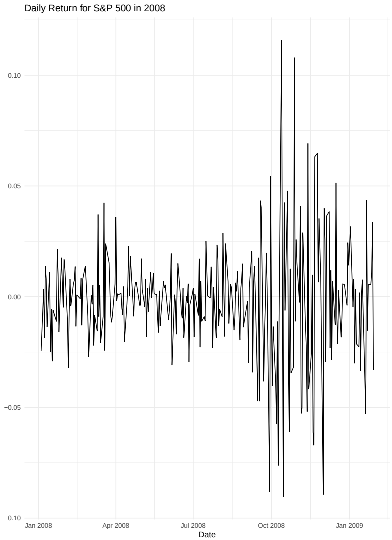
- When expanded to the full universe of control firms, all estimated effects are effectively zero
  - Why?

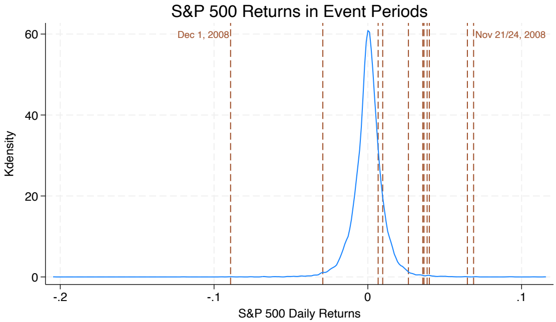


Reason 1: differences in factors loadings



Reason 2: non-random timing

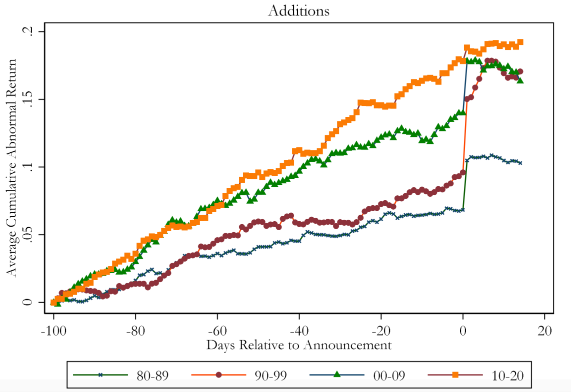




- Timing of event is correlated with significant risk factors

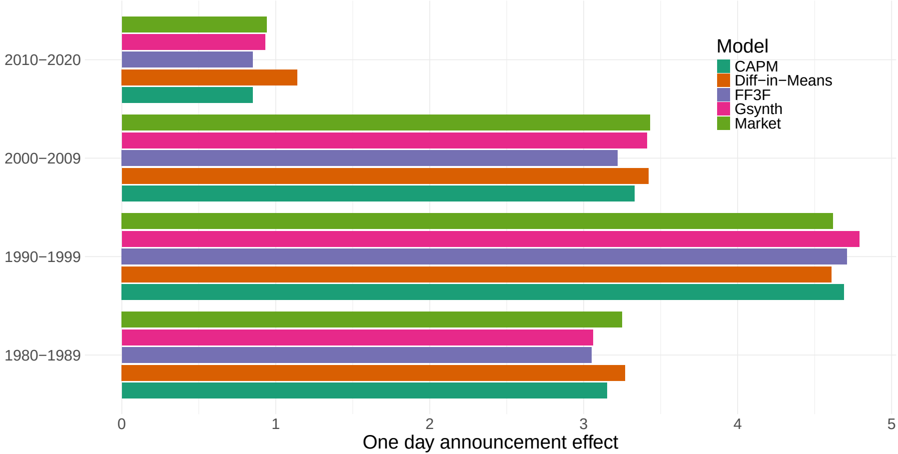
Empirical example 2: S&P Index Inclusion Effect

- S&P 500 index inclusion effect: firms added to the index experience a large positive return on the day of inclusion
- Replicate analysis from Greenwood and Sammon (2025)
  - S&P inclusions from 1976-2020
  - Use announcement dates from Siblis Research, if missing, use day prior to effective day
- Key features:
  - Many events
  - As-if random timing
  - Short- and long- horizon

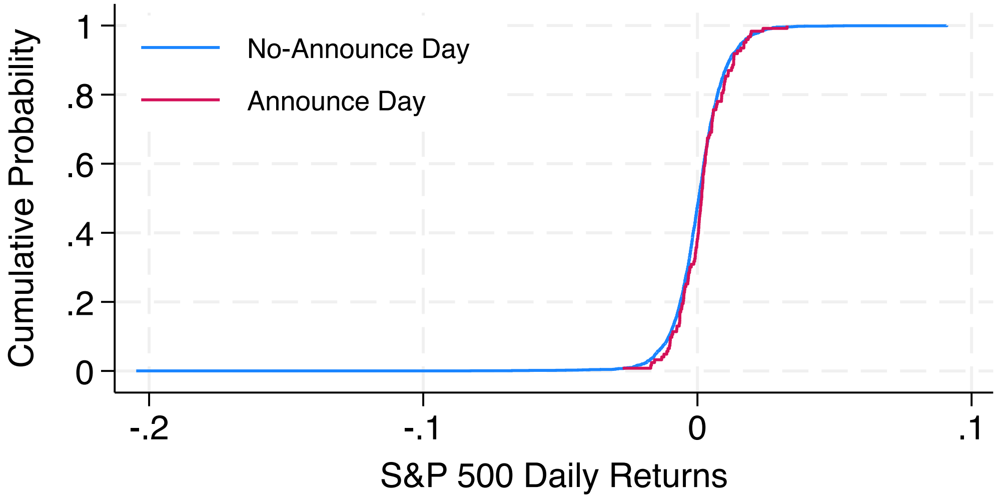


Empirical example 2: S&P Index Inclusion Effect

- Method for short-run estimation does not matter

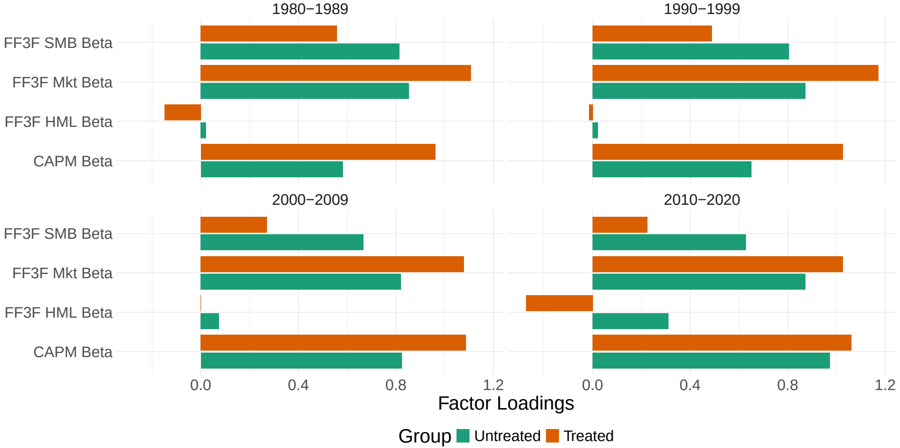


One-day event effect is roughly consistent because of random timing



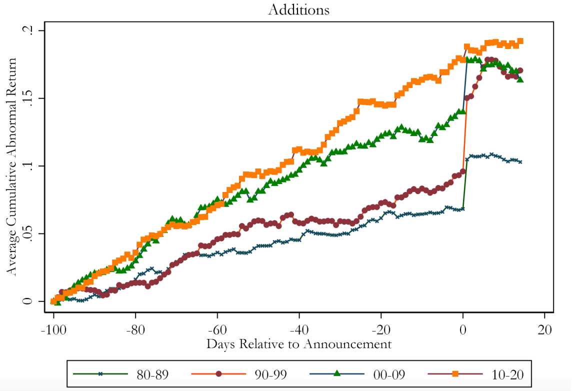
However, not randomly assigned to firms

- Treated firms are significantly different than untreated firms

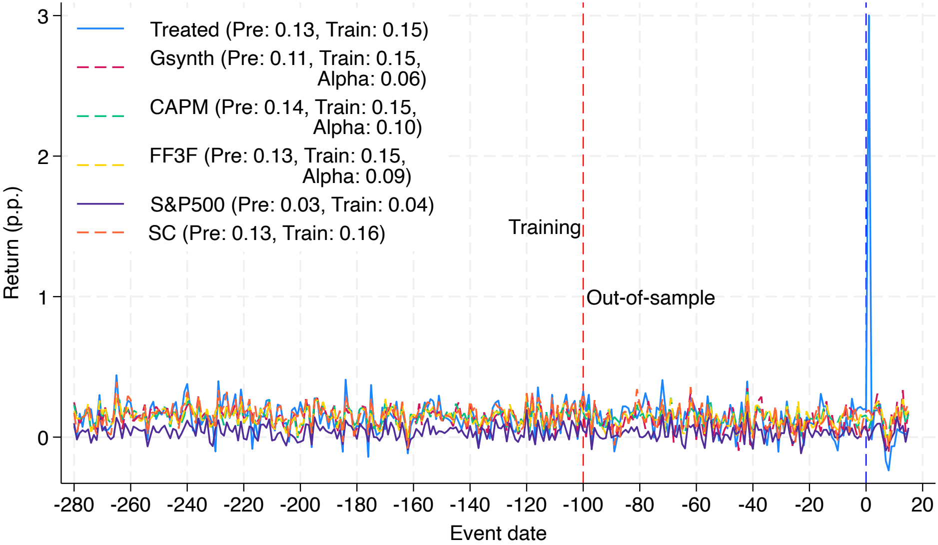


Pre-inclusion drift as a long-run effect

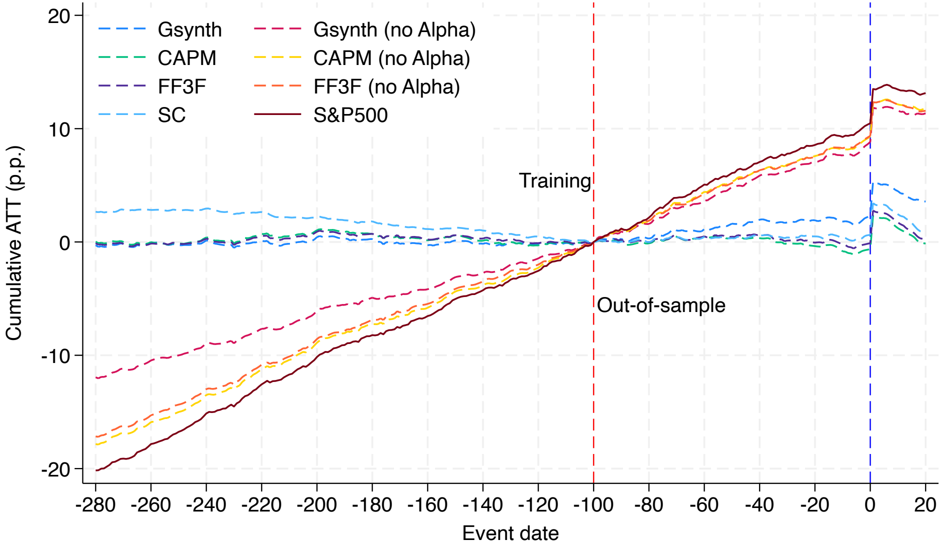
- We study the pre-inclusion "drift" as a form of long-run bias
- Often, drift is pointed to as a puzzle, evidence of potential front-running, or other market activity



Per-Period ATTs for Index Inclusion

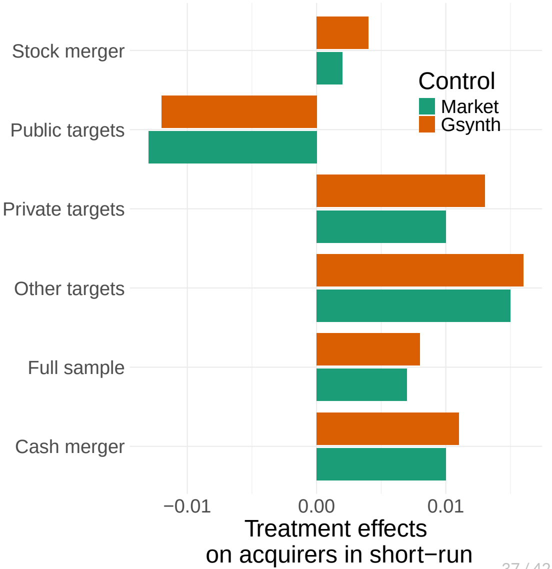


Factor methods remove trends prior to index inclusion

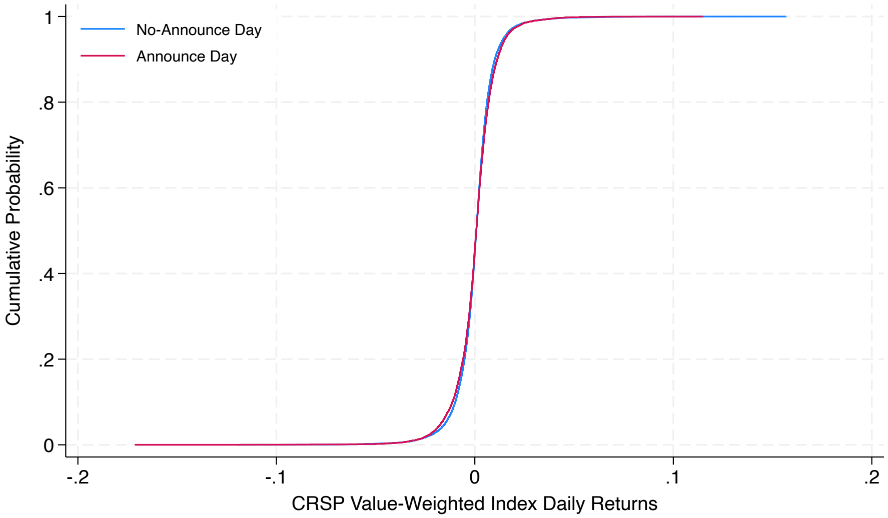


Empirical example 3: M&A (Malmandier (2017))

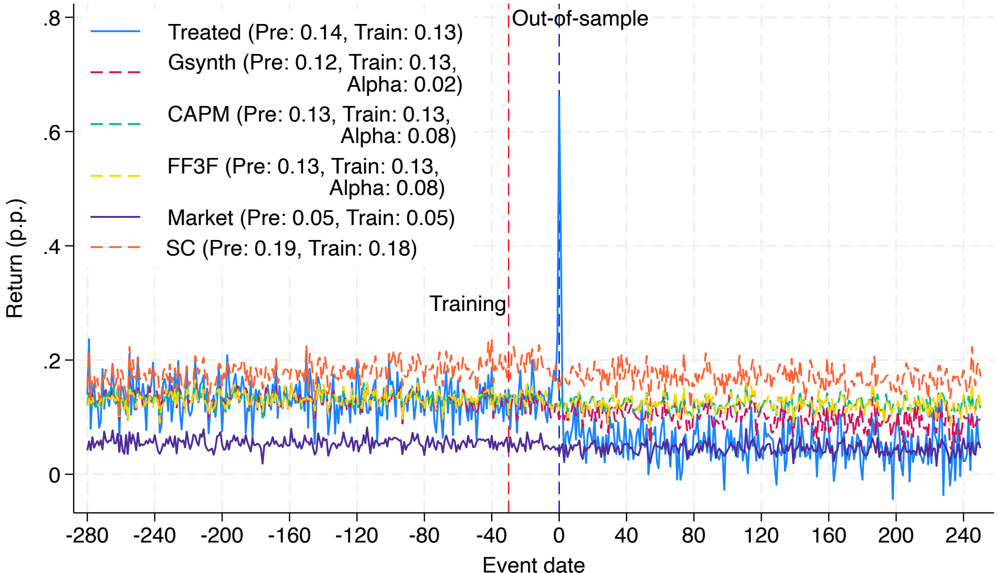
- What is impact of acquisitions on *acquirer* returns?
- Replicate analysis from Malmandier (2017)?
  - All acquisitions in SDC from 1980-20204
- Key features:
  - Many events
  - Quasi-random timing
  - Short- and long- horizon
- Short-run effects quite similar, but...



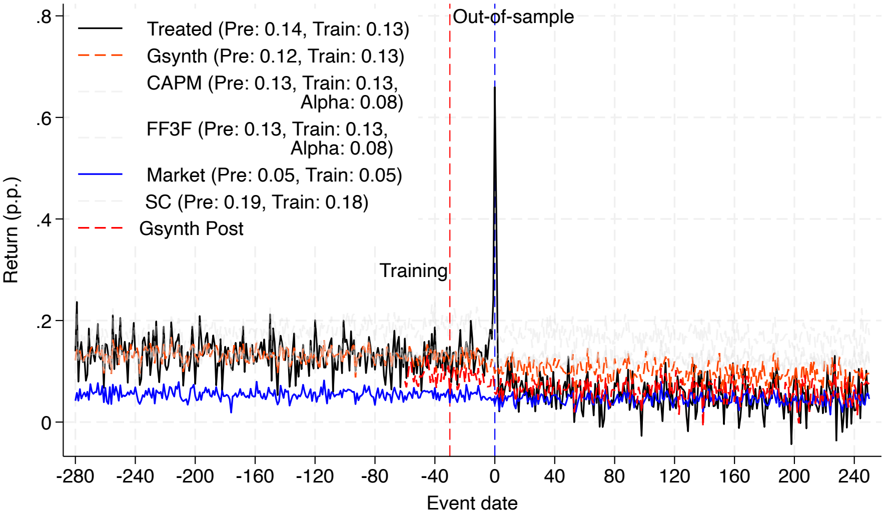
Short-term effects are similar because of random timing



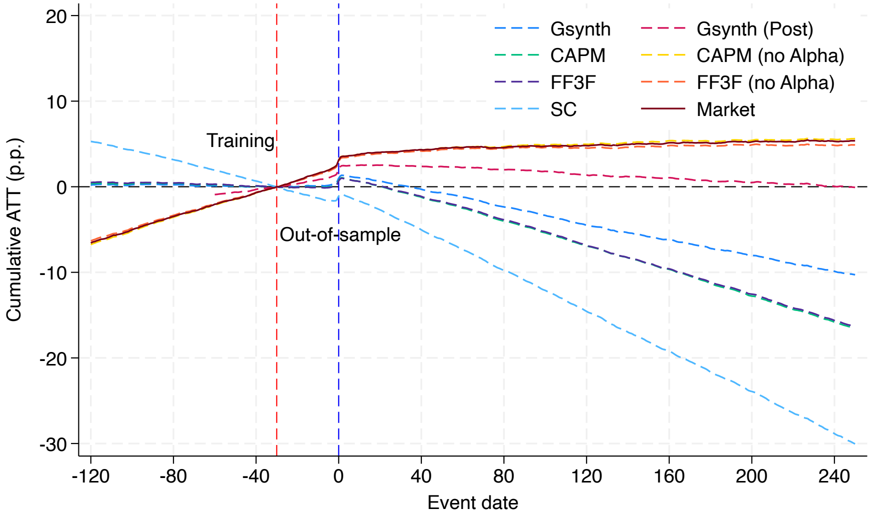
Per-Period ATTs for M&A



Per-Period ATTs for M&A (Gsynth Pre vs Post Period)



Synthetic methods match on overpricing in pre-period, leading to negative post M&A returns



Take-home messages

- Positive results in short-run are consistent with folk knowledge of event studies

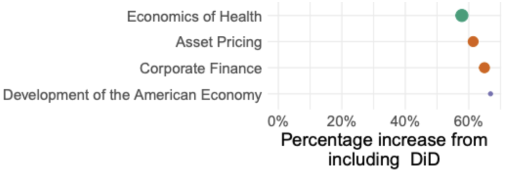
  The results were not materially different when returns were not corrected for market movements. [Shleifer (1986)]

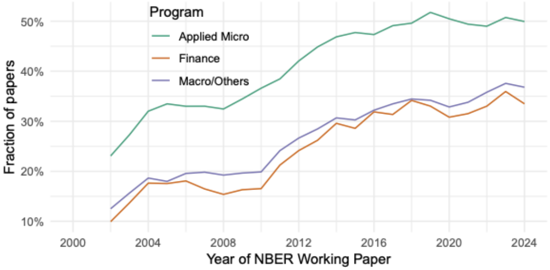
- Short-run estimates work well under random timing

- Long-run estimates needs a careful counterfactual model
  - Best to use synth or gsynth with many firms

Causal inference in finance as an agenda

- These are issues that show up for panel data studies using difference-in-differences!
  - Asset prices incorporate information much faster than other economic outcomes
- Finance has lagged behind many other econ fields in causal inference tools, but we have a powerful set of outcomes and experiments that other fields do not
  - Financial event studies can be important tool for this!





(a) Identification

Goldsmith-Pinkham (2024)