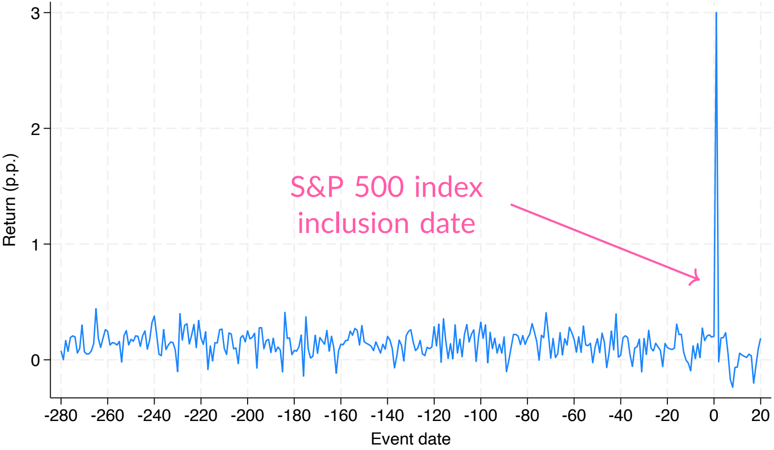
Causal Inference in Financial Event Studies

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

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

Tying together two literatures, and extending an old debate

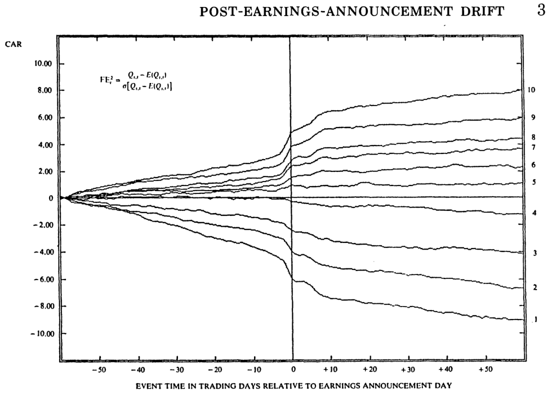
- Finance literature studying the impact of events on asset prices
- Econometrics literature estimating causal effects



Historically, financial event studies are an important tool

What types of financial events? Examples...

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



Historically, financial event studies are an important tool

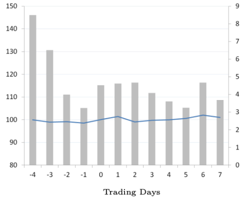
What types of financial events? Examples...

- Index Inclusion
- Earnings Announcements
- 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."

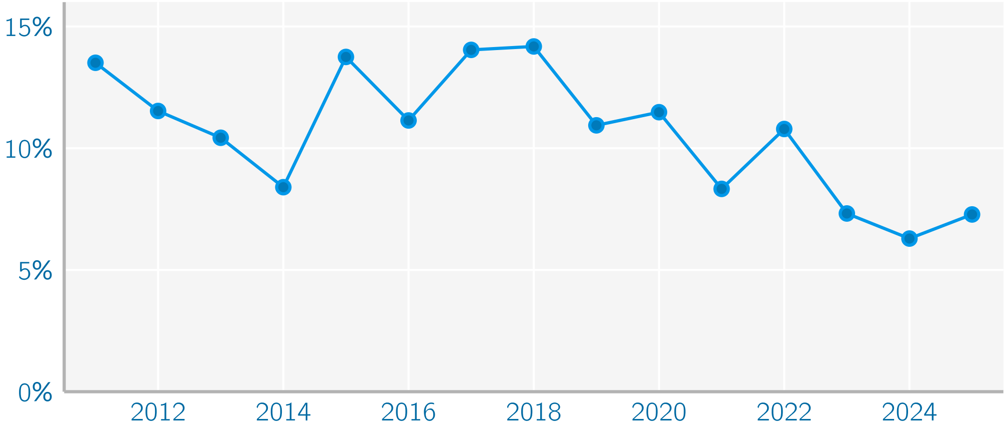
(c) Patent 6,317,722 granted to Amazon.com on Nov 13, 2001, "Use Of Electronic Shopping Carts To Generate Personal Recommendations."



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

Financial event studies continue to be used!

% of JF/JFE/RFS articles



Source: https://paulgp.com/econlit-pipeline/search.html search of JF, JFE, RFS articles mentioning “cumulative abnormal returns” or “announcement returns”

Contribution of this paper

- Reframe event studies as a causal inference problem
  - What is the counterfactual return?
- Characterize when abnormal return estimators are biased
  - Short-run, many events: usually fine
  - Short-run, one event: depends on volatility
  - Long-run: almost always biased
- Propose alternative estimators
  - Synthetic control, PCA regression (GSynth)

- Four applications:
  1. Geithner / political connections
  2. S&P 500 Index Inclusion
  3. Mergers and acquisitions
  4. Close merger contests
- Revisiting established findings, we show some may reflect model misspecification rather than true treatment effects

Roadmap

1. The problem: abnormal returns and misspecification

2. Estimators and when bias matters

3. Four applications:
   - Geithner: single event, extreme volatility
   - S&P 500 Index Inclusion: many events, random timing
   - M&A: many events, long-run bias
   - Close contests: quasi-experimental benchmark

4. What should researchers do?

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

- Unit of analysis: a path of stock returns $R_i = \{R_{i1}, \dots, R_{iT}\}$
  - 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:

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

- Note: assumes SUTVA — one firm's event does not affect other firms' returns. Ignores aggregate spillovers (e.g. index rebalancing, information contagion). Key consideration for future work.

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?

$$
\begin{aligned}
\widehat{AR}_{it} &= R_{it} - \hat{\alpha}_i - \hat{\beta}_{i,1} F_{1,t} \\
&= [\alpha_i - \hat{\alpha}_i] + [\beta_{i,1} F_{1,t} - \hat{\beta}_{i,1} F_{1,t}] + \beta_{i,2} F_{2,t} + \tau_t + \varepsilon_{i,t}
\end{aligned}
$$

Misspecification in the abnormal return estimator

- What happens when a factor is omitted?

$$
\begin{aligned}
\widehat{AR}_{it} &= R_{it} - \hat{\alpha}_i - \hat{\beta}_{i,1} F_{1,t} \\
&= [\alpha_i - \hat{\alpha}_i] + [\beta_{i,1} F_{1,t} - \hat{\beta}_{i,1} F_{1,t}] + \beta_{i,2} F_{2,t} + \tau_t + \varepsilon_{i,t}
\end{aligned}
$$

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

where $\overline{\beta}_2 = n_1^{-1} \sum_{i: T_i = s} \beta_{i,2}$ and $(\bar{\alpha} - \bar{\hat{\alpha}})$ is the intercept estimation error

The short-and-long consequences of misspecification

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

- 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
- Intercept error: OLS absorbs mean omitted factor premium into $\hat{\alpha}$
  - $(\bar{\alpha} - \bar{\tilde{\alpha}}) = -\overline{\beta}_2(E[F_2 \mid \text{pre}])$ – constant, typically small
  - **On average, this is equal to the other piece of misspecification error!**
- Loading error scales with $F_{2,t}$: if τ_t is large relative, bias is second order

Bias is amplified during volatile periods

- For a single event, bias from misspecification is:

$$|\text{Bias}| = |a + b \cdot F_t|$$

  where $a = \alpha_s - \tilde{\alpha}_s$ (intercept error) and $b = \beta_s - \tilde{\beta}_s$ (loading error)

- Bias is monotonically increasing in $|F_t|$ for large $|F_t|$
  - Small loading misspecification $\times$ extreme factor realizations $=$ large bias
  - This is precisely when events often occur! (financial crises, policy announcements)

- Example: $b = 0.6$ (treated vs. control beta gap), $F_t = 6.9\%$ (Nov 21, 2008)
  - Intercept error: $a = -b \cdot \bar{F}^{\text{pre}} \approx -0.6 \times 0.04\% = -0.02\%$
  - Full bias: $|-0.02\% + 0.6 \times 6.9\%| \approx 4.1\%$ – loading term dominates

Aside: asset pricing models $\neq$ counterfactual prediction

- AP models (CAPM, FF3, etc.) predict *expected* returns: compensation for priced risk
  - With many random events, realized factor returns average out $\rightarrow$ model "works"
- But counterfactual prediction requires matching *contemporaneous* returns
  - Contemporaneous returns include both priced *and unpriced* risks
  - A firm's return today depends on factors that are *not* compensated in expectation
- This is the fundamental disconnect:
  - $E[R_{it}]$ can be well-approximated by CAPM/FF3 $\rightarrow$ works for many-event averages
  - R_{it} on a *specific day* depends on unpriced factors too $\rightarrow$ fails for single events or volatile periods
- Synthetic control / gsynth sidestep this by matching on *realized* return paths, not expected return models

Linear factor model assumption for theory results

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)
- Could generalize factor loadings β_s to account for time-varying characteristics (Daniel, Mota, Rottke, and Santos (2020)), Kelly, Pruitt, and Su (2019)

Four approaches to building a counterfactual return

Estimator	Counterfactual	Key assumption
Abnormal returns (CAPM, FF3)	$\hat{\alpha}_i + \hat{\beta}_i F_t^o$ estimated pre-event	Researcher specifies correct factors
Difference-in-means	Average return of control stocks	Treated and control have same avg. loadings
Synthetic control	Weighted portfolio of controls matching pre-event path	Replicating portfolio exists (convex weights)
Gsynth / PCA	PCA-estimated factors + pre-event loadings	Linear factor structure (factors learned from data)

- First two: researcher specifies the model. Last two: data-driven
- Synthetic control and gsynth do not require knowing which factors matter

Why synthetic methods help: they match on realized paths

- Abnormal return estimators specify F_t^o and estimate $\hat{\beta}_i$
  - If factors are wrong or loadings are wrong $\rightarrow$ bias proportional to $(\beta_s - \hat{\beta}_s)F_t$
- Synthetic control constructs a *replicating portfolio*:
  - Weighted combination of control stocks that tracks treated stocks pre-event
  - Implicitly matches on *all* factor loadings, not just the ones you specify
  - Under a linear factor model, this exactly recovers the true counterfactual
- Gsynth uses PCA to *learn* the factor structure from control stocks
  - Cross-validates the number of factors
  - Then estimates treated loadings from pre-event data
- GSynth and synth are consistent as n, $T \rightarrow \infty$ formal results

When does bias matter? Three cases

- **Case 1: Many events, random timing, short horizon** ✓ Standard methods work
  - Factor realizations average out across event dates
  - OLS intercept absorbs mean factor premium, canceling bias
  - This is why "the model doesn't matter" (Brown & Warner (1985), Shleifer (1986))
- **Case 2: Single event or extreme volatility** $\times$ Standard methods fail
  - No averaging — bias is $(\beta_s - \hat{\beta}_s) \times F_t$
  - Events during crises: small loading error $\times$ huge factor realization $=$ large bias
- **Case 3: Long horizon** $\times$ Standard methods fail
  - Even "small" daily bias *compounds*: 0.02%/day $\rightarrow$ 5% over one year
  - Factor distributions may shift between pre and post periods (non-stationarity)

Synthetic control and gsynth: consistent in all three cases formal results

Key takeaway from the theory

1. Short-run event studies with many randomly-timed events are robust — the textbook approach works
2. Single-event studies during volatile periods are fragile — use synthetic methods
3. Long-run event studies are almost always biased — even with many events, daily misspecification compounds
4. When in doubt, report synthetic control or gsynth alongside standard methods

What about calendar-time portfolios?

- Calendar-time portfolio (Fama (1998)): form portfolio of firms currently in event window, regress on factors, estimate α
- We show: calendar-time α is a *reweighted* version of event-time ATT
  - Downweights event-clustered periods; coincides with event-time only if event flow is uniform
- Same bias structure as event-time — same three cases apply
- One key distinction: when treatment *changes* factor loadings (e.g., M&A raises beta)
  - Event-time ATT $=$ total return impact
  - Calendar-time $\alpha =$ risk-adjusted anomaly (relative to new loadings)
  - Neither is inherently correct — depends on research question

details

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 (financial crisis)
  - 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

A R T I C L E I N F O

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

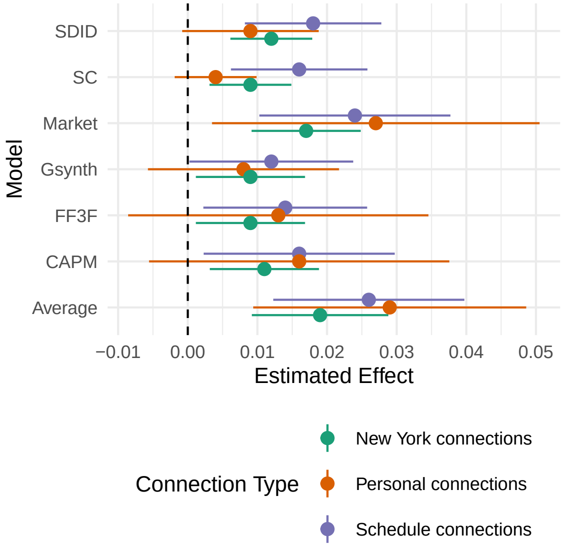
Keywords:
Political connections
Economic crises
Institutions

A B S T R A C T

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.

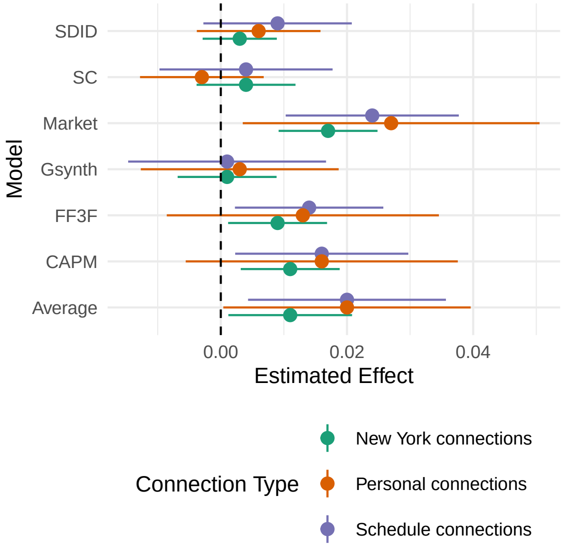
Results are much closer to zero using synthetic methods

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

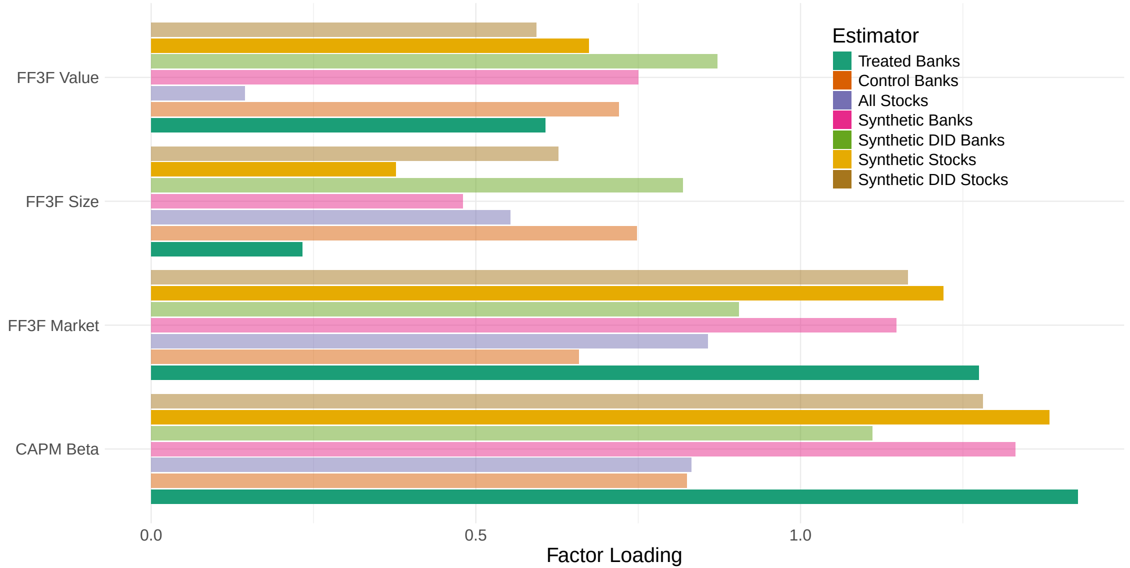


Results are much closer to zero using synthetic methods

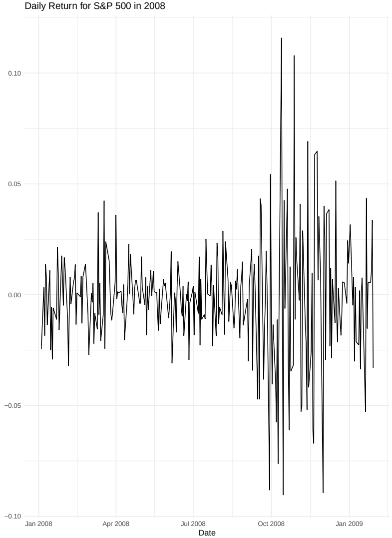
- Effect sizes are large in the paper, shrink significantly with synthetic approach, even when focused on just banks as controls
- When expanded to the full universe of control firms, all estimated effects are effectively zero
  - Why?



Reason 1: differences in factor loadings



Reason 2: non-random timing



- Timing of event is correlated with significant risk factors
- S&P returned +6.6% on Nov 21 – in extreme tail of distribution
- Beta gap of 0.6 $\times$ 6.9% return $\approx$ 4.1% mechanical bias

Geithner takeaway

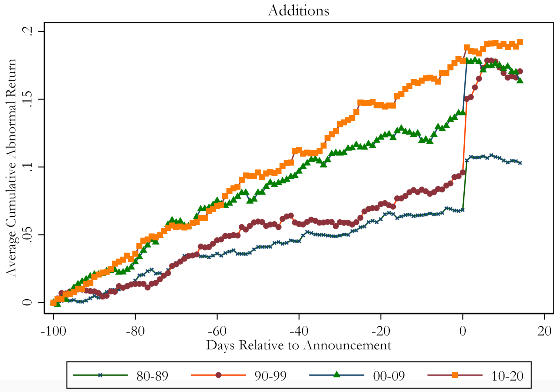
- With a **single event** during **extreme volatility**, the standard approach can generate large spurious effects
- A beta gap of 0.6 $\times$ a 6.9% market day $=$ 4.1% mechanical bias *on one day*
- Synthetic methods eliminate the apparent effect entirely
- For single-event studies,check for balance using factor loading comparisons and synthetic control estimates

Empirical example 2: S&P Index Inclusion Effect

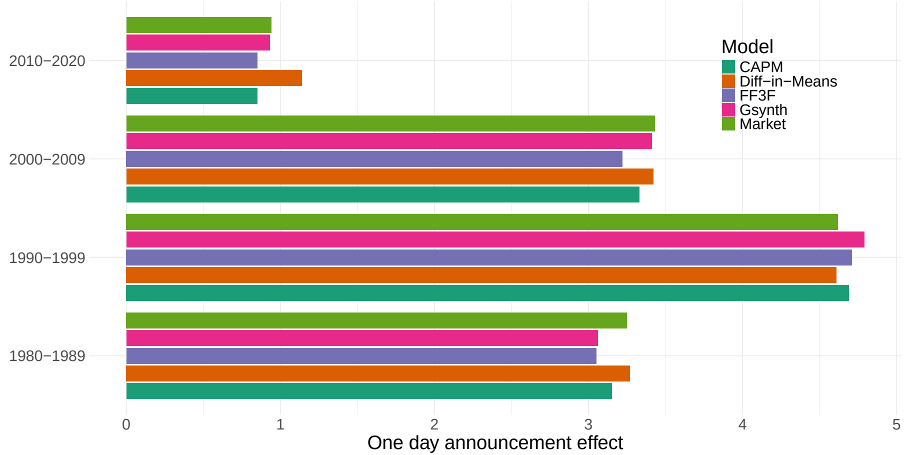
- S&P 500 index inclusion effect: firms added to the index experience a large positive return on the day of inclusion

- 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 ⏩ evidence
  - Short- and long- horizon

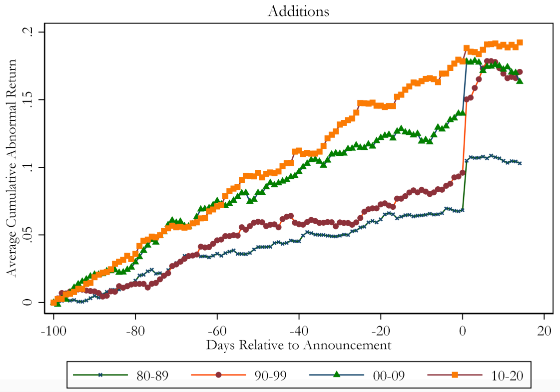


S&P Index Inclusion Effect: Method for short-run estimation does not matter

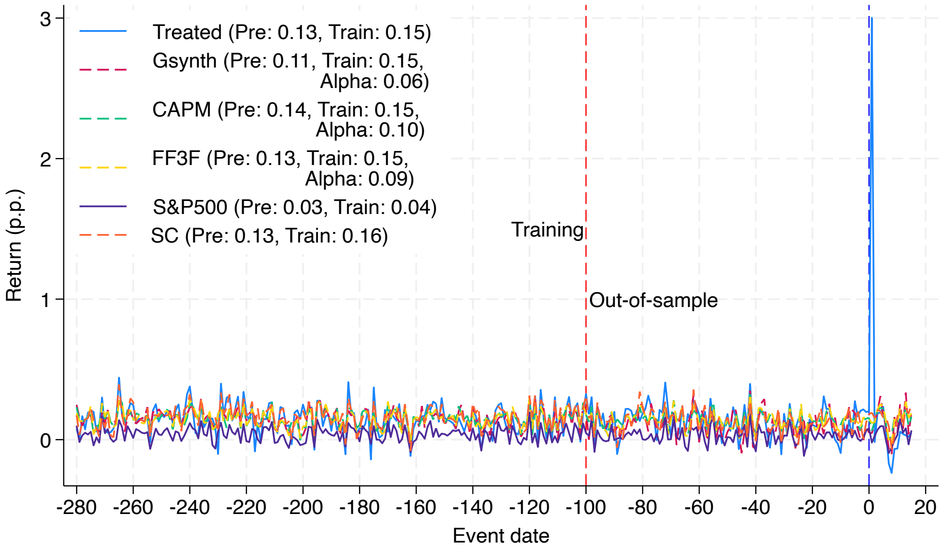


Pre-inclusion drift as a long-run effect

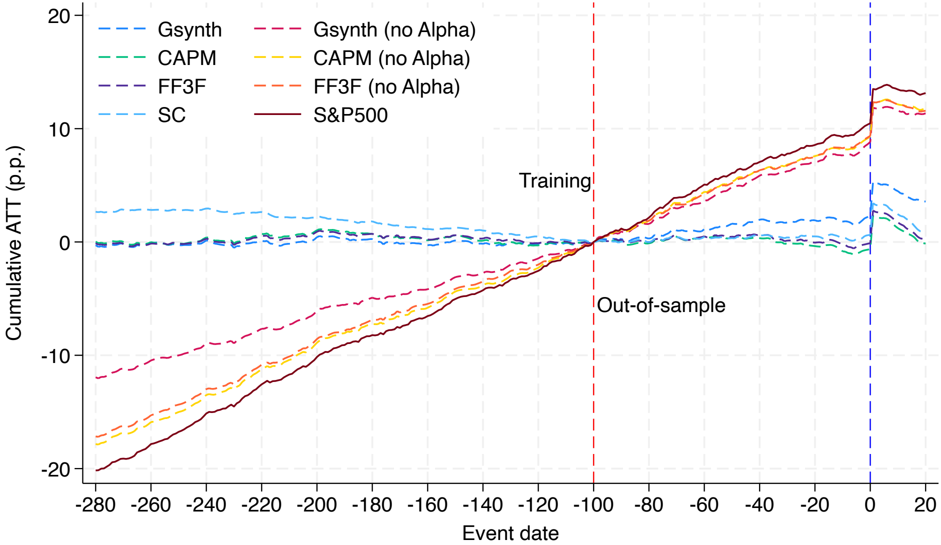
- We study the pre-inclusion "drift" as a form of long-run bias
- Often, drift is pointed to as a puzzle, evidence of potential front-running, or other market activity
- But: firms included in S&P 500 are *different* from other firms
  - Alpha in factor models is large and positive even 280 days pre-event
  - Reflects model misspecification: $\hat{\alpha} = \alpha_{true} + \beta^{unobs} E(F^{unobs})$



Per-Period ATTs for Index Inclusion

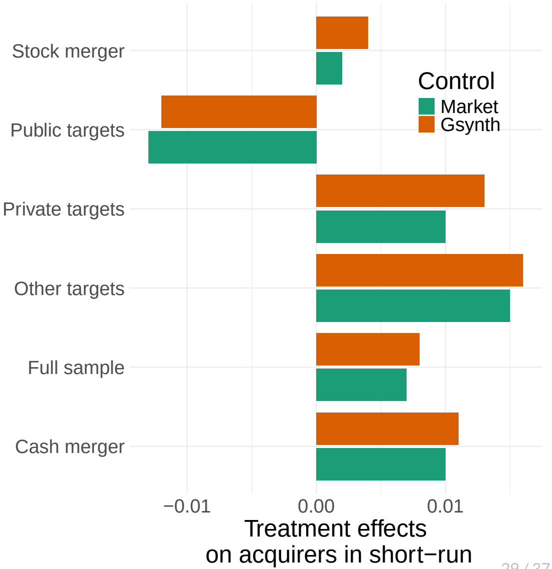


Factor methods remove trends prior to index inclusion

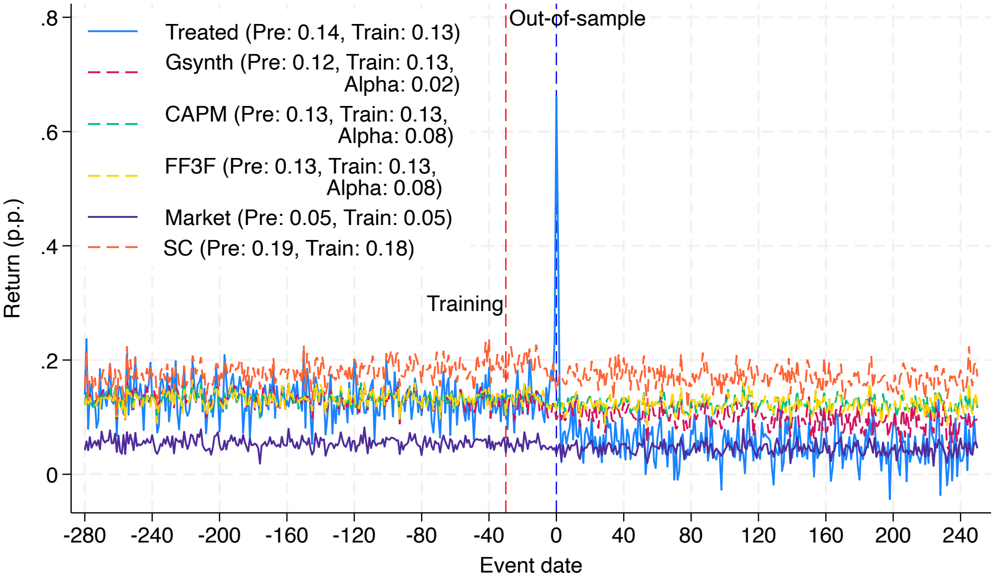


Empirical example 3: M&A

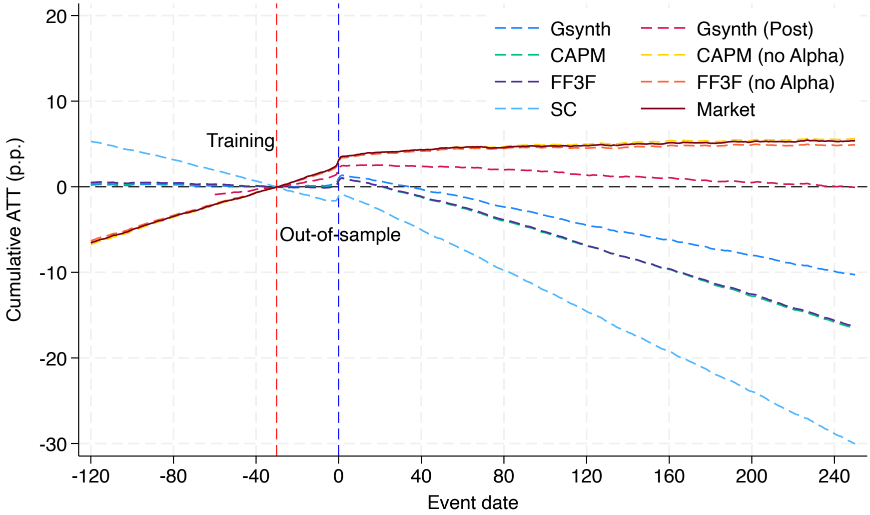
- What is impact of acquisitions on *acquirer* returns?
- Replicate analysis from Malmendier (2017)
  - All acquisitions in SDC from 1980-2024
- Key features:
  - Many events
  - Quasi-random timing ⏩ evidence
  - Short- and long- horizon
- Short-run effects quite similar, but...



Per-Period ATTs for M&A

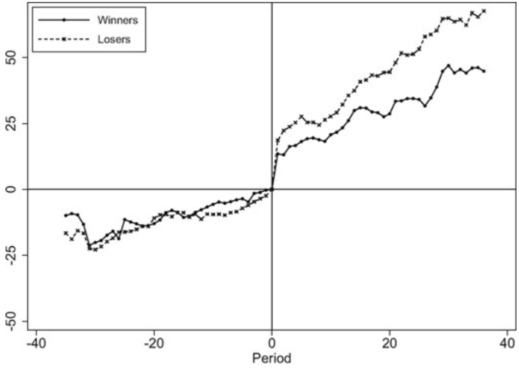


Synthetic methods match on overpricing in pre-period, leading to negative post M&A returns



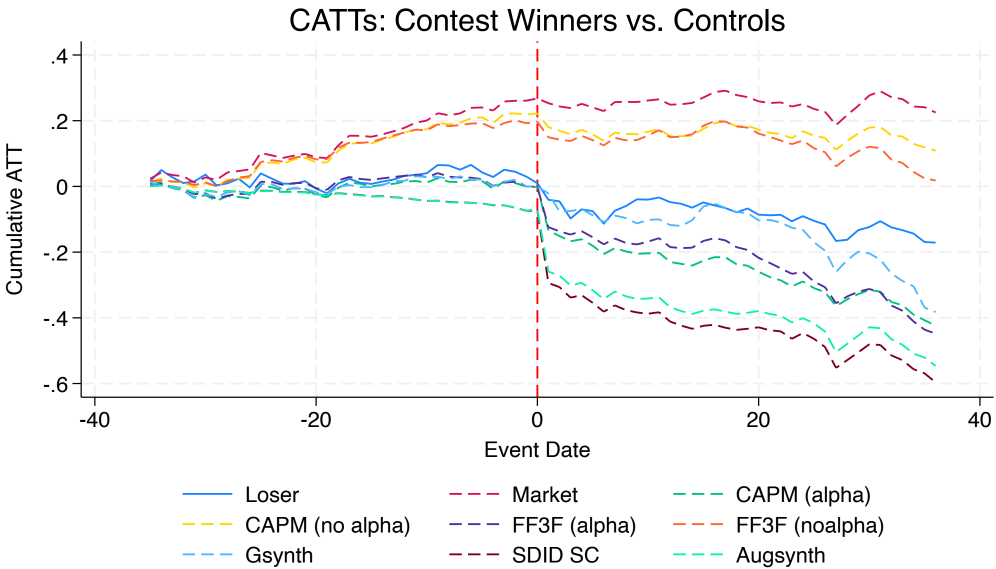
Empirical Example 4: Close merger contests as design-based comparison

- Malmendier, Moretti, and Peters (2018) use “close” mergers — bidding contests
  - Goal is to find as-if coinflip mergers
  - Many events (100 different mergers)
  - Short- and long- horizon
- Striking takeaway in paper is losers doing better
- Compare model based approach to design-based approach
  - LaLonde-style exercise: contest losers serve as quasi-experimental counterfactual



(d) Unadjusted buy-and-hold returns

Design-based approach gives very different answers



Key lessons from the close contest comparison

- The design-based loser counterfactual (benchmark truth) shows:
  - Smallest, least-trending pre-period
  - Small negative post-contest ATT — consistent with mild harm to acquirers
- Model-based approaches:
  - **Gsynth**: closest to loser counterfactual
  - **Synthetic control**: larger declines during and after contest
  - **CAPM, FF3F**: fit well pre-period but predict severe long-run declines not supported by loser comparison
- For long-run analysis: quasi-experimental variation is likely superior to model-based approaches

What should you take away from this paper?

- The "model doesn't matter" folk wisdom is right — but only for short-run, many-event, random-timing studies

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

- It is wrong for:
  - Single events during volatile periods (Geithner)
  - Long-horizon analyses (M&A, pre-inclusion drift)
  - Settings where treated firms have distinctive factor loadings

- The core issue: asset pricing models predict *expected* returns well, but not *contemporaneous* returns — and counterfactual prediction needs the latter

What should researchers do?

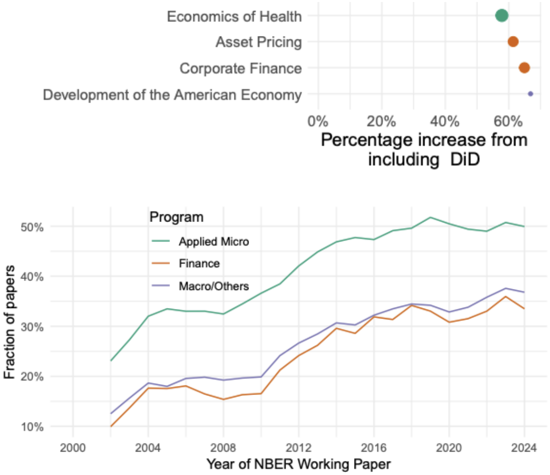
- **Always:** report which estimator you use and why
- **Short-horizon, many events, random timing**: standard methods are fine
  - *Test* the random timing assumption: compare factor return distributions on event vs. non-event days
- **Short-horizon, single event or volatile period**: use synthetic control or gsynth
  - Report loading comparisons between treated and control groups
  - If loadings differ substantially, the standard approach is suspect
- **Long-horizon**: synthetic methods as complement, design-based if possible
  - Daily biases compound — even “small” misspecification matters at 250+ days
  - Close merger contests show design-based approaches outperform all model-based ones
- **Robustness checklist:**
  1. Compare CAPM / FF3 / synthetic control / gsynth side by side
  2. Test random timing (CDF plots of factor returns)
  3. Report pre-event and post-event betas of treated 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!

Thank you!



(a) Identification

Goldsmith-Pinkham (2024)

Formal bias results: single event

As $n_s, n_c, T \to \infty$:

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

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

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

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

- Even with $n_s, n_c, T \to \infty$, AR and DiD are biased if the factor model is wrong
- Synthetic control and gsynth are consistent

Formal bias results: multiple staggered events

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

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

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

Under random timing with stationary factors:

- The OLS intercept absorbs mean factor premium: $\alpha_i - \tilde{\alpha}_i = -\beta_i E[F_t] + \tilde{\beta}_i E[F_t^o]$
- This cancels bias exactly $\to$ theoretical foundation for "the model doesn't matter"

Under random timing with **non-stationary** factors: bias remains $\propto E(\beta_i)(E[F_t \mid \text{post}] - E[F_t \mid \text{pre}])$

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.

ATT building block

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

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

- 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 $\Theta_H^{\text{CATT}} = \sum_{\kappa=0}^{H} \theta_\kappa$.

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*

Buy-and-hold abnormal returns: formal result

- Geometric (BHAR) ATT differs from arithmetic (CAR) ATT:

$$\theta_H^{geo,ATT} \approx \theta_H^{ATT} - \sum_{\kappa=0}^{H} E\left[R_{i,s+\kappa}(\infty)\tau_i(s,s+\kappa) + \frac{1}{2}\tau_i(s,s+\kappa)^2 \mid T_i = s\right]$$

- Even a portfolio with the *correct* expected return is a bad counterfactual for BHAR if it has different variance (volatility drag)

- Recommendation: focus on arithmetic ATT for tractability and unbiasedness

Calendar-time portfolio: formal decomposition

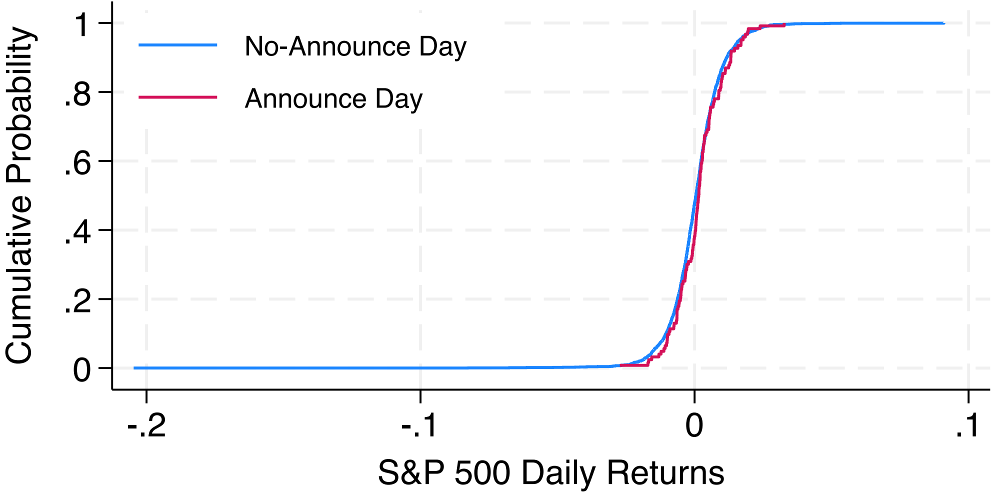
- Calendar-time α decomposes as:

$$\alpha^{cal} = \sum_{s \in \mathcal{S}} \sum_{\kappa=0}^{H} \omega_{s,\kappa}^{cal} \cdot \tau^{ATT}(s, s+\kappa) + \bar{\alpha}^{cal} + \text{bias}^{cal}$$

where $\omega_{s,\kappa}^{cal} = \frac{n_s}{T_{cal} \cdot N_{s+\kappa}}$

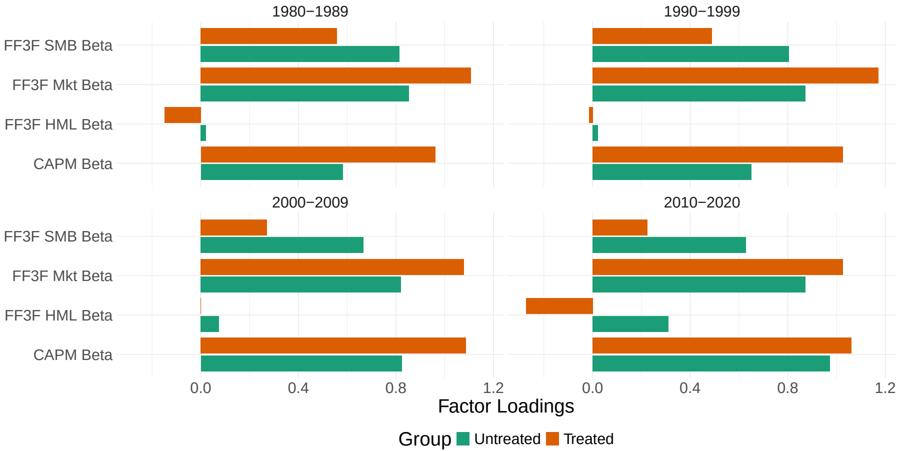
- When treatment changes factor loadings:
  - Event-time: τ^{ATT} includes $(\beta_i^{post} - \beta_i^{pre})F_t$
  - Calendar-time: estimates anomaly relative to *new* risk profile

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

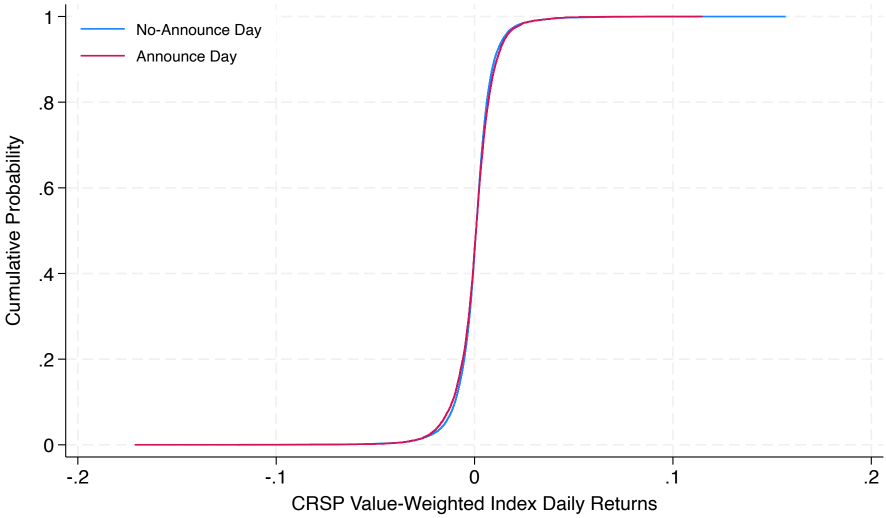


However, not randomly assigned to firms

- Treated firms are significantly different than untreated firms



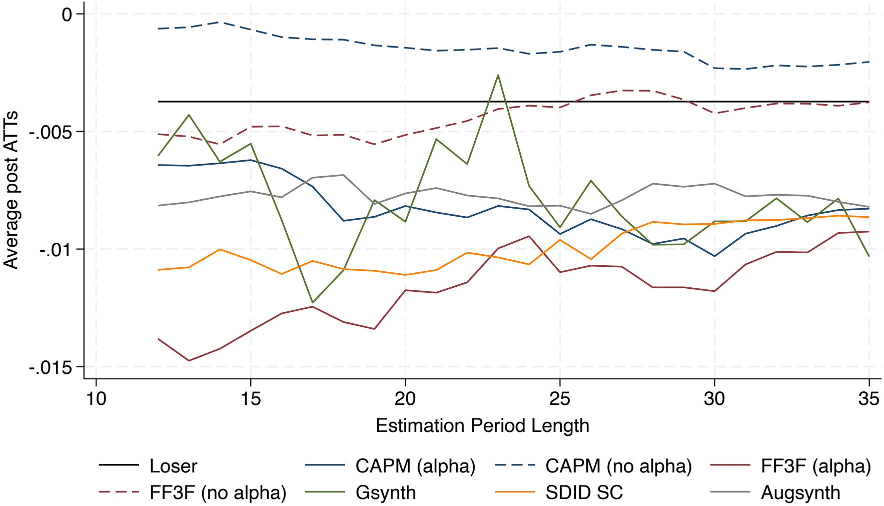
Short-term effects are similar because of random timing



Individual firm estimates: noisy but not necessarily biased

- For a single treated firm, three sources of randomness:
  1. Estimated parameters $(\hat{\alpha}_i, \hat{\beta}_i)$
  2. Factor realizations (F_t)
  3. Idiosyncratic error (ε_{it})
- With many firms, ε_{it} averages away; with one firm it does not
  - Relevant for litigation/damages context (Baker (2020))
- With many individual treatment effects (e.g. patent issuance, Kogan et al. (2017)):
  - Averages follow the same bias patterns as cohort estimates
  - Shrinkage estimators or portfolio groupings can help

Sensitivity to estimation window length (close contests)



Extension: Testing for over- and underreaction

- Common test: regress long-run returns on short-run returns

$$\tau_i^{long} = \alpha + \gamma \cdot \tau_i^{short} + \varepsilon_i$$

- $\gamma = 0$: efficient; $\gamma > 0$: underreaction; $\gamma < 0$: overreaction
- Two problems under misspecification:
  1. Measurement error in $\hat{\tau}^{short}$ attenuates $\hat{\gamma}$ toward zero
  2. Spurious positive correlation from common factor loading errors:

$$Cov(\widehat{CAR}_i^{short}, \widehat{CAR}_i^{long}) = Cov(\tau_i^{short}, \tau_i^{long}) + (\beta_i - \tilde{\beta}_i)^2 \cdot \Sigma_F$$

- Buy-and-hold can generate spurious overreaction: higher-volatility firms $\rightarrow$ larger volatility drag

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