Department of Economics

MSc Extended Essay Cover Sheet

CANDIDATE NUMBER *(This is a **5 digit number** found on LFY. It is **not** your 9 digit student number)*	6	0	0	5	4

COURSE CODE	E	C	4	7	5

TITLE

The Sectoral Effects of Volatility Shocks: A Residual Synthetic Control Analysis of Equity Market Reallocation

WORD COUNT	5	6	5	3

Double spaced, typed. Maximum 6,000 words (as a guideline this is 14-21 double spaced pages). Abstract, footnotes, references and appendices do not count toward the word count, provided such additions are brief and do not contain information that rightly belongs in the body of the essay. Equations are included in the word count and counted as the page equivalent (i.e., as the number of words that would occupy the same amount of space in text).*

Examiners are recommended to penalise excessively long essays (+ 10% of word count) by adjusting the marks by 6000/x, where x is the word count.

ARE YOU REGISTERED WITH THE DISABILITY AND WELL-BEING SERVICE FOR AN INCLUSION PLAN?

Yes

No

DATE OF SUBMISSION:

15/05/2026

Late submissions will be penalised. If a student fails to submit by the set deadline the following penalty will apply: Five marks out of 100 (5%) will be deducted for coursework submitted within 24-hours of the deadline and a further five marks (5%) will be deducted for each subsequent 24-hour period until the course work is submitted.

Table of contents

Abstract 3

1 Introduction 4

2 Literature Review 6

3 Data and Empirical Design 8
- 3.1 Data, Units, and Event Construction . . . 8
- 3.2 Outcomes, Windows, and Identification Logic . . . 10
- 3.3 Frozen Feature Protocol . . . 10
- 3.4 Feature-Based rSCM, SCM, and AugSynth . . . 13

4 Results 14
- 4.1 Event Set, Mean Effects, and Fit Diagnostics . . . 14
- 4.2 rSCM Cross-Sectional Reallocation . . . 15
- 4.3 Cross-Sectional Sensitivity Regression . . . 16
- 4.4 Benchmark Comparison . . . 17
- 4.5 Shock-Channel Checks . . . 19
- 4.6 Robustness Checks . . . 20

5 Discussion 24
- 5.1 Analysis . . . 24
- 5.2 Limitations . . . 25

6 Conclusion 27

7 Appendices 28
- 7.1 Appendix A. Empirical Summary Tables . . . 28
- 7.2 Appendix B. Consistency of Permutation-Importance Screening for a Generic Consistent Learner . . . 30
- 7.3 Appendix C. Consistency of Screened Residual Synthetic Control . . . 34

8 References 41

Abstract

I study whether extreme daily implied-volatility shocks reorganize relative sector performance inside the U.S. equity market. The empirical design uses weekly sector-minus-SPY outcomes for nine U.S. sector ETFs from January 5th 2001 through 2026, with volatility-shock event weeks running through March 6th, 2026. Events are identified from the upper tail of daily $\Delta \log(\text{VIX})$ in the CBOE Volatility Index (VIX), mapped to week-ending Fridays, and spaced at least a quarter apart. In each event-sector fit, the treated sector is compared with a synthetic counterfactual built from the other sector ETFs.

I find that volatility shocks do not produce one fixed ranking of sector winners and losers, but rather reorganize the sector cross-section in recurring ways. The average event-week treatment effect is -0.004%, indicating that the main evidence does not lie in a large pooled mean effect, but rather in the cross-sectional distribution of residualized sector responses. Sector-level averages are therefore descriptive rather than the main inferential object: materials is the clearest weaker sector, energy and financials also lean weaker, while technology, utilities, and consumer staples appear relatively more resilient. Stronger evidence instead comes from within-event sensitivity regressions using the same estimated effects as the dependent variable. A one-standard-deviation increase in pre-event market beta is associated with an event-week residualized ATT that is -0.459% lower, with an event-clustered p-value of 1.77e-06 and a sector-label permutation p-value of 2.00e-04. Market beta, downside beta, and credit-spread vulnerability provide the strongest evidence, while VIX vulnerability points in the same direction but is less stable across inference schemes.

I also develop a feature-based residual synthetic control method (rSCM) for the counterfactual problem this setting creates. A risk-consistent learner estimates the component of returns predictably tied to lagged macro-financial and sector-state variables, and synthetic control is applied to the residual. If the learner removes nothing, the procedure collapses back to the standard synthetic control method (SCM) of Abadie and Gardeazabal (2003). In the 5-percent event sample, mean pre-treatment root mean squared prediction error (pre-RMSPE) is 0.01367 for rSCM, 0.01334 for standard SCM, and 0.01400 for the augmented synthetic control method (AugSynth) of Ben-Michael et al. (2021). Thus, rSCM does not mechanically win on every pre-fit metric, but it is a useful way to construct a conditional cross-sectional treatment-effect object before testing for reallocation.

1 Introduction

In April 2001, during the aftermath of the dot-com collapse, in September 2008 with the failure of Lehman Brothers, in March 2020 at the onset of COVID-19, and again in more recent stress episodes, expected equity-market volatility surged and markets fell sharply as investors rushed to reduce their exposure to risk. The common description of these moments is familiar: panic rises, capital retreats, and asset correlations go to one as the market crashes together.

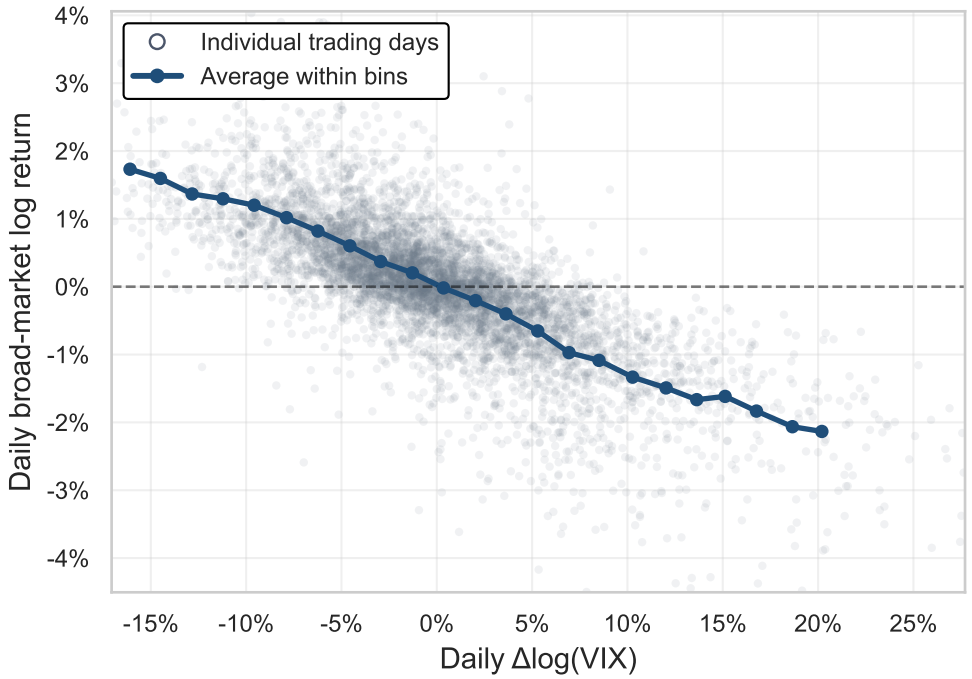

Figure 1: Relationship between volatility shocks and market returns. VIX is the standard index for expected volatility in the broad U.S. equity market over the next 30 days. The binned averages show a strong, nearly monotonic negative association, indicating that larger increases in implied volatility are systematically associated with lower equity returns.

Even in broad selloffs, sectors do not move equally, so I focus on relative sector performance and ask whether leadership changes after the same volatility shock hits every sector. If the shock only raises required returns, the response should look like a common discount-rate shock. If it also tightens financing conditions, hedging demand, or risk-bearing capacity, relative performance should rotate away from more cyclical or financing-sensitive sectors and toward sectors with cash flows investors are more willing to hold under stress. The object of interest is the short-run reordering of sector performance, not the common market drawdown.

The empirical design is a weekly quasi-causal event study built from large daily $\Delta \log(\text{VIX})$ shocks in the CBOE Volatility Index (VIX). Events are defined ex ante, mapped to week-ending Fridays, spaced by 13 weeks, and restricted to the top 5 percent of the historical shock distribution. I use relative sector outcomes rather than raw returns, because raw post-shock returns combine broad market repricing with the cross-sectional sector reallocation studied here. The shock is identified at the daily frequency but estimated in a weekly panel to capture the sector adjustment that follows the volatility event rather than the intraday shock itself.

I estimate volatility shocks' relative sector effects with a residual synthetic control method (rSCM) framework that conditions on nonlinear market state variables. In rSCM, a small set of primitive state variables is expanded into theoretically motivated transformations, a nonlinear learner removes the predictable component of returns, and synthetic control (SCM) is then applied to the residual path. My implementation here uses a Random Forest learner, although the framework extends to any risk-consistent estimator and the appendix proofs are stated at that level of generality.

In canonical SCM applications, untreated outcomes can often be approximated directly by a convex combination of donor paths, so matching in levels is sufficient and covariates typically enter only through the matching objective. Here, however, observable state variables explain part of market-relative sector returns even after subtracting the broad market component. Standard SCM can therefore assign donor weights partly to reproduce the prevailing macro-financial state rather than the cross-sectional structure of interest. I use two popular frameworks as my baseline for comparison: standard SCM, or the canonical synthetic control estimator introduced by Abadie & Gardeazabal, 2003, and the augmented synthetic control method (AugSynth) of Ben-Michael et al., 2021.

rSCM first removes the component of returns predicted by observed state variables and then applies SCM to the residual path; if the first stage removes nothing, the procedure collapses to standard SCM. The remaining residual captures cross-sectional structure not explained by the observed macro-financial state, including sector-specific sensitivities and co-movement patterns. This decomposition is useful in financial markets, where state-dependent return dynamics are difficult to parameterize even when relative sector relationships are observable.

I use rSCM to construct residualized event-sector treatment effects, sector rankings, and weaker-side frequency measures, then estimate event fixed-effect sensitivity regressions relating those outcomes to pre-event exposure to market downside, volatility, and credit-spread conditions. The main distinction is between broad market repricing and cross-sectional sector reallocation, because a volatility shock can lower equity prices across the market through discount-rate adjustment alone or shift relative performance away from more financing-sensitive sectors toward sectors investors are more willing to hold under stress.

This implies two empirical expectations. First, volatility shocks should produce recurring lower-tail sector ordering even if the pooled mean effect is small. Second, residualizing common state dependence before SCM should change the counterfactual object relative to raw relative returns. I therefore study whether severe implied-volatility shocks reorganize the sector cross-section rather than merely generating a common market drawdown, and whether rSCM is a useful way to isolate that residual cross-sectional structure.

The empirical contribution is to show that volatility shocks induce systematic cross-sectional reallocation that is not visible in pooled mean effects but is recoverable through within-event and distributional inference. The methodological contribution is to show why residualizing lagged macro-financial and sector-state variables before SCM can be useful when the object of interest is cross-sectional treatment heterogeneity in an environment with strong state-dependent comovement.

2 Literature Review

Large VIX moves matter because they mark bad states, not statistical noise. Bloom, 2009, Bloom et al., 2018, and Jurado et al., 2015 show that uncertainty shocks coincide with sharp changes in firms' and households' decisions, contractions in activity, and wider, less predictable distributions of future outcomes. Bekaert et al., 2013 and Adrian et al., 2014 make the same point through asset prices, where VIX-like state variables move with risk appetite, expected return variation, and compensation for bearing bad states. Fama & French, 1989, Cochrane, 2011, Stock & Watson, 2012, and Jorda et al., 2013 all push the same conclusion from a different angle, since expected returns, covariances, credit conditions, and leverage all vary materially with business conditions. Altogether, this literature yields the central logic here, that a large daily $\Delta \log(\text{VIX})$ move is a credible marker of a macro-financial stress state.

These large volatility shocks hit the whole market at once, but they need not leave every part of the market in the same relative position. Bernanke et al., 1999 and Kiyotaki & Moore, 1997 provide the classic credit-frictions version of that argument, while Brunnermeier & Pedersen, 2009, Adrian & Shin, 2010, Shin, 2010, He & Krishnamurthy, 2013, and Garleanu & Pedersen, 2011 show more directly how financing capacity, collateral values, margin requirements, and intermediary risk-bearing ability move unevenly across borrowers and assets. Credit spreads matter in exactly this setting because they compress several of those forces into observable prices (Merton, 1974; Duffie & Singleton, 2003; Elton et al., 2001; Collin-Dufresne et al., 2001; Chen et al., 2009; Gertler & Lown, 1999; Gilchrist et al., 2009; Gilchrist & Zakrajsek, 2012; Philippon, 2009). Bernanke & Kuttner, 2005, and Chodorow-Reich, 2014 show the same mechanism in real outcomes, where credit disruptions feed through to employment, investment, and risk taking. Thus, a common volatility spike should not imply uniform sector repricing; instead, sectors that are more financing-sensitive, more levered, or more exposed to distress should react differently from sectors whose cash flows are easier to hold through stress (Campbell et al., 2008; Avramov et al., 2009; Friewald et al., 2014).

Existing literature on cross-sectional asset pricing points the same way. Moskowitz & Grinblatt, 1999 show that return persistence is materially organized at the industry level rather than being purely idiosyncratic noise, while Barroso & Santa-Clara, 2015 show that momentum is state dependent enough for its crash risk to be managed with volatility information. The broader point is that expected returns vary with leverage, profitability, investment, and other observable characteristics through state dependence and nonlinear interaction (Frazzini & Pedersen, 2014; Asness et al., 2019; Hou et al., 2015; Kelly et al., 2019; Gu et al., 2020). Distress and systemic vulnerability also vary materially across firms and sectors rather than loading on one common market factor alone (Campbell et al., 2008; Avramov et al., 2009; Friewald et al., 2014; Acharya et al., 2017; Engle et al., 2015; Kelly et al., 2016). The natural implication is that a volatility shock can operate as a common discount-rate shock and as a cross-sectional reallocation shock at the same time.

There is also an important warning in the crisis and contagion literature. Longin & Solnik, 2001, and Forbes & Rigobon, 2002 show why rising correlations in bad markets do not, by themselves, tell us very much about structure inside the cross-section. That warning matters here because the obvious fact that sectors move together during stress is not the claim of interest. The claim of interest is narrower and harder: after removing the broad market movement and conditioning

on observable state, do sectors change relative position in a recurring way? That is why I use sector-minus-SPY performance rather than raw sector returns, and why I frame the evidence in terms of sector rankings, sector-level relative effects, and bottom-tail frequencies.

The methodological side follows from the same problem. Event-study methods impose timing discipline: the shock rule is observable, the window is transparent, and treatment timing is fixed outside sector outcomes. Synthetic control treats the counterfactual as an object to be built rather than assumed, while later extensions broaden that logic through balancing, matrix completion, synthetic difference-in-differences, and augmentation (MacKinlay, 1997; Kothari & Warner, 2007; Sun & Abraham, 2021; Abadie & Gardeazabal, 2003; Abadie et al., 2010; Abadie, 2021; Doudchenko & Imbens, 2016; Xu, 2017; Arkhangelsky et al., 2021; Athey et al., 2021; Ben-Michael et al., 2021). In this application the shock is common, the donor pool is small and correlated, and raw sector-minus-SPY outcomes still contain predictable nonlinear exposure to volatility, credit, rates, lagged performance, and drawdowns. Direct SCM can therefore force donor weights to match common state dependence and residual sector co-movement at once.

That is why I develop rSCM, a residual-centric SCM design that separates the two tasks. A first-stage learner estimates the predictable state-dependent component using donor information, and synthetic control is then applied to the residual path. The object being reconstructed is therefore not the full raw outcome, but the part left after common macro-financial state dependence has been stripped out. Athey & Imbens, 2017, Athey & Imbens, 2019, Chernozhukov et al., 2018, Wager & Athey, 2018, and Gu et al., 2020 make clear why a fixed parametric first stage is hard to defend here, since the relation between sector returns and state variables such as volatility, credit, rates, and lagged performance is almost certainly nonlinear and state dependent. Random Forests are a practical choice because they accommodate nonlinearities and interactions without requiring them to be specified in advance, they admit a consistency route, and they integrate naturally with permutation-based feature screening (Breiman, 2001; Genuer et al., 2010; Gregorutti et al., 2017; Hapfelmeier et al., 2014; Ramosaj & Pauly, 2019; Scornet et al., 2014).

The gap is specific: existing work shows that bad states, financing constraints, and cross-sectional structure matter, but it does not provide a fixed-feature event design for testing whether extreme volatility shocks reorganize the sector cross-section after common state dependence has been stripped out.

3 Data and Empirical Design

3.1 Data, Units, and Event Construction

The panel consists of the nine standard U.S. sector exchange-traded funds (ETFs), namely consumer discretionary (XLY), consumer staples (XLP), energy (XLE), financials (XLF), healthcare (XLV), industrials (XLI), materials (XLB), technology (XLK), and utilities (XLU), observed weekly from the start of 2001. SPY, the oldest and most liquid S&P 500 ETF, is retained as the broad-market benchmark used to construct relative outcomes and as a market-state variable, but the donor pool is restricted to the sector ETFs themselves.

Data object	Source	Weekly construction	N observations (weeks)
Sector ETF prices	FMP	Week-end close log weekly return	1,322
SPY price	FMP	Week-end close log weekly return	1,322
VIXCLS	FRED	Daily $\Delta \log(\text{VIX})$ weekly maximum	1,322
ICE BofA high-yield option-adjusted spread	FRED	Week-end spread weekly change	1,287
Treasury yields	FRED	Week-end yields changes and slope	1,322
WTI oil spot price	FRED	Week-end WTI level log weekly change	1,264

Table 1: Main data sources, weekly construction, and full-sample observation counts. Sector ETF prices form the treated and donor sector return series. SPY price forms the broad-market benchmark for sector-minus-SPY outcomes and market-state controls. VIXCLS defines daily $\Delta \log(\text{VIX})$ shock weeks and lagged volatility state. High-yield OAS, Treasury yields, and WTI oil enter the lagged macro-financial state vector. The observation count records usable Friday-ending weeks after calendar alignment.

Weekly ETF and SPY returns are measured using the last available close in each Friday-ending week, so holiday weeks use the final trading close available before Friday. I use close-based returns rather than dividend-adjusted total returns to keep the construction uniform across all sector ETFs and SPY, then macro-financial series are aligned to the same Friday-ending weekly grid using the most recent observation available by that week. This preserves the information set available at the time without forcing all data sources onto an identical reporting calendar; as a result, observation counts differ across inputs in Table 1, and the panel uses the aligned weekly value whenever it is available rather than restricting the sample to perfectly synchronized raw calendars.

The shock variable is daily $\Delta \log(\text{VIX})$ computed from VIXCLS, the FRED series corresponding to the CBOE Volatility Index (VIX), which serves as the standard implied-volatility measure for the S&P 500:

$$\Delta \log(\text{VIX})_t = \log(\text{VIX}_t) - \log(\text{VIX}_{t-1}).$$

Because the volatility shock is defined at the daily frequency while the event study is estimated on weekly data, the daily series must first be translated into event weeks. Each daily shock is assigned to its corresponding week-ending Friday, and if multiple large $\Delta \log(\text{VIX})$ moves occur within the same week, only the largest positive move is kept. Candidate event weeks are then ordered chronologically, with a new event admitted only if it occurs at least 13 weeks after the previous accepted event. I refer to this procedure as the 5-percent event design: accepted events are weeks whose largest daily $\Delta \log(\text{VIX})$ realization falls in the upper 5 percent of the full-sample distribution, subject to the 13-week spacing rule. The cutoff isolates extreme implied-volatility shocks while still preserving enough events to study cross-sectional sector reordering, and loosely corresponds to a two-standard-deviation tail event under a Gaussian benchmark.

The accepted event set keeps weeks whose largest daily VIX log change falls in the upper 5 percent of the sample, subject to the 13-week spacing rule. Shock-weighted summaries weight event effects by the event's VIX log-change size. The resulting sample spans both named crises and isolated stress weeks, which helps separate recurring sector behavior from one-off crisis narratives.

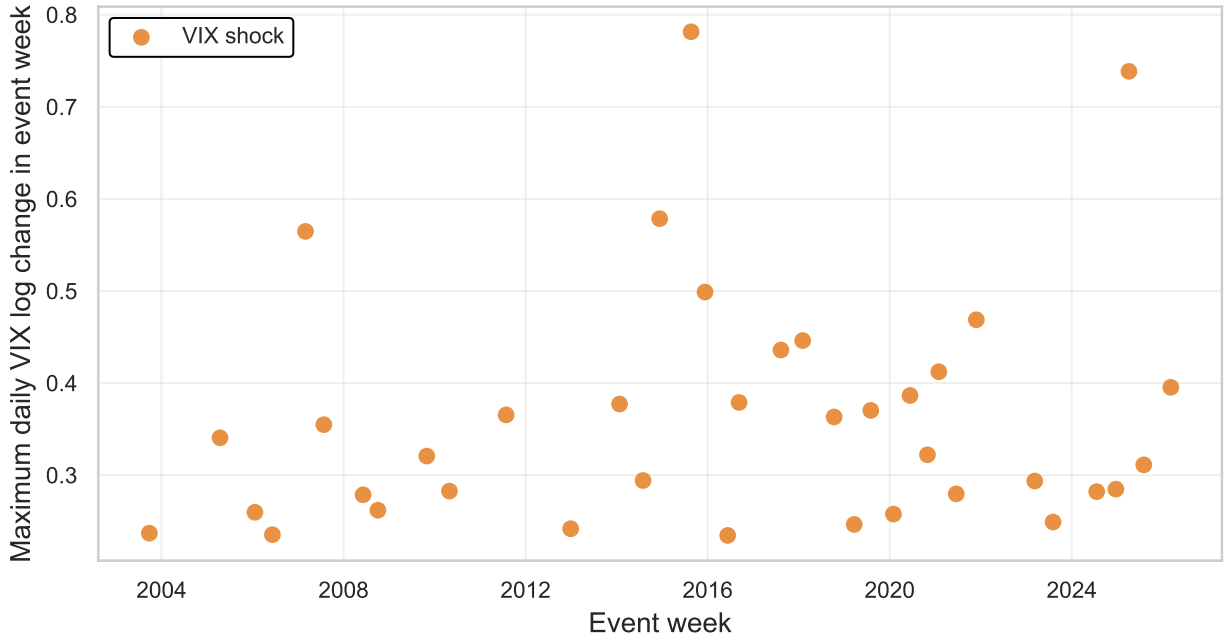

Figure 2: Volatility-shock event timeline. The y-axis reports the maximum daily change in log(VIX), $\Delta \log(\text{VIX})$, within each accepted event week. The accepted spikes cluster in familiar crisis periods but still cover a broad set of stress episodes across the sample.

3.2 Outcomes, Windows, and Identification Logic

The measured outcome is sector-minus-market relative performance:

$$Y_{it} = r_{it}^{1w} - r_{SPY,t}^{1w}.$$

A positive value means the sector outperformed SPY that week, while a negative value means it lagged the broad market. I use raw weekly sector returns as the base series and subtract the contemporaneous weekly SPY return, so the estimand is explicitly relative to the broad equity market. Each event-sector fit treats one sector at a time and uses the other eight sector ETFs as donors for the synthetic counterfactual, with a 52-week pre-treatment window. For an event assigned to week-ending Friday e, the pre-period is weeks $e - 52$ through $e - 1$, and the primary post-treatment outcome is the event week e itself.

The design is quasi-causal rather than fully randomized, since large VIX spikes often coincide with broader macro-financial disturbances. Still, the identifying variation is externally observable, mechanically defined, and largely outside the control of any individual sector. Because every sector ETF is exposed to the same volatility shock at the same time, the estimand is not the total no-shock counterfactual for the equity market as a whole, but the differential residual sector response relative to a donor-implied cross-sectional benchmark under a common stress event. This interpretation is most credible in the present setting because the shock rule is defined independently of sector outcomes, the synthetic-control procedure enforces pre-treatment fit directly in the observed data, and the sector-minus-SPY outcome isolates relative sector reallocation rather than aggregate market repricing.

The object of interest is the cross-sectional distribution of residualized sector responses within each volatility-shock event, not the sector-level mean ATT. The primary reported objects are therefore within-event sector rankings, lower-tail frequencies, and sensitivity-slope regressions linking pre-event exposures to residualized treatment effects, with sector-level ATT relative to SPY reported only as a descriptive summary.

3.3 Frozen Feature Protocol

The predictor set is frozen before the final event study, because the goal is comparability across sectors instead of adaptive tuning to each event window. A small set of primitive macro-financial variables is expanded into theoretically motivated transformations (moving averages, realized volatility, changes, trends, momentum, and drawdowns), with said transformations forming a flexible basis for approximating the unknown conditional mean $m_0(X)$ without committing to a functional form ex ante.

Feature selection then proceeds using permutation-importance-ranked forward stepwise selection, starting from 74 lagged candidate features and 17,316 pre-event sector-week observations; in simpler terms, I rank variables by how

much prediction worsens when I scramble them. Donor-only cluster representatives are then added sequentially, with out-of-bag prediction error recorded after each addition to identify where additional variables stop improving predictive performance in a meaningful way.

The one-standard-deviation out-of-bag (OOB) frontier then identifies the smallest part of the stepwise path whose prediction error is statistically close to the best observed model. Equivalently, the screen is meant to preserve predictive risk- here, risk means squared prediction loss, not financial risk, so the selected feature set should reduce dimensionality without materially worsening prediction of untreated outcomes. This lets the learner explore a rich but theory-structured transformation space and retain what is predictively relevant, reducing ad hoc specification while producing a fixed, interpretable feature set (Appendix B gives a formal proof of permutation-importance screening under a generic consistent learner).

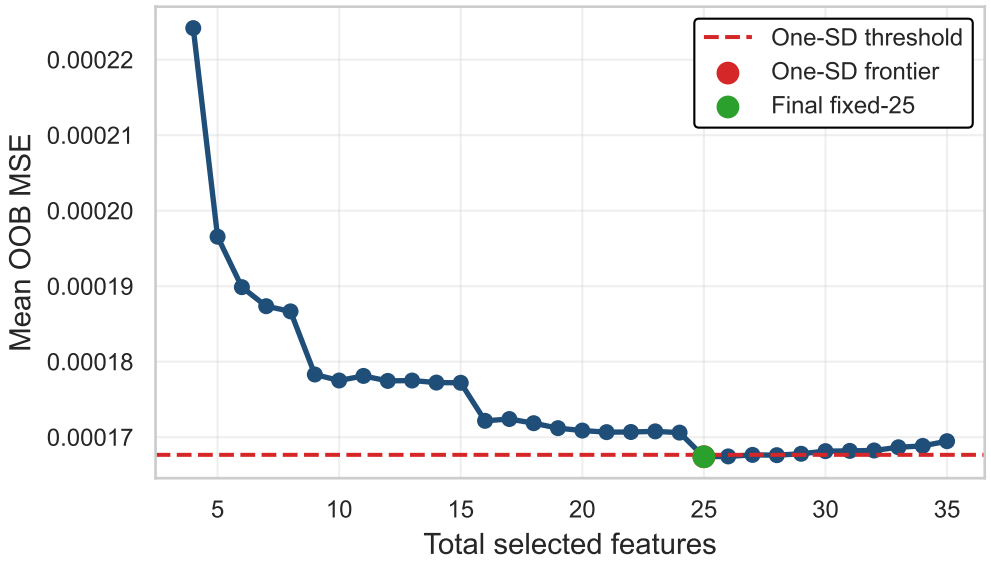

Figure 3: Feature-selection frontier, or pooled pre-event out-of-bag prediction error by feature count. The one-standard-deviation frontier occurs at 25 features, while the minimum-OOB frontier is nearby at 26 features. The reported specification uses the smaller one-standard-deviation set.

The fixed-feature choice matters as much for interpretation as for implementation, because keeping the same predictors across events makes pooled evidence interpretable. If one event is matched using mostly volatility variables, another using return momentum, and another using credit-state variables, then differences across estimated effects may reflect differences in the covariate set rather than differences in sector behavior. Freezing the feature set keeps the event windows, donor pools, and covariates fixed, so differences across rSCM, standard SCM, and AugSynth can actually be read as pure estimator differences.

Feature	Economic Block	Role	Importance Share
Lagged VIX level	Implied volatility	Forced	NA
Lagged 1-week high-yield OAS change	Credit conditions	Forced	NA
Lagged 1-week SPY return	Market return state	Forced	NA
Lagged 1-week sector return	Sector return history	Forced	NA
Lagged 3-month VIX deviation	Implied volatility	Discretionary	3.5%
Lagged 1-month high-yield OAS change	Credit conditions	Discretionary	2.5%
Lagged 3-month average SPY return	Market return state	Discretionary	3.0%
Lagged 1-month SPY return	Market return state	Discretionary	2.7%
Lagged 12-to-1-month SPY momentum	Market return state	Discretionary	2.7%
Lagged 6-month average SPY return	Market return state	Discretionary	2.5%
Lagged 4-week SPY volatility	Market volatility	Discretionary	3.0%
Lagged 12-week SPY drawdown	Market drawdown	Discretionary	2.5%
Lagged 12-month average excess return	Sector return history	Discretionary	7.3%
Lagged 3-month average excess return	Sector return history	Discretionary	6.8%
Lagged 3-month average sector return	Sector return history	Discretionary	5.9%
Lagged 6-month average excess return	Sector return history	Discretionary	5.8%
Lagged 1-month excess return	Sector return history	Discretionary	5.7%
Lagged 1-week excess return	Sector return history	Discretionary	5.3%
Lagged 1-month sector log return	Sector return history	Discretionary	5.3%
Lagged 6-month average sector return	Sector return history	Discretionary	4.9%
Lagged 6-month sector realized volatility	Sector volatility	Discretionary	7.0%
Lagged 4-week sector volatility	Sector volatility	Discretionary	5.8%
Lagged 12-month sector realized volatility	Sector volatility	Discretionary	5.4%
Lagged 12-to-1-month sector momentum	Sector momentum	Discretionary	6.1%
Lagged 12-month sector drawdown	Sector drawdown	Discretionary	6.3%

Table 2: The fixed 25-feature set used in the reported design. Rows are ordered by role and economic block. Forced variables preserve core return, volatility, credit, and VIX information, while discretionary variables are selected from lagged Random Forest permutation-importance screening with correlation and theory-block pruning. Importance shares normalize each discretionary clustered representative's global RF risk increase by total retained discretionary importance; forced variables enter by design and are reported as NA.

3.4 Feature-Based rSCM, SCM, and AugSynth

For each event-sector panel, rSCM separates predictable state dependence from the residual component that SCM has to reconstruct. A Random Forest first learns the relation between lagged macro-financial and sector-state variables and untreated sector-minus-SPY returns using donor observations only. Synthetic control is then applied to the residual path rather than the raw outcome path. The estimand is therefore conditional on volatility, credit, rates, market drawdowns, and sector return history; the goal is to study how sectors reorder after predictable state dependence has been removed, not to estimate the total market effect of a VIX shock.

The first stage is cross-fitted on donor observations from the 52-week pre-treatment window, and treated-sector predictions are generated by donor-trained forests only. This keeps treated outcomes out of the nuisance stage. I then solve a ridge-regularized SCM problem on the residual pre-period path, which discourages brittle one-donor solutions while preserving the donor-based counterfactual logic. If the first stage removes nothing, rSCM collapses to standard SCM. The consistency argument is correspondingly simple: if screening preserves predictive content, the nuisance function is consistently estimated, and the residual path remains representable by stable donor weights, then rSCM approximates the untreated counterfactual.

The event-week ATT is the primary effect. I stack those estimates into an event-sector panel and summarize them using within-event ranks, bottom-five indicators, and sign indicators. The main slope tests relate event-week residualized ATT to four pre-event sensitivity measures: market beta, downside beta, VIX vulnerability, and credit-spread vulnerability. Each sensitivity is estimated from the 52 weeks before the event and standardized within event, so a coefficient reports the effect of moving one standard deviation higher in exposure inside the same shock cross-section.

Measure	Estimation window	Regression / formula	Sign convention
Market beta	52 weeks before event e	$\widehat{\beta}^{M}_{i,e} = \dfrac{\mathrm{Cov}(\ell_{i,t}, \ell_{\mathrm{SPY},t})}{\mathrm{Var}(\ell_{\mathrm{SPY},t})}$	Higher = more market-sensitive
Downside beta	Negative-SPY weeks in the same 52-week window	$\widehat{\beta}^{D}_{i,e} = \dfrac{\mathrm{Cov}(\ell_{i,t}, \ell_{\mathrm{SPY},t})}{\mathrm{Var}(\ell_{\mathrm{SPY},t})}$ $\ell_{\mathrm{SPY},t} < 0$	Higher = more down-market exposure
VIX vulnerability	52 weeks before event e	$\ell_{i,t} - \ell_{\mathrm{SPY},t} = a_{i,e} + b^{VIX}_{i,e} \Delta \log(\mathrm{VIX})_t + u_{i,t}$ $S^{VIX}_{i,e} = -\hat{b}^{VIX}_{i,e}$	Higher = worse when VIX rises
Credit-spread vulnerability	52 weeks before event e	$\ell_{i,t} - \ell_{\mathrm{SPY},t} = a_{i,e} + b^{OAS}_{i,e} \Delta OAS_t + u_{i,t}$ $S^{OAS}_{i,e} = -\hat{b}^{OAS}_{i,e}$	Higher = worse when HY spreads widen

Table 3: Sensitivity-measure definitions. Regression-based sensitivities are sign-flipped so that larger values indicate greater vulnerability to rising VIX or widening high-yield spreads.

rSCM is benchmarked against standard SCM and AugSynth on the same event-sector windows. Since the untreated post path is unobserved, the benchmark criterion is pre-treatment reconstruction, summarized by pre-RMSPE. Identification remains within-event: all sectors face the same volatility shock, so cross-sectional differences in residualized outcomes are interpreted as conditional differential responses rather than aggregate market effects.

4 Results

Unless a table states otherwise, regression standard errors are clustered by event, sector-level descriptive standard errors are computed across events, and pooled ATT standard errors are descriptive standard errors across event-sector fits.

4.1 Event Set, Mean Effects, and Fit Diagnostics

Mean pre-RMSPE in the admitted subset of 37 events from September 26th, 2003 to March 6th, 2026 is 0.0137 under rSCM, compared with 0.0133 under standard SCM and 0.0140 under AugSynth. Here, pre-treatment root mean squared prediction error (pre-RMSPE) is the square root of the average squared gap between the treated sector and its estimated counterfactual over the pre-event window.

The average event-week ATT is -0.004% on an equal-weight basis and -0.021% when weighted by shock size. To limit unstable synthetic-control fits, I impose two guardrails: no donor weight may exceed 0.95 and the effective donor count must be at least 2, and this yields 333 event-sector fits with a guardrail pass rate of 85.29% (pass-only summaries are reported in Appendix A as a fit-quality robustness check).

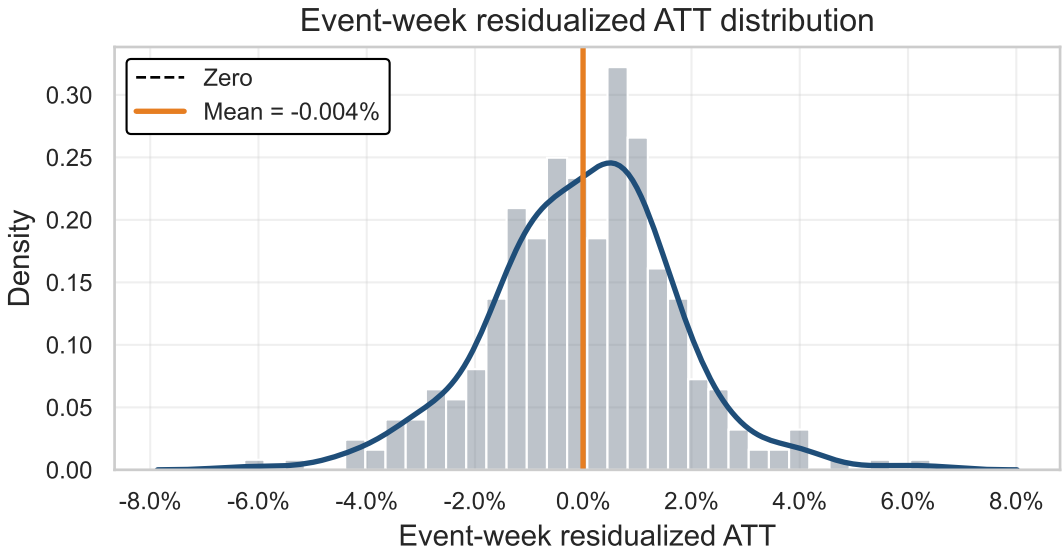

Figure 4: Distribution of event-week residualized ATT across 333 rSCM event-sector fits in the 5-percent event design. Vertical lines mark zero and the pooled mean.

4.2 rSCM Cross-Sectional Reallocation

Applying rSCM yields one residualized treatment effect for each sector in each volatility-shock event, and those event-sector estimates are summarized at the sector level in Table 4:

Rank	Sector	Event-week ATT	N events	Pre-RMSPE	% bottom 5
1	Materials (XLB)	-1.035% (0.251%)	37	0.0120	83.8%
2	Financials (XLF)	-0.466% (0.243%)	37	0.0141	59.5%
3	Energy (XLE)	-0.306% (0.423%)	37	0.0233	54.1%
4	Industrials (XLI)	-0.305% (0.165%)	37	0.0088	64.9%
5	Technology (XLK)	-0.105% (0.304%)	37	0.0136	59.5%
6	Consumer Discretionary (XLY)	0.066% (0.193%)	37	0.0112	51.4%
7	Healthcare (XLV)	0.171% (0.188%)	37	0.0125	62.2%
8	Consumer Staples (XLP)	0.862% (0.256%)	37	0.0107	35.1%
9	Utilities (XLU)	1.084% (0.268%)	37	0.0167	29.7%

Table 4: Sector-level event-week ATT in the 5-percent event design. The outcome is sector-minus-SPY performance, estimates use rSCM, and standard errors are in parentheses.

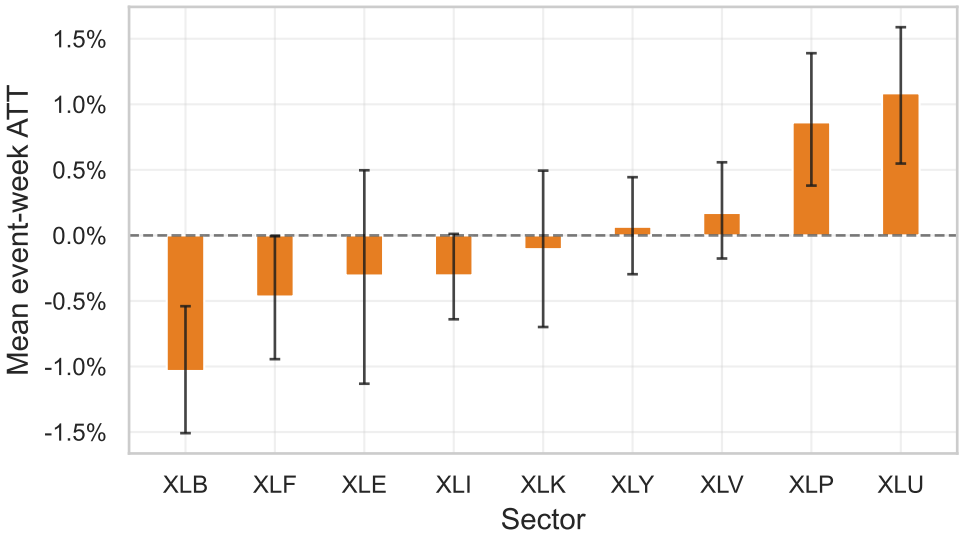

Figure 5: Sector-level event-week ATT in the 5-percent event design, sorted from most negative to most positive.

4.3 Cross-Sectional Sensitivity Regression

I formally test cross-sectional structure using the same rSCM-estimated residual ATT as the dependent variable. For each pre-event sensitivity measure m, I estimate

$$\hat{\tau}_{i,e}^{(1)} = \alpha_e + \beta_m \, \text{Sensitivity}_{i,e,m} + \varepsilon_{i,e},$$

where α_e is an event fixed effect and $\text{Sensitivity}_{i,e,m}$ is standardized inside event e. A negative β_m means that more stress-sensitive sectors have more negative residualized ATT within the same volatility-shock event.

Sensitivity measure	Event-week β	Cluster p-value	Within-event label p-value	Global-label p-value	N events	N event-sector obs.
Market beta	-0.459% (0.096%)	0.000002	0.0002	0.003	37	333
Downside beta	-0.384% (0.132%)	0.004	0.001	0.008	37	333
VIX vulnerability	-0.431% (0.108%)	0.0001	0.0002	0.003	37	333
Credit-spread vulnerability	-0.449% (0.099%)	0.00001	0.0002	0.003	37	333

Table 5: Event fixed-effect sensitivity regressions using event-week rSCM residualized sector ATT. Each sensitivity measure is estimated from 52-week pre-event windows, standardized within event, and entered one at a time; standard errors are clustered by event, and permutation p-values test the directional alternative that the slope is negative under random sector assignment.

Table 5 reports two randomization checks: a within-event permutation test which randomly reassigns sector labels separately inside each event, and a global-label permutation test which permutes the mapping between sector identities and sensitivity scores across the full sample while preserving repeated sector identity across events. Here, the global-label p-value is the stricter test because it preserves the panel structure of the data rather than reshuffling sectors independently within each event.

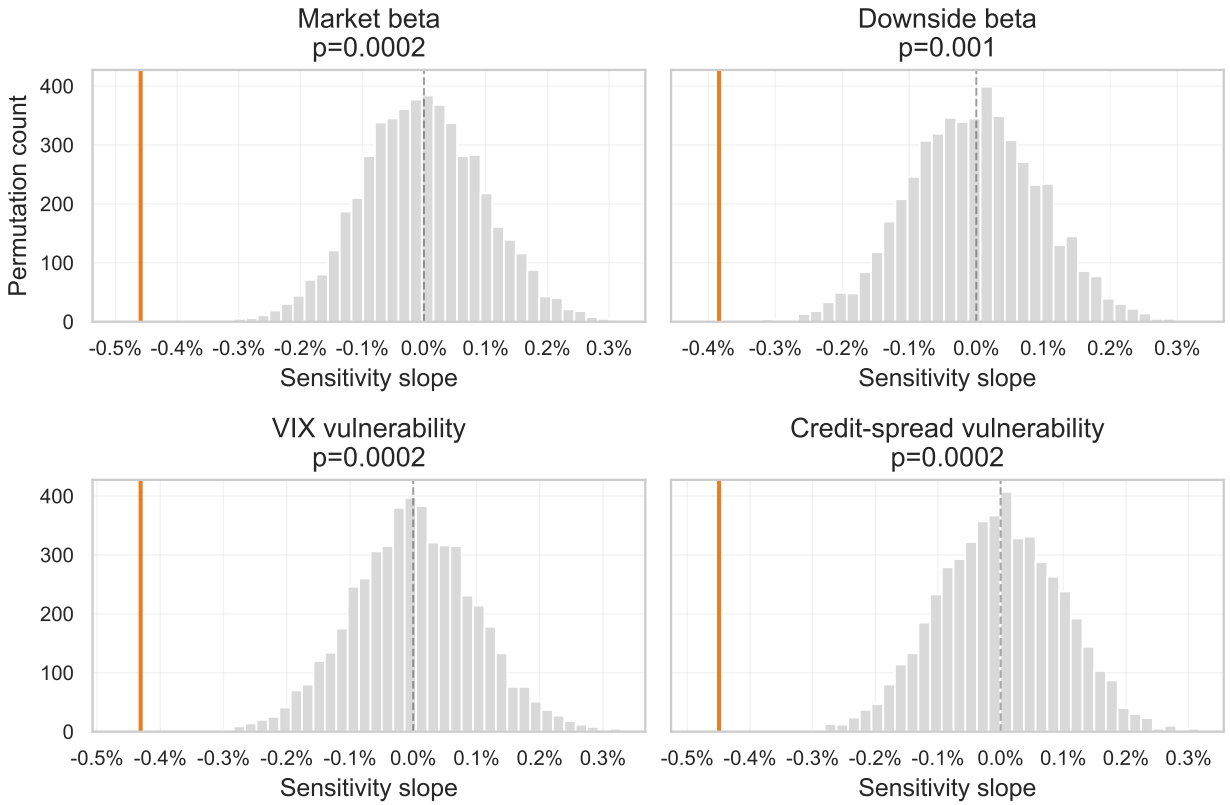

Figure 6: Sector-label permutation inference for the sensitivity slopes. Each histogram reports the within-event sector-label permutation null distribution from 5,000 random permutations; the vertical line is the observed rSCM slope.

The estimated slopes for market beta, downside beta, and credit-spread vulnerability are negative under both clustered and permutation inference, while the VIX vulnerability slope is also negative and has larger clustered and global-label p-values. A one-standard-deviation increase in market beta lowers the event-week residualized ATT by about 0.46%, while the corresponding downside-beta slope is about 0.38%. The median within-event cross-sectional standard deviation of event-week residualized ATT is 1.69%, so the market-beta slope corresponds to roughly 27% of a typical event's cross-sectional dispersion and about 25% of the unconditional weekly standard deviation of relative returns.

4.4 Benchmark Comparison

The benchmark comparison reports pre-treatment reconstruction and guardrail pass rates for rSCM, standard SCM, and AugSynth on the same event-sector windows, with estimator-specific sensitivity slopes reported in Table 7 and sector-level benchmark tables reported in Appendix A.

Method	N matched windows	Mean pre-RMSPE	Mean pass rate	Event-sector wins
rSCM	333	1.36709% (0.03531%)	85.29%	68
standard SCM	333	1.33367% (0.03478%)	77.48%	172
AugSynth	333	1.40038% (0.03396%)	73.57%	93

Table 6: Benchmark comparison on the full 5-percent event-sector sample. Lower pre-RMSPE indicates better pre-treatment reconstruction, and standard errors are in parentheses.

Mean pre-RMSPE is 0.01367 under rSCM, 0.01334 under standard SCM, and 0.01400 under AugSynth, with standard SCM having the lowest average pre-RMSPE and rSCM having the highest guardrail pass rate. On the same benchmark sample, the mean absolute event-week ATT is 1.31% under rSCM, 1.25% under standard SCM, and 1.37% under AugSynth.

Method	Market beta β	Downside beta β	VIX vulnerability β	Credit vulnerability β
rSCM	-0.459% (0.096%)	-0.384% (0.132%)	-0.431% (0.108%)	-0.449% (0.099%)
standard SCM	-0.401% (0.085%)	-0.349% (0.124%)	-0.383% (0.100%)	-0.387% (0.086%)
AugSynth	-0.463% (0.086%)	-0.418% (0.139%)	-0.414% (0.108%)	-0.499% (0.118%)

Table 7: Sensitivity slopes by estimator on the full 5-percent event-sector sample. The dependent variable is the event-week ATT produced by each estimator, and standard errors are clustered by event.

For the final diagnostic in this block, I report a simple linear check on the first-stage residualization step in Table 8, which compares the raw sector-minus-SPY outcome with the rSCM first-stage residual on the pre-treatment event-sector windows, reporting correlations with lagged VIX, lagged high-yield OAS changes, and lagged SPY returns, together with a linear state R^2 summary.

Object	N	Corr. lag VIX	Corr. lag OAS change	Corr. lag SPY return	Linear state R^2
Raw sector-minus-SPY outcome	17316	-0.011	-0.014	0.020	0.001
rSCM first-stage residual	17316	-0.009	-0.014	0.016	0.000

Table 8: Residualization diagnostic on pre-treatment event-sector windows. The table compares raw sector-minus-SPY outcomes with rSCM first-stage residuals using simple linear state summaries; lower absolute correlations and lower linear state R^2 indicate less remaining linear association with these measured lagged states.

In the residualization diagnostic, linear state R^2 is 0.001 for the raw sector-minus-SPY outcome and 0.000 for the rSCM first-stage residual.

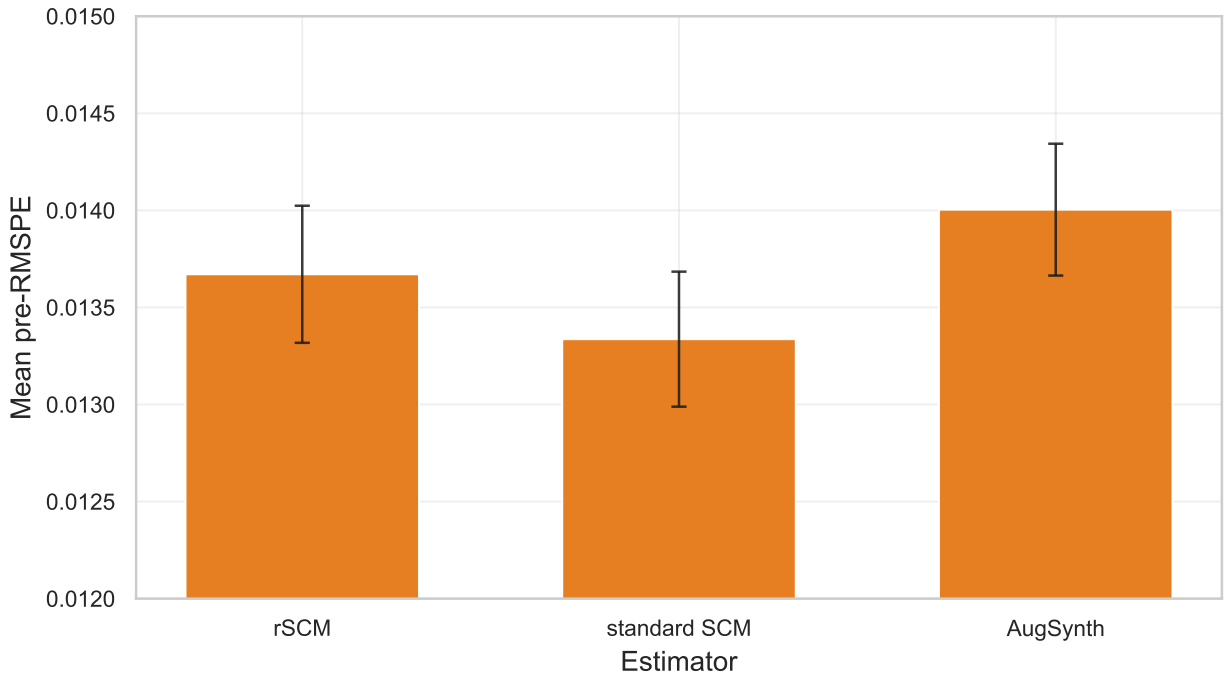

Figure 7: Benchmark fit comparison: mean pre-treatment RMSPE by estimator on the full 5-percent event-sector sample. Error bars are standard errors across matched event-sector windows.

4.5 Shock-Channel Checks

I split events by whether they are credit-intense, rate-intense, or among the largest VIX jumps. Credit-intense events are the top third of accepted events by weekly high-yield OAS widening, rate-intense events are the top third by absolute weekly moves in the 10-year Treasury yield, and largest-VIX-jump events are the top third by daily $\Delta \log(\text{VIX})$, yielding 13 credit-intense events, 14 rate-intense events, and 13 largest-VIX-jump events. These are separate binary splits rather than mutually exclusive shock-family assignments, so an event can belong to more than one group; in the accepted event set, 8 events are both credit-intense and rate-intense, 7 are both credit-intense and largest VIX jumps, 6 are both rate-intense and largest VIX jumps, and 4 events are classified in all three groups.

Shock split	Sensitivity measure	Intense β	Other events β	Interaction p-value
Rate-intense	Market beta	-0.618% (0.144%)	-0.362% (0.125%)	0.172
Rate-intense	Downside beta	-0.642% (0.275%)	-0.227% (0.125%)	0.161
Rate-intense	VIX vulnerability	-0.625% (0.196%)	-0.313% (0.123%)	0.169
Rate-intense	Credit-spread vulnerability	-0.692% (0.196%)	-0.301% (0.096%)	0.068
Credit-intense	Market beta	-0.634% (0.139%)	-0.364% (0.125%)	0.141
Credit-intense	Downside beta	-0.767% (0.255%)	-0.176% (0.134%)	0.036
Credit-intense	VIX vulnerability	-0.707% (0.176%)	-0.281% (0.128%)	0.046
Credit-intense	Credit-spread vulnerability	-0.711% (0.193%)	-0.307% (0.102%)	0.059
Largest VIX jumps	Market beta	-0.510% (0.151%)	-0.432% (0.126%)	0.686
Largest VIX jumps	Downside beta	-0.469% (0.272%)	-0.338% (0.146%)	0.664
Largest VIX jumps	VIX vulnerability	-0.516% (0.206%)	-0.385% (0.126%)	0.580
Largest VIX jumps	Credit-spread vulnerability	-0.579% (0.215%)	-0.379% (0.100%)	0.390

Table 9: Shock-channel sensitivity checks. The table reports event fixed-effect slopes separately for credit-intense, rate-intense, and largest-VIX-jump event splits; standard errors are clustered by event.

Across these channel splits, the signs remain mostly negative, but the interaction tests have large standard errors and the subgroup memberships overlap. The credit-intense and rate-intense splits show some more negative subgroup slopes, while the largest-VIX-jump split produces smaller differences across subgroups.

4.6 Robustness Checks

The primary sensitivity estimates use the event-week ATT from the fixed 25-feature rSCM specification together with the baseline daily-to-weekly VIX-shock mapping. I revisit the same negative sensitivity-slope pattern under alternative feature counts, event-cutoff choices, event-mapping conventions, short-lag conditioning choices, post-treatment horizons, and leave-one-crisis exclusions.

Event mapping	Features	Events	Market beta β	Downside beta β	VIX vulnerability β	Credit vulnerability β
Daily mapped	5	37	-0.886% (0.393%)	-0.967% (0.362%)	-0.703% (0.405%)	-1.118% (0.413%)
Daily mapped	10	37	-0.911% (0.421%)	-1.020% (0.426%)	-0.696% (0.420%)	-1.191% (0.461%)
Daily mapped	20	37	-0.960% (0.437%)	-0.947% (0.453%)	-0.723% (0.457%)	-1.213% (0.474%)
Daily mapped	25	37	-0.946% (0.452%)	-0.980% (0.456%)	-0.706% (0.459%)	-1.242% (0.488%)
Daily mapped	30	37	-0.974% (0.446%)	-1.038% (0.460%)	-0.753% (0.453%)	-1.248% (0.482%)
Weekly peak-to-peak	5	37	-0.818% (0.394%)	-0.893% (0.375%)	-0.571% (0.390%)	-1.006% (0.404%)
Weekly peak-to-peak	10	37	-0.887% (0.430%)	-0.962% (0.452%)	-0.678% (0.441%)	-1.174% (0.470%)
Weekly peak-to-peak	20	37	-0.893% (0.459%)	-0.937% (0.470%)	-0.661% (0.473%)	-1.154% (0.486%)
Weekly peak-to-peak	25	37	-0.978% (0.458%)	-0.979% (0.450%)	-0.734% (0.473%)	-1.261% (0.489%)
Weekly peak-to-peak	30	37	-0.896% (0.443%)	-0.991% (0.454%)	-0.673% (0.461%)	-1.175% (0.482%)

Table 10: Cumulative-horizon sensitivity-slope comparison across sparse feature-count and event-mapping choices. The dependent variable is cumulative eight-week rSCM residualized sector ATT; standard errors are clustered by event.

Across these checks, the qualitative pattern changes little. The negative sensitivity slopes remain visible across sparse and richer feature sets, under the weekly peak-to-peak VIX mapping, and under broader top-decile and narrower top-percentile shock definitions.

Shock cutoff	Events	Market beta β	Downside beta β	VIX vulnerability β	Credit vulnerability β
Top 10%	60	-0.865% (0.293%)	-0.589% (0.328%)	-0.659% (0.298%)	-0.983% (0.306%)
Top 5%	37	-0.459% (0.096%)	-0.384% (0.132%)	-0.431% (0.108%)	-0.449% (0.099%)
Top 1%	12	-0.985% (0.577%)	-1.434% (0.386%)	-0.754% (0.566%)	-1.152% (0.568%)

Table 11: Event-cutoff robustness for headline event-week sensitivity slopes. The top 5-percent row is the primary specification; top 10-percent and top 1-percent rows rerun the same fixed 25-feature daily-mapped core pipeline. Standard errors are clustered by event and shown in parentheses.

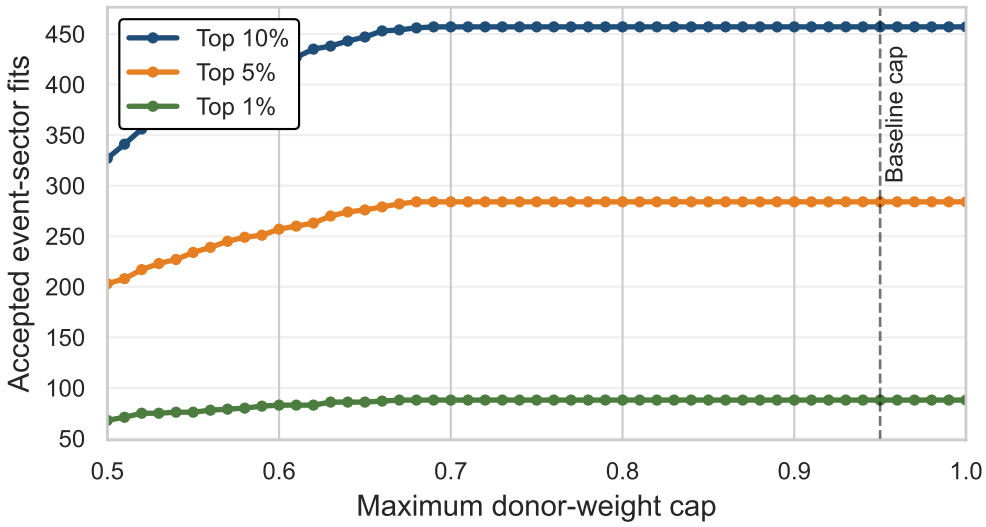

Figure 8: Donor-weight cap sensitivity. Lines report the number of event-sector fits that pass the full guardrail screen as the maximum donor-weight cap is tightened in 0.01 increments from 0.50 to 1.00, holding the effective-donor floor at 2 and the pre-RMSPE cap at 0.08. The top-percentile line is included for consistency but is based on only 12 accepted events. The baseline cap of 0.95 is far from the binding region in all three event definitions.

Sensitivity	β	**Cluster p**	**Within-event p**	**Global-label p**	**Features kept**
Market beta	-0.914% (0.447%)	0.041	0.004	0.015	21
Downside beta	-1.024% (0.451%)	0.023	0.001	0.006	21
VIX vulnerability	-0.727% (0.446%)	0.103	0.015	0.054	21
Credit-spread vulnerability	-1.212% (0.477%)	0.011	0.0002	0.003	21

Table 12: Cumulative-horizon no-short-lag comparison. The rSCM specification drops the most treatment-proximate lagged sector-state features and re-estimates the eight-week sensitivity slopes; p-values are directional negative-slope tests.

Horizon	**Market beta** β	**Downside beta** β	**VIX vulnerability** β	**Credit vulnerability** β
1w	-0.459% (0.096%)	-0.384% (0.132%)	-0.431% (0.108%)	-0.449% (0.099%)
2w	-0.541% (0.138%)	-0.571% (0.217%)	-0.453% (0.139%)	-0.571% (0.135%)
4w	-0.656% (0.229%)	-0.424% (0.317%)	-0.468% (0.249%)	-0.695% (0.231%)
8w	-0.946% (0.452%)	-0.980% (0.456%)	-0.706% (0.459%)	-1.242% (0.488%)

Table 13: Horizon comparison for the sensitivity slopes. The event-week estimate is the primary specification; later cumulative horizons are reported for comparison.

Exclusion	Dropped accepted events	Events	Market beta β	Downside beta β	VIX vulnerability β	Credit vulnerability β
Exclude 2022 inflation/rates	0	37	-0.946% (0.452%)	-0.980% (0.456%)	-0.706% (0.459%)	-1.242% (0.488%)
Exclude 2023 regional-bank stress	1	36	-0.957% (0.464%)	-1.007% (0.468%)	-0.683% (0.471%)	-1.263% (0.501%)
Exclude COVID crash	0	37	-0.946% (0.452%)	-0.980% (0.456%)	-0.706% (0.459%)	-1.242% (0.488%)
Exclude GFC core	1	36	-0.644% (0.328%)	-0.731% (0.382%)	-0.453% (0.382%)	-0.971% (0.405%)

Table 14: Cumulative-horizon leave-one-crisis-out comparison. Each row drops events in the named crisis window and re-estimates the eight-week rSCM sensitivity slope.

The no-short-lag check keeps the slopes negative after removing the sector-return lags most likely to be contaminated by early shock responses. Horizon and leave-one-crisis-out checks likewise show that the result is not confined to the event week or driven by one named stress episode.

5 Discussion

5.1 Analysis

rSCM produces pre-RMSPE close to standard SCM and AugSynth, indicating that residualization does not substantially weaken pre-treatment fit. Its higher guardrail pass rate also suggests that comparable fits are less often produced by concentrated donor solutions. The procedure therefore mainly changes the interpretation of the estimated effect by removing predictable variation associated with observable market states.

The sector-level estimates show the reallocation margin directly. Materials is the clearest weak-side sector, with an average event-week ATT of -1.03%, and it appears in the bottom five in 83.8% of accepted events. Financials and energy also lean negative on average, with event-week ATTs of -0.47% and -0.31%, respectively. Utilities and consumer staples sit at the resilient end of the ordering, with average event-week ATTs of 1.08% and 0.86%, respectively. The ordering is economically coherent, as stress episodes tend to move relative performance away from sectors tied more closely to cyclical demand, credit conditions, and commodity-linked risk, and toward sectors whose earnings or balance sheets are easier to hold through volatility shocks.

Aggregating across sectors, the equal-weight event-week ATT is -0.004%, and the shock-weighted event-week ATT is -0.021%. Both averages are small, and the full ATT distribution is approximately symmetric around zero. Once SPY is subtracted, the remaining signal looks less like an additional market-wide residual loss or gain and more like rotation inside the sector cross-section. Nonzero pooled means can still arise when observed sector-relative returns deviate systematically from their synthetic counterfactuals, but the central pattern is cross-sectional rather than pooled.

A one-standard-deviation increase in market beta is associated with an event-week ATT that is -0.46% lower, while the downside-beta and credit-spread vulnerability slopes are -0.38% and -0.45%, respectively. These are meaningful movements relative to the cross-section: the market-beta slope is about 27% of a typical event's within-sector dispersion, and the credit-spread vulnerability slope is about 27%. The corresponding within-event sector-label permutation p-values are 0.0002, 0.001, and 0.0002, so the pattern is difficult to explain as random reshuffling of sector identities.

This is the main empirical point, because the weak-side sectors are not just weak labels in one sample; instead, weaker residualized performance is concentrated among sectors that already looked more exposed to downside market conditions and credit stress before the shock. VIX vulnerability points in the same direction, with a slope of -0.43%, but it is less stable under the stricter inference checks. Most relevant from all of this is the clear alignment between post-shock residualized performance and pre-event market beta, downside beta, and credit-spread vulnerability.

The mechanism evidence points most clearly toward downside-risk and financing-condition channels. Market beta and downside beta capture exposure to broad repricing and bad-market states, while credit-spread vulnerability captures whether a sector tends to weaken when high-yield spreads widen. The fact that all three measures predict weaker residualized performance suggests that volatility shocks do not only reshuffle sectors mechanically; they push relative

performance away from sectors that are harder to hold when risk-bearing capacity tightens. The shock-channel splits are consistent with that interpretation, but they should be read cautiously. The credit-intense and rate-intense splits produce more negative slopes for some measures while the largest-VIX-jump split is less informative, suggesting that the reallocation pattern is more closely tied to financing and rate-stress conditions than to VIX jump size alone. However, the small and overlapping event groups make this only suggestive.

rSCM is not rescuing a raw outcome that was obviously dominated by simple linear state exposure: sector-minus-SPY is already close to orthogonal to the simple lagged state variables reported in Table 8. That is plausible after subtracting SPY and moving to weekly relative returns because the outcome has already removed the broad market component. The value of rSCM is narrower: it tests whether residual cross-sectional structure remains after a more flexible nonlinear conditioning step.

The robustness checks show that the result is not dependent on one design choice. The pattern remains visible across alternative feature counts, VIX mappings, event cutoffs, horizons, no-short-lag specifications, and leave-one-crisis-out exclusions. The estimates move in magnitude, but the qualitative interpretation changes very little.

5.2 Limitations

The main identifying assumption is that the residual sector path remains representable by a stable donor combination after conditioning on state variables, but this is not directly testable. It is most plausible when sector comovement is strong enough to provide useful donor information, but not so strong that stress periods collapse the donor space into one nearly common factor. I try to make this assumption more credible with fit diagnostics and guardrail checks, but this is a core identifying issue in synthetic control designs generally: the untreated counterfactual must be well represented by the donor pool in the relevant outcome space.

A second limitation comes from dynamic conditioning. Some nuisance features are lagged sector-state variables, so later post-event predictions can condition on lagged returns, volatility, or drawdowns that may already reflect the early shock response. This is the main motivation as to why the event-week estimate is the primary estimand, and why cumulative horizons are only treated as comparisons. The no-short-lag robustness check helps, because the slopes remain negative after the most treatment-proximate sector-return lags are removed, but it does not eliminate the issue completely.

Power is also limited at the sector level. The primary design has 37 accepted events and nine sectors, and sector ETFs are highly correlated precisely in the stress states studied here, so sector means can be noisy. This makes it hard to treat any one sector's average ATT as a precise standalone estimate, especially when a few crisis episodes can move sector averages meaningfully. For that reason, I put more weight on repeated lower-tail membership, within-event rankings, and pooled sensitivity slopes than on isolated sector-level means.

Finally, sector ETFs are imperfect proxies for clean economic sectors. Their holdings are determined by index classification rules, and those classifications do not always line up with underlying business models or risk exposures.

For example, Tesla and Amazon currently comprise nearly half of the consumer discretionary sector ETF even though they are members of the Magnificent 7, an informal group of megacap American corporations that most investors would classify as part of the tech sector. Similar classification ambiguity can arise when conglomerates, platform firms, or firms with mixed revenue sources are assigned to a single sector bucket. The main way I address this is by studying events over a long sample period: the composition of XLY in 2006 is very different from its composition in 2026, so averaging across many stress episodes should reduce the influence of any one sector composition at a single point in time and leave more persistent sector-level risk patterns.

Several extensions follow from these limitations. On the econometric side, the natural next step is a finite-sample Monte Carlo design built around nonlinear untreated outcome processes. That would make the tradeoff more concrete: residualization should help most when nonlinear state dependence biases standard SCM, and much less when the raw outcome is already well represented by donor weights. A more ambitious methodological extension would be an orthogonalized SCM-style estimator in which first-stage nuisance error is made locally irrelevant to the target estimand in the Neyman-orthogonality sense. Empirically, one could also define stress using realized volatility, variance-risk measures, volatility-term-structure measures, or broader cross-asset stress indicators; stability of the ordering across those definitions would strengthen the interpretation.

A further extension could classify stress episodes into richer shock families, potentially using principal components or related dimension-reduction tools, and then test whether sector ordering differs across geopolitical shocks, financial-system shocks, inflation-and-rates shocks, growth scares, or policy-driven events. I avoid doing that here because it would broaden the scope too much and take focus away from the central question of whether volatility shocks systematically reorganize relative sector performance.

6 Conclusion

I began with a simple question: after a large daily volatility shock, does the market only fall together, or does sector leadership reorganize in a systematic way? The evidence points to reorganization. Once SPY is subtracted, the pooled event-week ATT is small and the ATT distribution is roughly centered around zero, so I do not find a large additional market-wide residual loss. The important movement is instead inside the sector cross-section, where some sectors lose relative ground, others hold up, and that ordering is tied to economically meaningful pre-event exposures.

The strongest evidence comes from the sensitivity regressions. Sectors with higher market beta, downside beta, and credit-spread vulnerability tend to have more negative residualized event-week responses after volatility shocks. Sector-label permutation tests show that this pattern is difficult to explain as random reshuffling, and the robustness checks show that it does not depend on one feature count, event cutoff, VIX mapping rule, horizon, or named crisis episode. The result is not a universal sector ranking after every shock, but a repeated tendency for volatility shocks to rotate relative performance away from sectors with greater downside and financing-condition exposure.

My methodological contribution is residual synthetic control method (rSCM), which separates two problems that are otherwise conflated in this setting: predictable state-dependent sector movement, and the residual cross-sectional response to the shock itself. The counterfactual is built in residual space, so the estimated treatment effect is explicitly conditional on observable macro-financial state rather than absorbing it into the donor weights. Despite the difference in approach, pre-RMSPE under rSCM is close to both standard SCM and AugSynth, so the procedure does not sacrifice pre-treatment fit. Estimated ATTs are also fairly close across the three methods, but the main difference is that rSCM has a higher guardrail pass rate, suggesting that its comparable fits are less often produced by highly concentrated or low-breadth donor solutions.

7 Appendices

7.1 Appendix A. Empirical Summary Tables

Episode	Event week	Spearman vs main ranking	Weakest sector	Weakest ATT	Strongest sector	Strongest ATT
Early GFC funding stress	2007-11-09	0.45	Technology (XLK)	-5.34%	Consumer Staples (XLP)	2.92%
Lehman week	2008-09-19	-0.70	Utilities (XLU)	-3.60%	Financials (XLF)	8.82%
Euro-area sovereign stress	2011-11-04	0.43	Financials (XLF)	-4.10%	Utilities (XLU)	1.83%
Volmageddon	2018-02-09	0.03	Energy (XLE)	-3.76%	Utilities (XLU)	2.99%
COVID volatility aftershock	2020-06-12	0.53	Energy (XLE)	-4.09%	Technology (XLK)	3.38%
Inflation / rates stress	2022-04-22	0.37	Energy (XLE)	-2.49%	Consumer Staples (XLP)	3.18%
Regional-bank stress	2023-03-10	0.37	Financials (XLF)	-3.65%	Technology (XLK)	2.59%
April 2025 'Liberation Day' shock	2025-04-04	0.72	Energy (XLE)	-6.27%	Consumer Staples (XLP)	4.50%

Table 15: Recognizable stress episodes and formal alignment with the main sector ordering. Spearman correlations compare each full nine-sector event ranking with the average ranking from the 5-percent event design. This table is descriptive and not used for inference.

Total features	Mean OOB MSE	SD across seeds	One-SD frontier point
18	0.000172	0.000000	No
20	0.000171	0.000000	No
25	0.000167	0.000000	Yes
26	0.000167	0.000001	No
27	0.000168	0.000000	No
30	0.000168	0.000001	No

Table 16: Selected points along the OOB screening frontier.

The event-week estimate is the main specification. For comparison, the table below lists sectors from most negative to most positive at later cumulative horizons as well.

Horizon	Top 10% ordering most negative to most positive	Top 5% ordering most negative to most positive
1w	XLB, XLK, XLF XLY, XLV, XLI XLE, XLP, XLU	XLB, XLF, XLE XLI, XLK, XLY XLV, XLP, XLU
2w	XLB, XLF, XLY XLV, XLK, XLE XLI, XLU, XLP	XLF, XLE, XLB XLV, XLY, XLI XLK, XLU, XLP
4w	XLV, XLY, XLB XLF, XLI, XLE XLK, XLU, XLP	XLF, XLB, XLE XLY, XLI, XLV XLK, XLU, XLP
8w	XLY, XLF, XLB XLE, XLV, XLI XLU, XLK, XLP	XLE, XLF, XLB XLY, XLI, XLV XLK, XLU, XLP

Table 17: Sector ordering by horizon, listed from most negative to most positive relative performance.

This table is a timing comparison only, and the main text treats the event-week response as the primary estimand.

7.1.1 Two Comparative SCM Cases

To put the residualized design in context, I also rerun two well-known SCM applications with three estimators: standard SCM, OLS-residualized SCM, and AugSynth. These cases are useful because they are much closer to the canonical SCM setting, where the untreated path can often be matched directly in outcome levels. I wnated to check how much independent post-treatment gap remains for synthetic control to explain once the covariates emphasized in the original studies are partialled out first.

I find that if first-stage covariates are treatment-proximate or already carry most of the post-treatment object, residualization can mechanically wipe out the effect one is trying to study. The thesis application is different because the first stage is built from lagged macro-financial and sector-state variables, not from contemporaneous sector outcomes.

Case	Method	Pre-RMSPE	Avg post gap	Avg post % gap	Final gap
Basque Country	AugSynth	0.0521	-1.30	-13.6%	-1.75
Basque Country	OLS-rSCM	0.0328	-0.04	-0.6%	0.04
Basque Country	standard SCM	0.0642	-0.98	-11.0%	-1.11
Sweden Carbon Tax	AugSynth	0.6624	6.10	37.5%	15.40
Sweden Carbon Tax	OLS-rSCM	0.6031	1.24	11.4%	-0.28
Sweden Carbon Tax	standard SCM	0.6576	6.42	40.4%	16.12

Table 18: External SCM case-study comparisons using standard SCM, OLS-rSCM, and AugSynth.

The Basque Country case of Abadie and Gardeazabal (2003) is the cleanest illustration. Standard SCM reproduces the familiar negative post-treatment gap, with an average post gap of -0.98 and an average percent gap of -11.0%, while OLS-rSCM cuts pre-RMSPE from 0.0642 to 0.0328 and drives the average residual gap down to only -0.04. Once the Abadie-style covariates are partialled out, very little residual counterfactual problem is left. AugSynth estimates a larger negative gap here but fails the donor-weight guardrails.

The Swedish carbon-tax study of Moore et al. (2025) yields the same pattern. Standard SCM and AugSynth both recover a sizable positive post-reform gap, with average post effects of 6.42 and 6.10 clean patents per year, but OLS-rSCM improves pre-fit only modestly, from 0.6576 to 0.6031, while shrinking the average post gap to 1.24 and the final gap to -0.28. The residual effect is again close to zero once that study's predictor set is used in the first stage.

In these canonical applications, the variables that the original papers show to be relevant already carry most of the economically meaningful structure, so residualization leaves very little independent donor-representation problem behind. The thesis setting is different because its first stage is built from lagged macro-financial and sector-state variables such as volatility, credit, rates, and return history, not from contemporaneous sector outcomes. When residualization still leaves a meaningful cross-sectional sector-ordering problem after conditioning on those state variables, that is evidence that the method is doing nontrivial work rather than simply re-labeling the original SCM fit.

7.2 Appendix B. Consistency of Permutation-Importance Screening for a Generic Consistent Learner

7.2.1 Setup

Let (X, Y) be a generic untreated observation with

$$Y = m_0(X) + \varepsilon, \qquad \mathbb{E}[\varepsilon \mid X] = 0,$$

and assume $\mathbb{E}[Y^4] < \infty$. Write $X = (X_1, \dots, X_p)$. For any feature j, define the single-coordinate-permuted covariate vector

$$X^{\pi_j} = (X_1, \dots, X_{j-1}, \widetilde{X}_j, X_{j+1}, \dots, X_p),$$

where $\widetilde{X}_j$ is an independent draw from the marginal law of X_j, independent of (X, Y).

Let $\mathcal{A}_n$ be any learning algorithm trained on an independent training sample D_n^{tr}, and let

$$\widehat{m}_n = \mathcal{A}_n(D_n^{tr}).$$

The empirical permutation-importance estimator is computed on an independent evaluation sample $D_m^{ev} = \{(X_i^{ev}, Y_i^{ev}, \widetilde{X}_{ij}^{ev})\}_{i=1}^m$:

$$\widehat{\Delta}_{j,n,m} = \frac{1}{m} \sum_{i=1}^{m} \left[\left(Y_i^{ev} - \widehat{m}_n((X_i^{ev})^{\pi_j})\right)^2 - \left(Y_i^{ev} - \widehat{m}_n(X_i^{ev})\right)^2 \right].$$

Breiman introduced permutation importance for Random Forests (Breiman 2001), and Gregorutti, Michel, and Saint-Pierre discuss how correlated predictors complicate feature-importance interpretation (Gregorutti et al. 2017). The narrower result I prove below is that with a consistent learner and an independent evaluation sample, empirical permutation importance converges to the population increase in predictive risk from permuting a feature. Random Forests are then one practical learner choice because a consistency route exists through Scornet et al. (2014).

7.2.2 Population Permutation Importance

For any square-integrable predictor g, define

$$R(g) = \mathbb{E}[(Y - g(X))^2], \qquad R_j^{\pi}(g) = \mathbb{E}[(Y - g(X^{\pi_j}))^2],$$

and

$$\Delta_j(g) = R_j^{\pi}(g) - R(g).$$

The key identity is that for the true conditional mean m_0, permutation importance is exactly the lost signal from breaking the j-th coordinate.

Proposition B.1. Under the setup above,

$$\Delta_j(m_0) = \mathbb{E}\big[(m_0(X) - m_0(X^{\pi_j}))^2\big] \geq 0.$$

Proof. Expand the squared losses:

$$R_j^{\pi}(m_0) - R(m_0) = \mathbb{E}\left[(Y - m_0(X^{\pi_j}))^2 - (Y - m_0(X))^2\right].$$

Substitute $Y = m_0(X) + \varepsilon$:

$$(Y - m_0(X^{\pi_j}))^2 - (Y - m_0(X))^2 = (m_0(X) - m_0(X^{\pi_j}) + \varepsilon)^2 - \varepsilon^2.$$

Therefore

$$\Delta_j(m_0) = \mathbb{E}\big[(m_0(X) - m_0(X^{\pi_j}))^2\big] + 2\mathbb{E}[\varepsilon\,(m_0(X) - m_0(X^{\pi_j}))].$$

The cross term vanishes because $\mathbb{E}[\varepsilon \mid X] = 0$ and $\widetilde{X}_j$ is independent of (X, ε), so conditional on X, the quantity $m_0(X) - m_0(X^{\pi_j})$ is measurable with finite second moment and has zero covariance with ε. Hence

$$\Delta_j(m_0) = \mathbb{E}\big[(m_0(X) - m_0(X^{\pi_j}))^2\big] \geq 0.$$

□

7.2.3 Sample Estimator and Consistency

The next result reduces permutation-importance consistency to learner consistency. The learner must approximate the conditional mean under both the original covariate law and the single-feature-permuted law, and the evaluation sample must be independent of training data.

Assumption B.1. The learner is risk-consistent under both the original and single-coordinate-permuted covariate laws:

$$\|\widehat{m}_n - m_0\|_{L_2(P_X)} \xrightarrow{p} 0, \qquad \|\widehat{m}_n - m_0\|_{L_2(P_{X^{\pi_j}})} \xrightarrow{p} 0.$$

Assumption B.2. The evaluation sample is independent of the training sample and satisfies a law of large numbers for the square-loss differences.

Theorem B.1. Under Assumptions B.1 and B.2,

$$\widehat{\Delta}_{j,n,m} \xrightarrow{p} \Delta_j(m_0)$$

as $n, m \to \infty$.

Proof. Write

$$\widehat{\Delta}_{j,n,m} - \Delta_j(m_0) = \Big(\widehat{\Delta}_{j,n,m} - \Delta_j(\widehat{m}_n)\Big) + \Big(\Delta_j(\widehat{m}_n) - \Delta_j(m_0)\Big).$$

The first bracket is an empirical-process term. Conditional on the training sample, the evaluation observations are iid and the square-loss difference has finite expectation. Therefore, by the conditional law of large numbers,

$$\widehat{\Delta}_{j,n,m} - \Delta_j(\widehat{m}_n) \xrightarrow{p} 0.$$

It remains to show $\Delta_j(\widehat{m}_n) - \Delta_j(m_0) \to_p 0$. Since

$$\Delta_j(g) - \Delta_j(h) = (R_j^\pi(g) - R_j^\pi(h)) - (R(g) - R(h)),$$

it is enough to show both risk differences vanish when $g = \widehat{m}_n$ and $h = m_0$. Consider the ordinary risk:

$$R(\widehat{m}_n) - R(m_0) = \mathbb{E}\left[(m_0(X) - \widehat{m}_n(X))(m_0(X) + \widehat{m}_n(X) - 2Y)\right].$$

By Cauchy-Schwarz,

$$|R(\widehat{m}_n) - R(m_0)| \le \|\widehat{m}_n - m_0\|_{L_2(P_X)} \cdot \|m_0 + \widehat{m}_n - 2Y\|_{L_2(P_X)}.$$

The first factor converges to zero by Assumption B.1. The second is $O_p(1)$ because Y has finite fourth moment and $\widehat{m}_n$ is L_2-consistent. Hence

$$R(\widehat{m}_n) - R(m_0) \xrightarrow{p} 0.$$

The same argument under the permuted law gives

$$R_j^{\pi}(\widehat{m}_n) - R_j^{\pi}(m_0) \xrightarrow{p} 0.$$

Therefore

$$\Delta_j(\widehat{m}_n) - \Delta_j(m_0) \xrightarrow{p} 0,$$

which combined with the evaluation-sample law of large numbers yields

$$\widehat{\Delta}_{j,n,m} \xrightarrow{p} \Delta_j(m_0).$$

□

Thus, for any learner satisfying the stated L_2 consistency conditions, permutation importance estimates the population loss from breaking a feature. This justifies using the screen to preserve predictive content before the rSCM first stage.

7.3 Appendix C. Consistency of Screened Residual Synthetic Control

7.3.1 Setup

Fix one event-sector panel. Let unit 1 be treated after period T_0, and let donor units be $j = 2, \dots, N$. The result formalizes the case where untreated outcomes combine observable nonlinear state dependence with a donor-representable residual component. Observed outcomes satisfy

$$Y_{it} = \begin{cases} Y_{it}(0), and i \neq 1, \\ Y_{1t}(0), and i = 1,\ t \leq T_0, \\ Y_{1t}(1), and i = 1,\ t > T_0. \end{cases}$$

Assume untreated outcomes admit the decomposition

$$Y_{it}(0) = m_0(X_{it}) + r_{it}(0).$$

Let S_n denote the screened feature set. The nuisance learner, not necessarily RF, produces $\widehat{m}_n$ from donor data only. Define residuals

$$\hat{r}_{it} = Y_{it} - \widehat{m}_n(X_{it,S_n}).$$

The residual SCM weights solve

$$\widehat{w} = \arg\min_{w \in \Delta^{N-1}} \frac{1}{T_0} \sum_{t \le T_0} \left(\hat{r}_{1t} - \sum_{j=2}^{N} w_j \hat{r}_{jt} \right)^2 + \lambda_n \|w\|_2^2,$$

where Δ^{N-1} is the donor simplex. The post-treatment counterfactual is

$$\widehat{Y}_{1t}(0) = \widehat{m}_n(X_{1t,S_n}) + \sum_{j=2}^{N} \widehat{w}_j \hat{r}_{jt}.$$

7.3.2 Assumptions

The assumptions separate the two jobs: creening and nuisance estimation remove predictable untreated structure, while SCM still has to approximate the remaining residual path with stable donor weights.

Assumption C.1 (No anticipation).

$$Y_{1t}(1) = Y_{1t}(0), \qquad t \le T_0.$$

Assumption C.2 (Screening preserves predictive risk).

$$\mathbb{E}\left[(m_0(X) - m_0(X_{S_n}))^2 \right] \to 0.$$

That is, the selected feature set S_n should approximate the same untreated conditional mean as the full candidate set; the screening step may reduce dimensionality, but it should not materially change the squared-loss prediction target.

Assumption C.3 (Nuisance consistency). For every treated and donor unit used in the panel,

$$\frac{1}{T_0} \sum_{t \le T_0} \left(\widehat{m}_n(X_{it,S_n}) - m_0(X_{it,S_n}) \right)^2 \xrightarrow{p} 0$$

and

$$\frac{1}{T_1} \sum_{t > T_0} \left(\widehat{m}_n(X_{it,S_n}) - m_0(X_{it,S_n}) \right)^2 \xrightarrow{p} 0.$$

Assumption C.4 (Pre-period residual approximability). There exists $w^\star \in \Delta^{N-1}$ such that

$$\frac{1}{T_0}\sum_{t\le T_0}\left(r_{1t}(0)-\sum_{j=2}^{N}w_j^\star r_{jt}(0)\right)^2\to 0.$$

Assumption C.5 (Approximate-minimizer stability). Any sequence of simplex weights that approximates the treated residual path in the pre-period also approximates it in the post-period. Formally, if $w_n \in \Delta^{N-1}$ satisfies

$$\frac{1}{T_0}\sum_{t\le T_0}\left(r_{1t}(0)-\sum_{j=2}^{N}w_{n,j} r_{jt}(0)\right)^2\to 0,$$

then

$$\frac{1}{T_1}\sum_{t> T_0}\left(r_{1t}(0)-\sum_{j=2}^{N}w_{n,j} r_{jt}(0)\right)^2\to 0.$$

This is stronger than requiring one oracle vector to work before and after treatment. It rules out cases where many weights fit the pre-period equally well but imply very different post-period residual paths.

Assumption C.6 (Compact weight space and vanishing ridge). Weights are chosen on the simplex and $\lambda_n \to 0$.

7.3.3 Theorem

Theorem C.1 (Consistency of screened residual synthetic control). Under Assumptions C.1–C.6,

$$\frac{1}{T_1}\sum_{t>T_0}\left(\widehat{Y}_{1t}(0)-Y_{1t}(0)\right)^2\xrightarrow{p}0.$$

Consequently,

$$\hat{\tau}_t = Y_{1t}-\widehat{Y}_{1t}(0)$$

is pointwise consistent for each post-treatment period, and the average post-treatment effect

$$\hat{\tau}=\frac{1}{T_1}\sum_{t>T_0}\hat{\tau}_t$$

is consistent for

$$\tau = \frac{1}{T_1} \sum_{t>T_0} \left(Y_{1t}(1) - Y_{1t}(0) \right).$$

7.3.4 Proof

Define nuisance errors

$$\delta_{it} = \widehat{m}_n(X_{it,S_n}) - m_0(X_{it}).$$

By Assumptions C.2 and C.3, δ_{it} is small in mean square for every relevant treated and donor path. Since

$$\widehat{r}_{it} = Y_{it}(0) - \widehat{m}_n(X_{it,S_n}) = r_{it}(0) - \delta_{it},$$

the pre-treatment residual objective can be written as

$$Q_n(w) = \frac{1}{T_0} \sum_{t \le T_0} \left(r_{1t}(0) - \sum_{j=2}^{N} w_j r_{jt}(0) - \delta_{1t} + \sum_{j=2}^{N} w_j \delta_{jt} \right)^2 \; + \; \lambda_n \|w\|_2^2.$$

Let the oracle objective be

$$Q(w) = \frac{1}{T_0} \sum_{t \le T_0} \left(r_{1t}(0) - \sum_{j=2}^{N} w_j r_{jt}(0) \right)^2.$$

Write

$$a_t(w) = r_{1t}(0) - \sum_{j=2}^{N} w_j r_{jt}(0), \qquad d_t(w) = -\delta_{1t} + \sum_{j=2}^{N} w_j \delta_{jt}.$$

Then

$$Q_n(w) - Q(w) = \frac{1}{T_0} \sum_{t \le T_0} \left(2 a_t(w) d_t(w) + d_t(w)^2 \right) + \lambda_n \|w\|_2^2.$$

Because w lies on the simplex,

$$|d_t(w)| \le |\delta_{1t}| + \sum_{j=2}^{N} w_j |\delta_{jt}| \le |\delta_{1t}| + \max_{2\le j\le N} |\delta_{jt}|.$$

Hence, if

$$\bar{\delta}_n = \max_{1\le i\le N} \left(\frac{1}{T_0} \sum_{t\le T_0} \delta_{it}^2 \right)^{1/2},$$

then uniformly over $w \in \Delta^{N-1}$,

$$\frac{1}{T_0} \sum_{t\le T_0} d_t(w)^2 = O_p(\bar{\delta}_n^2).$$

Similarly, by Cauchy-Schwarz,

$$\sup_{w\in\Delta^{N-1}} \left| \frac{1}{T_0} \sum_{t\le T_0} a_t(w) d_t(w) \right| \le \sup_{w\in\Delta^{N-1}} \left(\frac{1}{T_0} \sum_{t\le T_0} a_t(w)^2 \right)^{1/2} \cdot O_p(\bar{\delta}_n).$$

The first factor is $O_p(1)$ because the donor and treated residual paths are square-integrable, and the second vanishes by Assumptions C.2 and C.3. Together with $\lambda_n \to 0$, this yields

$$\sup_{w\in\Delta^{N-1}} |Q_n(w) - Q(w)| \xrightarrow{p} 0.$$

Since the simplex is compact and Q_n converges uniformly to Q, any minimizer $\widehat{w}$ of Q_n is asymptotically an approximate minimizer of Q:

$$Q(\widehat{w}) \le Q(w^\star) + o_p(1).$$

By Assumption C.4, $Q(w^\star) \to 0$. Therefore

$$\frac{1}{T_0} \sum_{t\le T_0} \left(r_{1t}(0) - \sum_{j=2}^{N} \widehat{w}_j r_{jt}(0) \right)^2 \xrightarrow{p} 0.$$

Now decompose the post-treatment counterfactual error:

$$\widehat{Y}_{1t}(0) - Y_{1t}(0) = \Big[\widehat{m}_n(X_{1t,S_n}) - m_0(X_{1t})\Big] + \Big[\sum_{j=2}^{N} \widehat{w}_j \hat{r}_{jt} - r_{1t}(0)\Big].$$

For the residual term, substitute $\hat{r}_{jt} = r_{jt}(0) - \delta_{jt}$:

$$\sum_{j=2}^{N} \widehat{w}_j \hat{r}_{jt} - r_{1t}(0) = \sum_{j=2}^{N} \widehat{w}_j r_{jt}(0) - r_{1t}(0) - \sum_{j=2}^{N} \widehat{w}_j \delta_{jt}.$$

The pre-period result above says that $\widehat{w}$ is an approximate residual minimizer before treatment. Assumption C.5 then carries that approximation to the post-period:

$$\frac{1}{T_1} \sum_{t>T_0} \left(r_{1t}(0) - \sum_{j=2}^{N} \widehat{w}_j r_{jt}(0) \right)^2 \xrightarrow{p} 0.$$

The nuisance-contamination term

$$\sum_{j=2}^{N} \widehat{w}_j \delta_{jt}$$

vanishes because weights are on the simplex and donor nuisance errors vanish in average squared norm.

Combining these pieces gives

$$\frac{1}{T_1} \sum_{t>T_0} \left(\sum_{j=2}^{N} \widehat{w}_j \hat{r}_{jt} - r_{1t}(0) \right)^2 \xrightarrow{p} 0.$$

The nuisance term

$$\frac{1}{T_1} \sum_{t>T_0} \left(\widehat{m}_n(X_{1t,S_n}) - m_0(X_{1t}) \right)^2$$

also vanishes by Assumptions C.2 and C.3. Therefore the full counterfactual error converges to zero in post-period mean square:

$$\frac{1}{T_1} \sum_{t>T_0} \left(\widehat{Y}_{1t}(0) - Y_{1t}(0) \right)^2 \xrightarrow{p} 0.$$

Pointwise ATT consistency follows from

$$\widehat{\tau}_t - \tau_t = Y_{1t}(0) - \widehat{Y}_{1t}(0),$$

and average ATT consistency follows by averaging over $t > T_0$. □

8 References

Abadie, Alberto. 2021. "Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects." *Journal of Economic Literature* 59 (2): 391–425. https://doi.org/10.1257/jel.20191450.

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2010. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." *Journal of the American Statistical Association* 105 (490): 493–505. https://www.nber.org/system/files/working_papers/t0335/t0335.pdf.

Abadie, Alberto, and Javier Gardeazabal. 2003. "The Economic Costs of Conflict: A Case Study of the Basque Country." *American Economic Review* 93 (1): 113–32. https://www.nber.org/system/files/working_papers/w8478/w8478.pdf.

Acharya, Viral V., Lasse Heje Pedersen, Thomas Philippon, and Matthew Richardson. 2017. "Measuring Systemic Risk." *Review of Financial Studies* 30 (1): 2–47. https://doi.org/10.1093/rfs/hhw088.

Adrian, Tobias, Erkko Etula, and Tyler Muir. 2014. "Financial Intermediaries and the Cross-Section of Asset Returns." *The Journal of Finance* 69 (6): 2557–96. https://doi.org/10.1111/jofi.12189.

Adrian, Tobias, and Hyun Song Shin. 2010. "Liquidity and Leverage." *Journal of Financial Intermediation* 19 (3): 418–37. https://doi.org/10.1016/j.jfi.2008.12.002.

Arkhangelsky, Dmitry, Susan Athey, David A. Hirshberg, Guido W. Imbens, and Stefan Wager. 2021. "Synthetic Difference-in-Differences." *American Economic Review* 111 (12): 4088–118. https://doi.org/10.1257/aer.20190159.

Asness, Clifford S., Andrea Frazzini, and Lasse Heje Pedersen. 2019. "Quality Minus Junk." *Review of Accounting Studies* 24 (1): 34–112. https://doi.org/10.1007/s11142-018-9470-2.

Athey, Susan, Mohsen Bayati, Nikolay Doudchenko, Guido W. Imbens, and Khashayar Khosravi. 2021. "Matrix Completion Methods for Causal Panel Data Models." *Journal of the American Statistical Association* 116 (536): 1716–30. https://doi.org/10.1080/01621459.2021.1891924.

Athey, Susan, and Guido W. Imbens. 2017. "The State of Applied Econometrics: Causality and Policy Evaluation." *Journal of Economic Perspectives* 31 (2): 3–32. https://doi.org/10.1257/jep.31.2.3.

Athey, Susan, and Guido W. Imbens. 2019. "Machine Learning Methods That Economists Should Know About." *Annual Review of Economics* 11 (1): 685–725. https://doi.org/10.1146/annurev-economics-080217-053433.

Avramov, Doron, Tarun Chordia, Gergana Jostova, and Alexander Philipov. 2009. "Credit Ratings and the Cross-Section of Stock Returns." *Journal of Financial Markets* 12 (3): 469–99. https://doi.org/10.1016/j.finmar.2009.01.005.

Barroso, Pedro, and Pedro Santa-Clara. 2015. "Momentum Has Its Moments." *Journal of Financial Economics* 116 (1): 111–20. https://doi.org/10.1016/j.jfineco.2014.11.010.

Bekaert, Geert, Marie Hoerova, and Marco Lo Duca. 2013. "Risk, Uncertainty and Monetary Policy." *Journal of Monetary Economics* 60 (7): 771–88. https://doi.org/10.1016/j.jmoneco.2013.06.003.

Ben-Michael, Eli, Avi Feller, and Jesse Rothstein. 2021. "The Augmented Synthetic Control Method." *Journal of the American Statistical Association* 116 (536): 1789–803. https://www.nber.org/system/files/working_papers/w28885/w28885.pdf.

Bernanke, Ben S., Mark Gertler, and Simon Gilchrist. 1999. "The Financial Accelerator in a Quantitative Business Cycle Framework." Chap. 21 in *Handbook of Macroeconomics*, edited by John B. Taylor and Michael Woodford, vol. 1. Elsevier. https://doi.org/10.1016/S1574-0048(99)10034-X.

Bernanke, Ben S., and Kenneth N. Kuttner. 2005. "What Explains the Stock Market's Reaction to Federal Reserve Policy?" *The Journal of Finance* 60 (3): 1221–57. https://doi.org/10.1111/j.1540-6261.2005.00760.x.

Bloom, Nicholas. 2009. "The Impact of Uncertainty Shocks." *Econometrica* 77 (3): 623–85. https://doi.org/10.3982/ECTA6248.

Bloom, Nicholas, Max Floetotto, Nir Jaimovich, Itay Saporta-Eksten, and Stephen J. Terry. 2018. "Really Uncertain Business Cycles." *Econometrica* 86 (3): 1031–65. https://doi.org/10.3982/ECTA10927.

Breiman, Leo. 2001. "Random Forests." *Machine Learning* 45 (1): 5–32. https://doi.org/10.1023/A:1010933404324.

Brunnermeier, Markus K., and Lasse Heje Pedersen. 2009. "Market Liquidity and Funding Liquidity." *Review of Financial Studies* 22 (6): 2201–38. https://doi.org/10.1093/rfs/hhn098.

Campbell, John Y., Jens Hilscher, and Jan Szilagyi. 2008. "In Search of Distress Risk." *The Journal of Finance* 63 (6): 2899–939. https://doi.org/10.1111/j.1540-6261.2008.01416.x.

Chen, Long, Pierre Collin-Dufresne, and Robert S. Goldstein. 2009. "On the Relation Between the Credit Spread Puzzle and the Equity Premium Puzzle." *Review of Financial Studies* 22 (9): 3367–409. https://doi.org/10.1093/rfs/hhn078.

Chernozhukov, Victor, Denis Chetverikov, Mert Demirer, et al. 2018. "Double/Debiased Machine Learning for Treatment and Structural Parameters." *The Econometrics Journal* 21 (1): C1–68. https://doi.org/10.1111/ectj.12097.

Chodorow-Reich, Gabriel. 2014. "The Employment Effects of Credit Market Disruptions: Firm-Level Evidence from the 2008–9 Financial Crisis." *Quarterly Journal of Economics* 129 (1): 1–59. https://doi.org/10.1093/qje/qjt031.

Cochrane, John H. 2011. "Discount Rates." *The Journal of Finance* 66 (4): 1047–108. https://doi.org/10.1111/j.1540-6261.2011.01671.x.

Collin-Dufresne, Pierre, Robert S. Goldstein, and J. Spencer Martin. 2001. "The Determinants of Credit Spread Changes." *The Journal of Finance* 56 (6): 2177–207. https://doi.org/10.1111/0022-1082.00402.

Doudchenko, Nikolay, and Guido W. Imbens. 2016. *Balancing, Regression, Difference-in-Differences and Synthetic Control Methods: A Synthesis*. NBER Working Paper No. 22791. National Bureau of Economic Research. https://doi.org/10.3386/w22791.

Duffie, Darrell, and Kenneth J. Singleton. 2003. *Credit Risk: Pricing, Measurement, and Management*. Princeton University Press. https://www.nber.org/papers/w9270.

Elton, Edwin J., Martin J. Gruber, Deepak Agrawal, and Christopher Mann. 2001. "Explaining the Rate Spread on Corporate Bonds." *The Journal of Finance* 56 (1): 247–77. https://doi.org/10.1111/0022-1082.00324.

Engle, Robert F., Eric Jondeau, and Michael Rockinger. 2015. "Systemic Risk in Europe." *Review of Finance* 19 (1): 145–90. https://doi.org/10.1093/rof/rfu012.

Fama, Eugene F., and Kenneth R. French. 1989. "Business Conditions and Expected Returns on Stocks and Bonds." *Journal of Financial Economics* 25 (1): 23–49. https://doi.org/10.1016/0304-405X(89)90095-0.

Forbes, Kristin J., and Roberto Rigobon. 2002. "No Contagion, Only Interdependence: Measuring Stock Market Comovements." *The Journal of Finance* 57 (5): 2223–61. https://doi.org/10.1111/0022-1082.00494.

Frazzini, Andrea, and Lasse Heje Pedersen. 2014. "Betting Against Beta." *Journal of Financial Economics* 111 (1): 1–25. https://doi.org/10.1016/j.jfineco.2013.10.005.

Friewald, Nils, Christian Wagner, and Josef Zechner. 2014. "The Cross-Section of Credit Risk Premia and Equity Returns." *The Journal of Finance* 69 (6): 2419–69. https://doi.org/10.1111/jofi.12143.

Gârleanu, Nicolae, and Lasse Heje Pedersen. 2011. "Margin-Based Asset Pricing and Deviations from the Law of One Price." *Review of Financial Studies* 24 (6): 1980–2022. https://doi.org/10.1093/rfs/hhq138.

Genuer, Robin, Jean-Michel Poggi, and Christine Tuleau-Malot. 2010. "Variable Selection Using Random Forests." *Pattern Recognition Letters* 31 (14): 2225–36. https://doi.org/10.1016/j.patrec.2010.03.014.

Gertler, Mark, and Cara S. Lown. 1999. "The Information in the High-Yield Bond Spread for the Business Cycle: Evidence and Some Implications." *Oxford Review of Economic Policy* 15 (3): 132–50. https://doi.org/10.1093/oxrep/15.3.132.

Gilchrist, Simon, Vladimir Yankov, and Egon Zakrajšek. 2009. "Credit Market Shocks and Economic Fluctuations: Evidence from Corporate Bond and Stock Markets." *Journal of Monetary Economics* 56 (4): 471–93. https://doi.org/10.1016/j.jmoneco.2009.03.017.

Gilchrist, Simon, and Egon Zakrajšek. 2012. "Credit Spreads and Business Cycle Fluctuations." *American Economic Review* 102 (4): 1692–720. https://doi.org/10.1257/aer.102.4.1692.

Gregorutti, Baptiste, Bertrand Michel, and Philippe Saint-Pierre. 2017. "Correlation and Variable Importance in Random Forests." *Statistics and Computing* 27 (3): 659–78. https://doi.org/10.1007/s11222-016-9646-1.

Gu, Shihao, Bryan T. Kelly, and Dacheng Xiu. 2020. "Empirical Asset Pricing via Machine Learning." *The Review of Financial Studies* 33 (5): 2223–73. https://doi.org/10.1093/rfs/hhaa009.

Hapfelmeier, Alexander, Torsten Hothorn, Kurt Ulm, and Carolin Strobl. 2014. "A New Variable Importance Measure for Random Forests with Missing Data." *Statistics and Computing* 24 (1): 21–34. https://doi.org/10.1007/s11222-012-9349-1.

He, Zhiguo, and Arvind Krishnamurthy. 2013. "Intermediary Asset Pricing." *American Economic Review* 103 (2): 732–70. https://doi.org/10.1257/aer.103.2.732.

Hou, Kewei, Chen Xue, and Lu Zhang. 2015. "Digesting Anomalies: An Investment Approach." *The Review of Financial Studies* 28 (3): 650–705. https://doi.org/10.1093/rfs/hhu068.

Jorda, Oscar, Moritz Schularick, and Alan M. Taylor. 2013. "When Credit Bites Back." *Journal of Money, Credit and Banking* 45 (s2): 3–28. https://doi.org/10.1111/jmcb.12069.

Jurado, Kyle, Sydney C. Ludvigson, and Serena Ng. 2015. "Measuring Uncertainty." *American Economic Review* 105 (3): 1177–216. https://doi.org/10.1257/aer.20131193.

Kelly, Bryan T., Seth Pruitt, and Yinan Su. 2019. "Characteristics Are Covariances: A Unified Model of Risk and Return." *Journal of Financial Economics* 134 (3): 501–24. https://doi.org/10.1016/j.jfineco.2019.05.001.

Kelly, Bryan, Hanno Lustig, and Stijn Van Nieuwerburgh. 2016. "Too-Systemic-to-Fail: What Option Markets Imply about Sector-Wide Government Guarantees." *American Economic Review* 106 (6): 1278–319. https://doi.org/10.1257/aer.20120389.

Kiyotaki, Nobuhiro, and John Moore. 1997. "Credit Cycles." *Journal of Political Economy* 105 (2): 211–48. https://doi.org/10.1086/262072.

Kothari, S. P., and Jerold B. Warner. 2007. "Econometrics of Event Studies." In *Handbook of Empirical Corporate Finance*, vol. 1. Elsevier. https://doi.org/10.1016/B978-0-444-53265-7.50015-9.

Longin, Francois, and Bruno Solnik. 2001. "Extreme Correlation of International Equity Markets." *The Journal of Finance* 56 (2): 649–76. https://doi.org/10.1111/0022-1082.00340.

MacKinlay, A. Craig. 1997. "Event Studies in Economics and Finance." *Journal of Economic Literature* 35 (1): 13–39. https://www.jstor.org/stable/2729691.

Merton, Robert C. 1974. "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates." *The Journal of Finance* 29 (2): 449–70. https://doi.org/10.1111/j.1540-6261.1974.tb03058.x.

Moore, Nils aus dem, Johannes Brehm, and Henri Gruhl. 2025. "Driving Innovation? Carbon Tax Effects in the Swedish Transport Sector." *Journal of Public Economics* 248: 105444. https://doi.org/10.1016/j.jpubeco.2025.105444.

Moskowitz, Tobias J., and Mark Grinblatt. 1999. "Do Industries Explain Momentum?" *The Journal of Finance* 54 (4): 1249–90. https://doi.org/10.1111/0022-1082.00146.

Philippon, Thomas. 2009. "The Bond Market's q." *Quarterly Journal of Economics* 124 (3): 1011–56. https://doi.org/10.1162/qjec.2009.124.3.1011.

Ramosaj, Burim, and Markus Pauly. 2019. "Asymptotic Unbiasedness of the Permutation Importance Measure in Random Forest Models." *arXiv Preprint arXiv:1912.03306*. https://arxiv.org/pdf/1912.03306.pdf.

Scornet, Erwan, Gérard Biau, and Jean-Philippe Vert. 2014. "Consistency of Random Forests." *Working Paper*. https://arxiv.org/abs/1405.2881.

Shin, Hyun Song. 2010. *Risk and Liquidity*. BIS Working Paper No. 313. Bank for International Settlements. https://www.bis.org/publ/work313.pdf.

Stock, James H., and Mark W. Watson. 2012. "Disentangling the Channels of the 2007–09 Recession." *Brookings Papers on Economic Activity* 43 (1): 81–156. https://doi.org/10.1353/eca.2012.0005.

Sun, Liyang, and Sarah Abraham. 2021. "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects." *Journal of Econometrics* 225 (2): 175–99. https://doi.org/10.1016/j.jeconom.2020.09.006.

Wager, Stefan, and Susan Athey. 2018. "Estimation and Inference of Heterogeneous Treatment Effects Using Random Forests." *Journal of the American Statistical Association* 113 (523): 1228–42. https://doi.org/10.1080/01621459.2017.1319839.

Xu, Yiqing. 2017. "Generalized Synthetic Control Method: Causal Inference with Interactive Fixed Effects Models." *Political Analysis* 25 (1): 57–76. https://doi.org/10.1017/pan.2016.2.