

7524987-cross-asset-shock-diffusion-a-reproducible-test-of-residual-underreacti on-shock-

SUMMARY: This paper examines whether differences in the speed at which traded assets respond to a common market shock can predict subsequent relative returns. The study combines a lagged rolling factor model with Absorption Gap (AG) and Shock Coherence (SC) to measure asset response errors and market states, respectively. The sample includes 24 ETFs from January 4, 2010, to August 28, 2026. The combined signal's executable next-open Rank IC is -0.00020, and its separate close-to-close residual Rank IC is -0.00592. The gross Sharpe ratio is 0.006 with a CAGR of -0.07%, and the net Sharpe ratio is -2.367 with a CAGR of -9.83%. The study rejects the disclosed public specification and identifies the failure modes that the separate proprietary layer was designed to address. The findings suggest that the public specification does not provide economically viable evidence of alpha.

METHOD: The study uses a sample of 24 ETFs from January 4, 2010, to August 28, 2026. The sample excludes eight factor proxies. The methodology combines a lagged rolling factor model with Absorption Gap (AG) and Shock Coherence (SC) to measure asset response errors and market states, respectively. The combined signal's executable next-open Rank IC is -0.00020, and its separate close-to-close residual Rank IC is -0.00592. The gross Sharpe ratio is 0.006 with a CAGR of -0.07%, and the net Sharpe ratio is -2.367 with a CAGR of -9.83%.

KEY FINDINGS:

- The combined signal's executable next-open Rank IC is statistically indistinguishable from chance.
- The standalone AG Rank IC is statistically indistinguishable from chance.
- The randomized-signal benchmark's empirical p-value is 0.968.
- The difference between the AG long-short spread on high- and low-coherence dates is -3.15 basis points, and its confidence interval includes zero.
- Logistic regression and gradient boosting remain close to random classification.

IMPLICATION FOR TRADING: The study suggests that the public specification does not provide economically viable evidence of alpha. The findings indicate that the disclosed daily ETF implementation does not isolate the underlying economic mechanism well enough to support a public trading claim. The study's results highlight the importance of understanding the limitations of the public specification and the need for a more thorough examination of the underlying economic mechanism to support trading strategies.

QUANTITATIVE FINANCE RESEARCH

Cross-Asset Shock Diffusion

A Reproducible Test of Residual Underreaction, Shock Coherence, and Trading Economics

Alina Khaybullina

Independent Researcher / 30 August 2026

Abstract

This paper examines whether differences in the speed with which traded assets respond to a common market shock can predict subsequent relative returns. The public framework combines a lagged rolling factor model with Absorption Gap (AG), which measures an asset's response error, and Shock Coherence (SC), which describes the market state on that date. SC has the same value for every asset and therefore serves only as an activation filter; it cannot change the cross-sectional ordering of securities. A signal formed after close t is executed at open $t+1$ and earns the return from open $t+1$ to open $t+2$.

The sample contains 24 ETFs from 4 January 2010 through 28 August 2026; eight factor proxies are excluded from the 16-asset traded cross-section. The combined signal's executable next-open Rank IC is -0.00020 ; its separate close-to-close residual Rank IC is -0.00592 and is treated only as a mechanism diagnostic. Zero-cost gross Sharpe is 0.006 with CAGR -0.07% ; at 5 bps, net Sharpe is -2.367 with CAGR -9.83% . Annualized arithmetic cost drag is approximately 10.28% . A within-date randomized-signal benchmark applied to the executable endpoint gives an empirical two-sided p-value of 0.968 . The public specification therefore provides no economically viable evidence of alpha.

On the separate close-to-close mechanism endpoint, standalone one-day AG Rank IC is -0.00265 (HAC $t=-0.62$; 95% moving-block interval $[-0.01025, 0.00539]$). The difference between the AG long-short spread on high- and low-coherence dates is -3.15 basis points at one day (HAC $t=-1.23$), and its confidence interval includes zero. Gross results change sign across chronological segments, while all six segments lose money after 5-bps costs. Logistic regression and gradient boosting remain close to random classification. These findings reject the disclosed public specification and define the failure modes that the separate private research layer was designed to address.

Keywords: return predictability; information diffusion; statistical arbitrage; factor models; exchange-traded funds; transaction costs

JEL: G11, G12, G14, C58

Research contribution

CONTRIBUTION The paper connects an economic hypothesis to an executable next-open portfolio, identifies why the disclosed implementation fails, and separates reproducible public evidence from the deliberately withheld proprietary alpha layer.

Results at a glance

Test	Estimate	Inference	Status
Combined signal, executable endpoint	Rank IC -0.00020	Indistinguishable from chance	Audited baseline
Gross portfolio	Sharpe 0.006 ; CAGR -0.07%	Approximately flat before costs	Audited baseline
Net portfolio, 5 bps	Sharpe -2.367 ; CAGR -9.83%	Cost drag $\sim 10.28\%$ p.a.	Audited baseline
Standalone AG, residual endpoint	Rank IC -0.00265	HAC $t=-0.62$; $p=0.536$	Diagnostic
AG by SC state, 1 day	-3.15 bps high-low	HAC $t=-1.23$; $p=0.218$	Diagnostic
Randomized-signal benchmark	Empirical $p=0.968$	Observed rank correlation is not unusual	Falsification
Chronological ML	ROC-AUC $0.501 / 0.494$	Near chance	Exploratory

1. Introduction

Financial information is incorporated through a decentralized market process. A broad economic shock can affect equity indices, sectors, rates, credit, commodities, currencies, and individual securities at different speeds and magnitudes. If those response differences contain structure rather than noise, relative adjustment may contain predictive information. That intuition connects momentum, lead-lag effects, cross-predictability, and price-delay research, but it does not by itself establish a trading edge.

The question studied here is deliberately narrow: conditional on a coherent common shock, does the gap between an asset's expected and observed reaction predict its subsequent relative return? A residual is economically ambiguous. It can represent slow information absorption, factor-model error, an offsetting idiosyncratic event, stale exposure estimates, microstructure noise, or an overshoot that later reverses. The surrounding market state must therefore be measured independently from the residual itself.

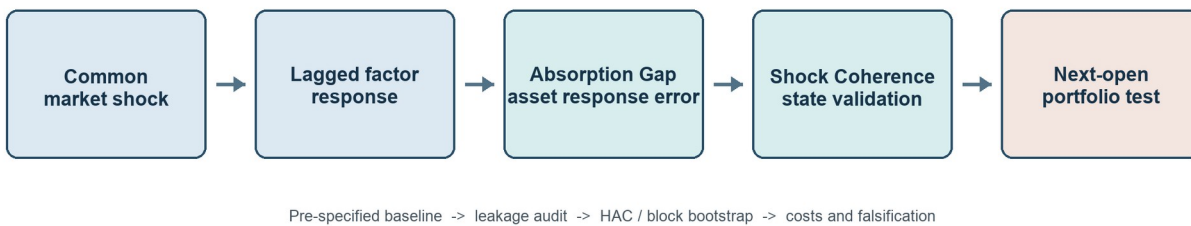


Figure 1. Research architecture. Observable close-t information forms AG and SC; positions become effective no earlier than the next open. The empirical layer freezes the specification before diagnostic analysis.

The study makes four contributions. It defines AG as a signed response error against exposures estimated only from prior observations; treats SC as a market-state filter rather than a source of cross-sectional rank; evaluates the signal with executable next-open timing and explicit transaction costs; and documents why the public specification fails. The diagnostic work separates the long and short books, compares the observed rank correlation with randomized signals, examines chronological subsamples, tests alternative normalizations, removes feature groups in sequence, and measures sensitivity to market regimes and portfolio breadth.

The public baseline is reproducible and timing-consistent, but its ranking performance is statistically indistinguishable from random ordering. The long book has positive gross performance, whereas the short book offsets it; daily rebalancing then creates sufficient turnover for modest trading costs to dominate the result. These findings do not rule out the underlying economic mechanism. They show that the disclosed daily ETF implementation does not isolate it well enough to support a public trading claim.

2. Related Literature and Economic Motivation

2.1 Gradual information diffusion and delayed response

The framework follows the literature on underreaction, momentum, and heterogeneous information processing. Jegadeesh and Titman (1993) establish intermediate-horizon momentum; Hong and Stein (1999) model gradual diffusion that can produce underreaction followed by overreaction; and Hou and Moskowitz (2005) connect price delay to market frictions. These findings motivate a conditional, event-level response measure rather than a universal assumption that all residuals should continue or reverse.

2.2 Cross-security transmission

Lead-lag and economic-link studies show that related securities need not adjust simultaneously. Lo and MacKinlay (1990) demonstrate economically important cross-autocovariances; Moskowitz and Grinblatt (1999) emphasize industry components; Cohen and Frazzini (2008) document predictability through customer-supplier links; and Menzly and Ozbas (2010) identify cross-predictability across economically related industries. The present study asks whether a public, traded-factor representation can operationalize the same broad idea at daily frequency.

2.3 Why a trading test is harder than an association test

Statistical association and investable performance are different tests. A signal with a positive coefficient can still fail when translated into a small cross-section with position limits, next-open execution, short exposure, frequent rebalancing, and transaction costs. Sharpe-ratio inference is sensitive to serial dependence and non-normality (Lo, 2002), while repeated strategy search creates selection bias (Bailey and Lopez de Prado, 2014). The analysis therefore labels the pre-specified public baseline separately from diagnostics performed after observing its result and from hypotheses reserved for future data.

3. Model and Hypotheses

3.1 Lagged common-response model

For asset i on date t , the observed return is written as a common-factor response plus an asset-specific residual:

$$r_{i,t} = \alpha_{i,t} + \beta_{i,t}' F_t + \varepsilon_{i,t}.$$

The coefficients used on date t are estimated from observations ending at $t-1$. The public baseline uses a 126-day rolling window, requires at least 60 observations, and applies a ridge penalty of 0.001. Ridge regularization stabilizes the coefficient estimates when market, equity, rates, credit, commodity, and currency proxies move together; it does not make the resulting coefficients causal or eliminate residual cross-asset dependence.

3.2 Absorption Gap

AG is the factor-implied common response minus the observed close-to-close return:

$$AG_{i,t} = \hat{r}_{i,t}^{\text{common}} - r_{i,t}.$$

AG is normalized using a trailing median and median absolute deviation whose window ends at $t-1$. Multiplication by the sign of the expected common response makes a positive signed gap represent underreaction for either shock direction. This definition is mechanically clear but not causally identified: a large gap can still reflect omitted factors or idiosyncratic news.

3.3 Shock Coherence

SC is the equal-weight average of seven historically normalized market-state variables: shock magnitude, directional breadth, inverse cross-sectional dispersion, abnormal volume, intraday confirmation, directional close location, and confirmation across related assets. Each component is limited to the interval $[-5, 5]$, and the score is calculated only when all seven components are available. Because every security receives the same SC value on a given date, a cross-sectional rank correlation for SC alone is undefined.

$$SC_t = (1/7) \sum_{k=1}^7 z_{k,t}.$$

3.4 Pre-specified public signal and hypotheses

The public score retains signed AG on positive-coherence dates and is zero otherwise:

$$s_{i,t} = \text{signed_}AG_{i,t} \cdot \mathbb{I}(SC_t > 0).$$

This specification treats SC as a market-state switch rather than a cross-sectional multiplier. The empirical tests ask whether AG predicts future residual returns, whether its long-short spread differs across coherence states, whether the resulting ranks are monotonic, whether apparent underreaction persists or reverses, how quickly any relationship decays, and whether it survives implementation costs. Variants introduced after the baseline result are reported only as diagnostics or future research proposals.

4. Data and Reproducible Research Design

4.1 Universe and sample

The cached adjusted-Yahoo sample spans 4 January 2010 to 28 August 2026 and contains 24 ETFs. Eight broad proxies—SPY, QQQ, IWM, IEF, HYG, UUP, GLD, and USO—enter the common-factor model and are excluded from the traded cross-section, leaving 16 assets. The universe provides transparent and liquid public instruments, but it is not a point-in-time institutional dataset. Adjusted daily OHLCV histories may be revised by the vendor and cannot reproduce opening-auction fills, bid-ask spreads, or historical product availability as observed in real time.

Instrument-level start and end dates, usable observation counts, and coverage percentages are retained in the machine-readable data-coverage table.

4.2 Timing and leakage controls

Every rolling exposure and normalization used on date t is estimated from information available by $t-1$. Features observed at close t determine a target position only after that close; the position becomes effective at open $t+1$ and earns the return from open $t+1$ to open $t+2$. The primary ranking endpoint uses that same executable open-to-open interval. Future close-to-close factor residuals are reported separately as mechanism diagnostics: they begin at close t and include the overnight move before the strategy can trade, so they are not presented as realizable alpha. Chronological folds include a one-day embargo, and missing observations are never filled from future values. This convention prevents look-ahead in the daily data, although it cannot recover the order of events within a daily bar.

4.3 Portfolio and costs

The portfolio buys the highest-scoring 20% of active assets and shorts the lowest-scoring 20%. Within each side, weights are proportional to signal strength and normalized to 0.5, producing a gross exposure target of 1.0 subject to a 0.15 position cap. The public constructor is dollar-neutral but does not impose market-beta or sector neutrality. Daily turnover is the sum of absolute changes in effective weights, and the primary net result deducts 5 basis points for each unit of turnover.

4.4 Correlation, dependence, and statistical inference

Market and asset returns are correlated by construction: equity sectors share the equity market, credit and rates respond to common macroeconomic news, and several factor proxies are themselves highly correlated. The rolling ridge model removes the component explained by the selected close-to-close factors before AG is measured. The remaining residual should therefore be read as 'unexplained by this factor set,' not as a pure idiosyncratic shock. The open-to-open factor-attribution design matrix has a condition number near 426, confirming material multicollinearity; ridge regularization in the signal model reduces numerical instability but does not remove model-selection risk or omitted common exposures.

Cross-sectional predictability is summarized first at the date level using Spearman rank correlation or a daily top-minus-bottom spread. Inference is then conducted across dates, which avoids treating correlated securities on the same day as independent observations. Heteroskedasticity-and-autocorrelation-consistent standard errors address serial dependence, and 5,000 moving-block bootstrap samples with block length 20 preserve short-run time clustering. The randomized-signal benchmark shuffles asset ranks within each date and asks whether the observed mean rank correlation is unusual relative to random ordering.

4.5 Research status

The complete historical sample was examined before the extended diagnostics, so no portion of it remains a genuinely untouched holdout. Chronological segments describe stability through time rather than new out-of-sample confirmation. Deflated Sharpe is not reported because the number of independent strategy trials cannot be established reliably. Any revised specification must be fixed before it is evaluated on observations after 28 August 2026 or on an independent point-in-time dataset.

INTEGRITY RULE The six chronological periods are stability diagnostics, not a new untouched holdout. Conditional variants, alternative horizons, and machine-learning results do not replace the conclusion for the pre-specified public baseline.

5. Empirical Results

5.1 Corrected baseline and primary rejection

The audit intentionally supersedes the legacy portfolio metrics because the previous engine attached to an effective open-t weight the return that ended at open t. Under corrected timing, executable combined-signal Rank IC is -0.000195 , zero-cost Sharpe is 0.006 , 5-bps annualized return is -9.83% , net Sharpe is -2.367 , maximum drawdown is -82.56% , and average daily turnover is 81.58% . The legacy discrepancy is recorded rather than concealed.

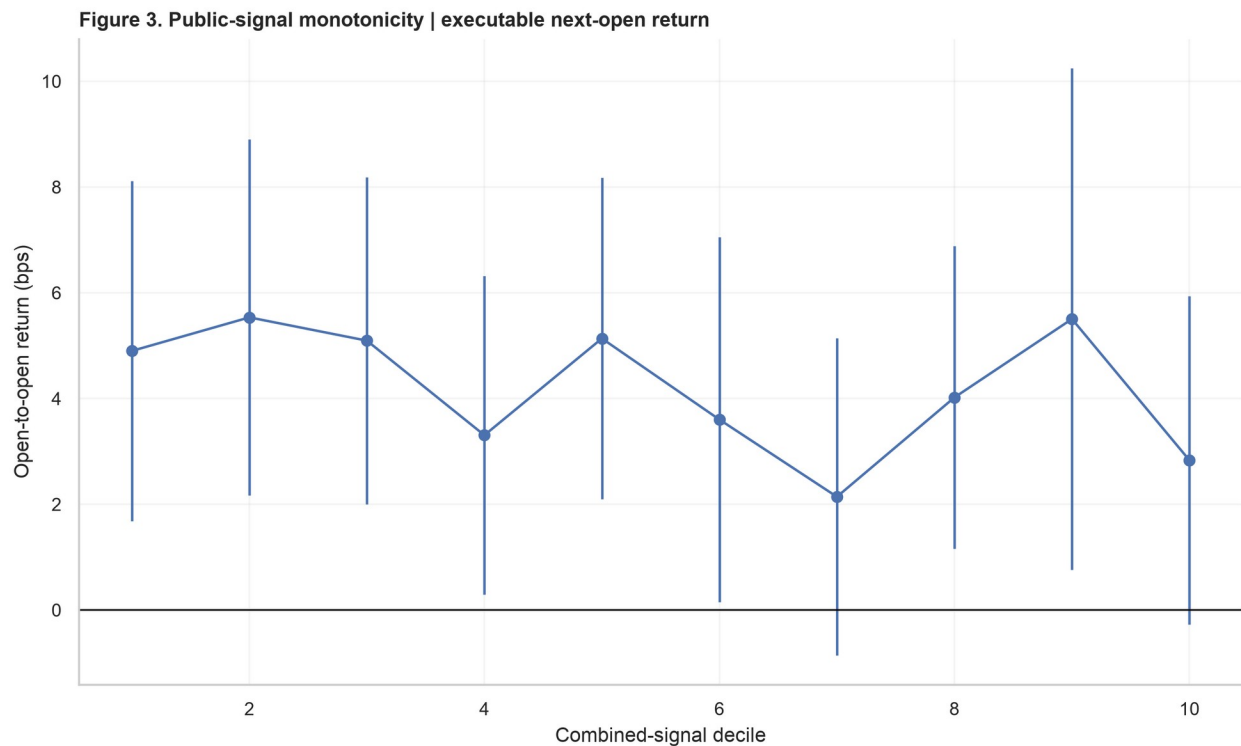


Figure 2. Average executable open $t+1$ -to-open $t+2$ return by public-signal decile. A useful ranking signal should produce an orderly increase from the lowest to the highest decile; this pattern is absent. Source: figure_data/combined_signal_deciles.csv.

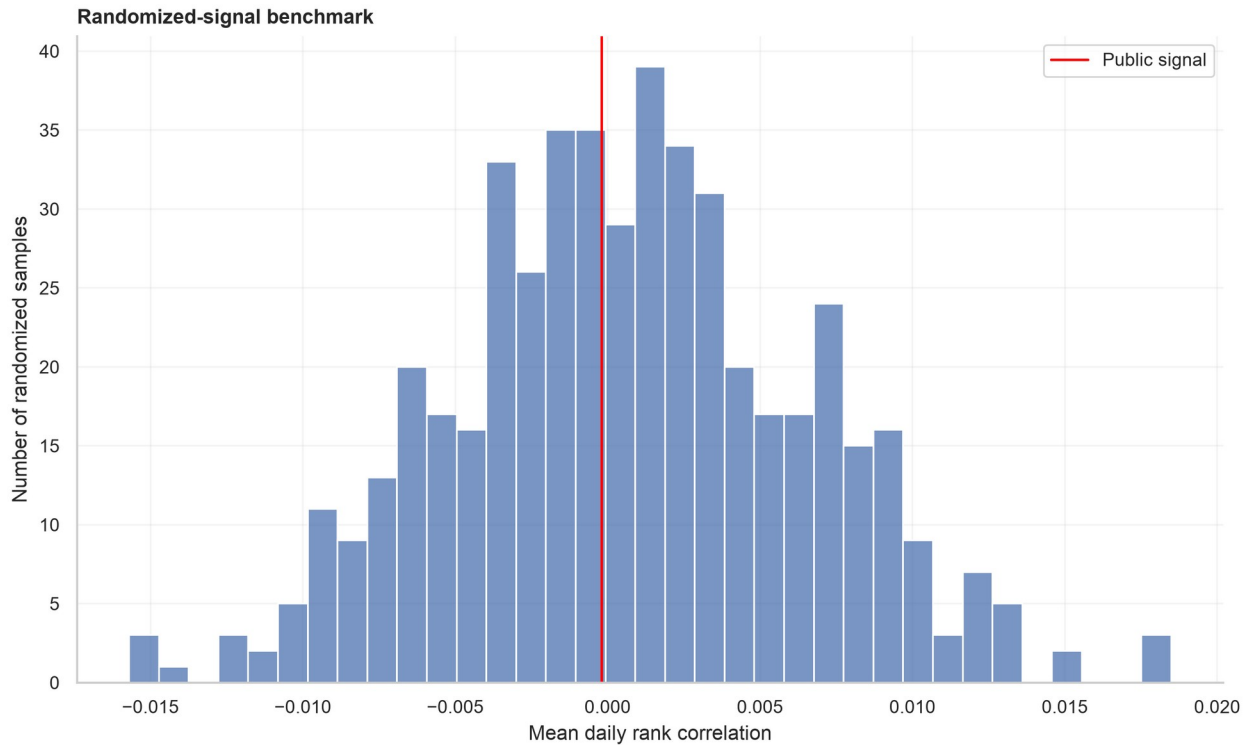


Figure 3. Randomized-signal benchmark. The histogram shows mean rank correlations obtained after randomly reordering assets within each date; the red line marks the public signal. The observed value is not unusual under random ordering (two-sided $p=0.968$). Source: figure_data/placebo_distributions.csv.

5.2 Absorption Gap as a standalone predictor

On the close-to-close residual endpoint, standalone signed AG has one-day Rank IC -0.00265 , HAC $t=-0.62$, $p=0.536$, and a 95% block-bootstrap interval of $[-0.01025, 0.00539]$. The interval spans economically small positive and negative effects. Because this endpoint starts at close t , it is evidence about the proposed response mechanism rather than realizable next-open performance. Longer-horizon estimates remain inconsistent with a stable positive propagation profile.

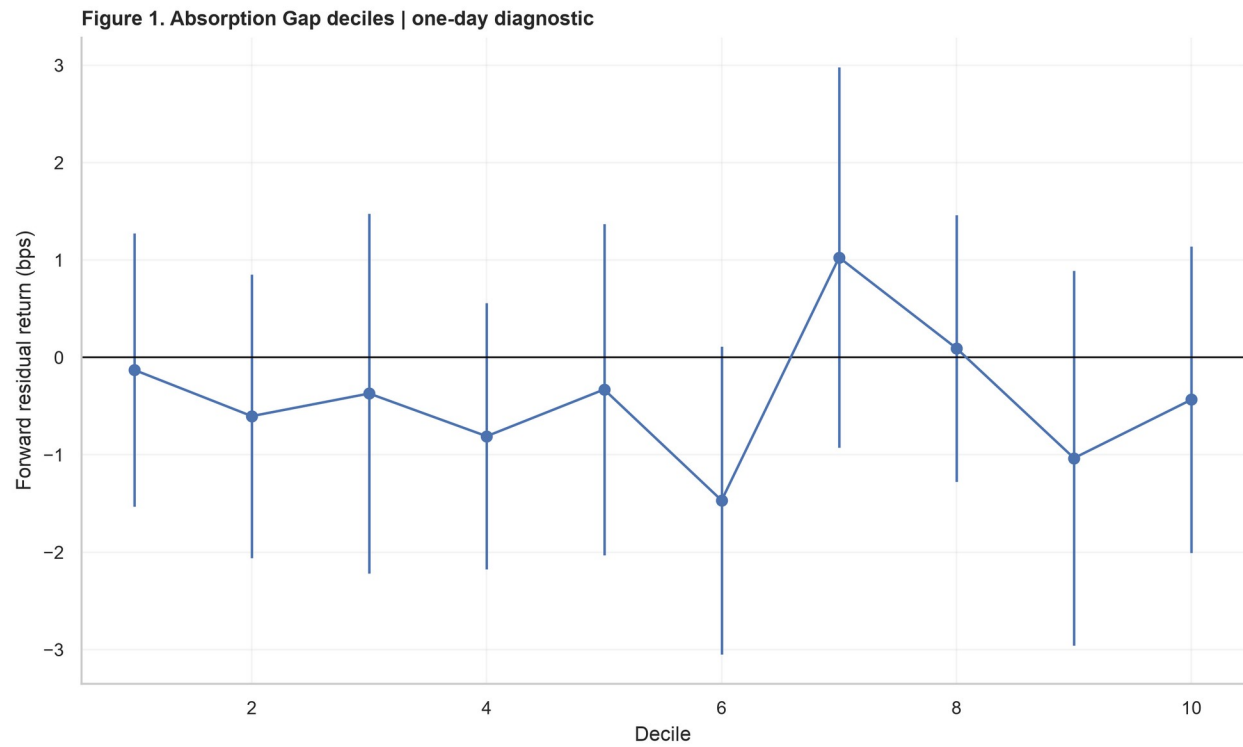


Figure 4. Average next-day residual return by Absorption Gap decile. Returns do not rise consistently from the lowest to the highest AG group. Source: figure_data/absorption_gap_deciles.csv.

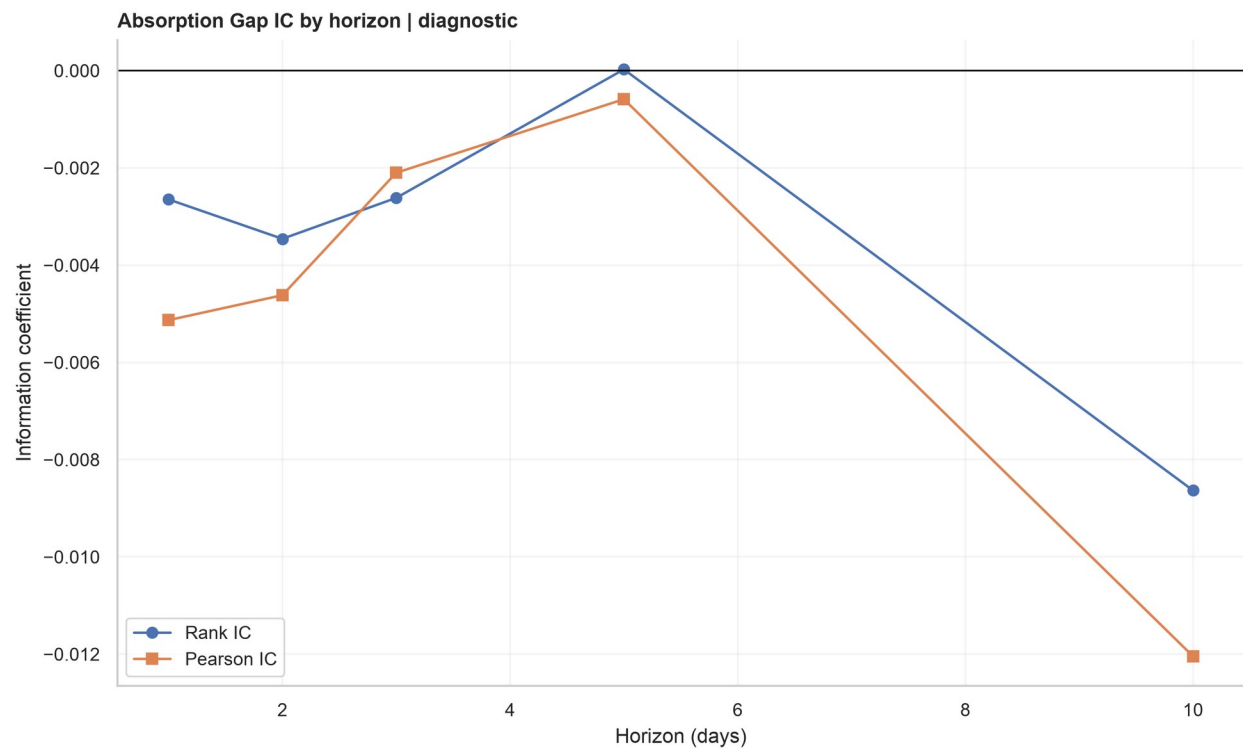


Figure 5. Correlation between Absorption Gap ranks and future residual-return ranks. Estimates remain close to zero across forecast horizons and do not form a stable positive pattern. Source: figure_data/absorption_gap_ic_by_horizon.csv.

5.3 Shock Coherence and the interaction hypothesis

SC is a date-level state measure rather than a cross-sectional ranking variable. Its standalone tests examine future absolute returns, volatility, and average directional returns; they do not establish independent directional information. The conditional test asks whether AG top-minus-bottom spreads differ between high- and low-coherence dates. Coherence groups are assigned from expanding historical distributions rather than full-sample thresholds. Low and high dates are treated as distinct samples using a HAC indicator regression and a block bootstrap. At one day, the high-minus-low difference is -3.15 bps (HAC $t=-1.23$, $p=0.218$); at two days it is -6.36 bps ($p=0.066$). Both signs oppose the proposed strengthening mechanism.

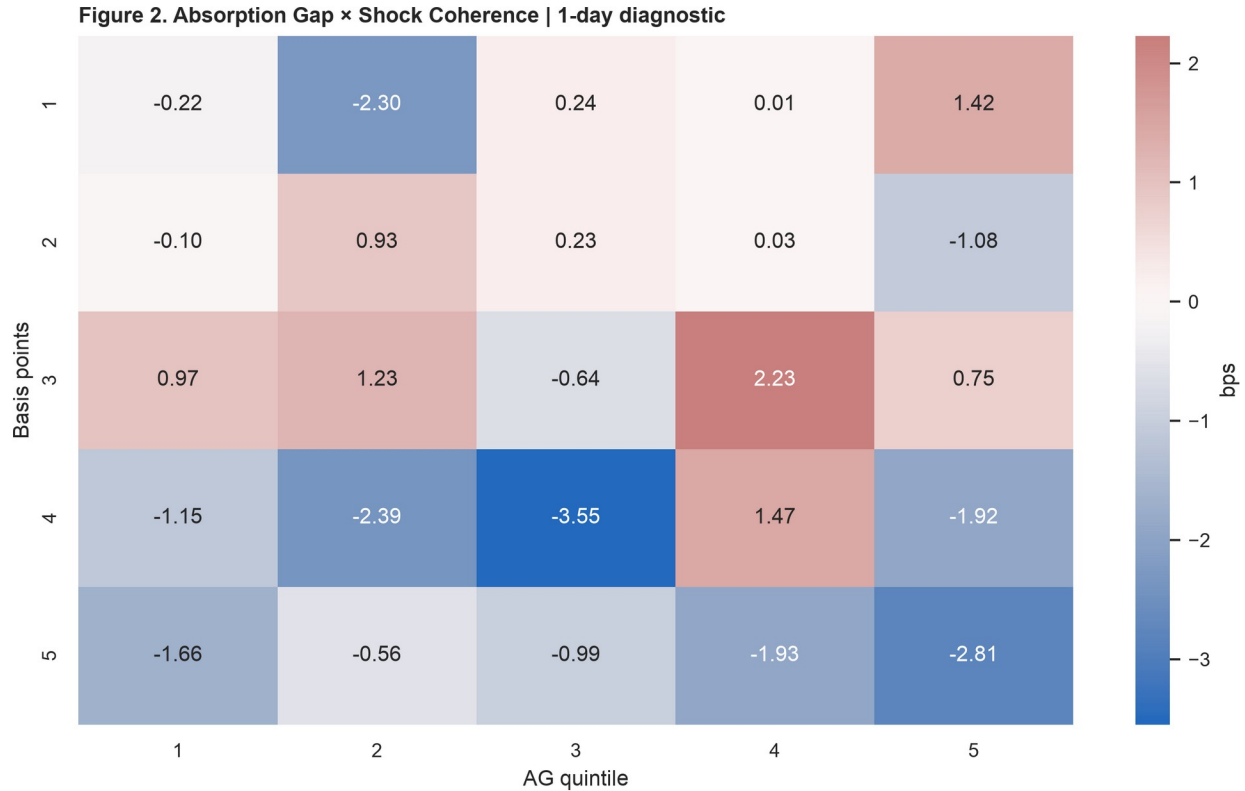


Figure 6. One-day residual returns by Absorption Gap group and market-coherence state. The long-short AG spread is weaker, not stronger, on high-coherence dates. Source: figure_data/ag_x_sc_h1.csv.

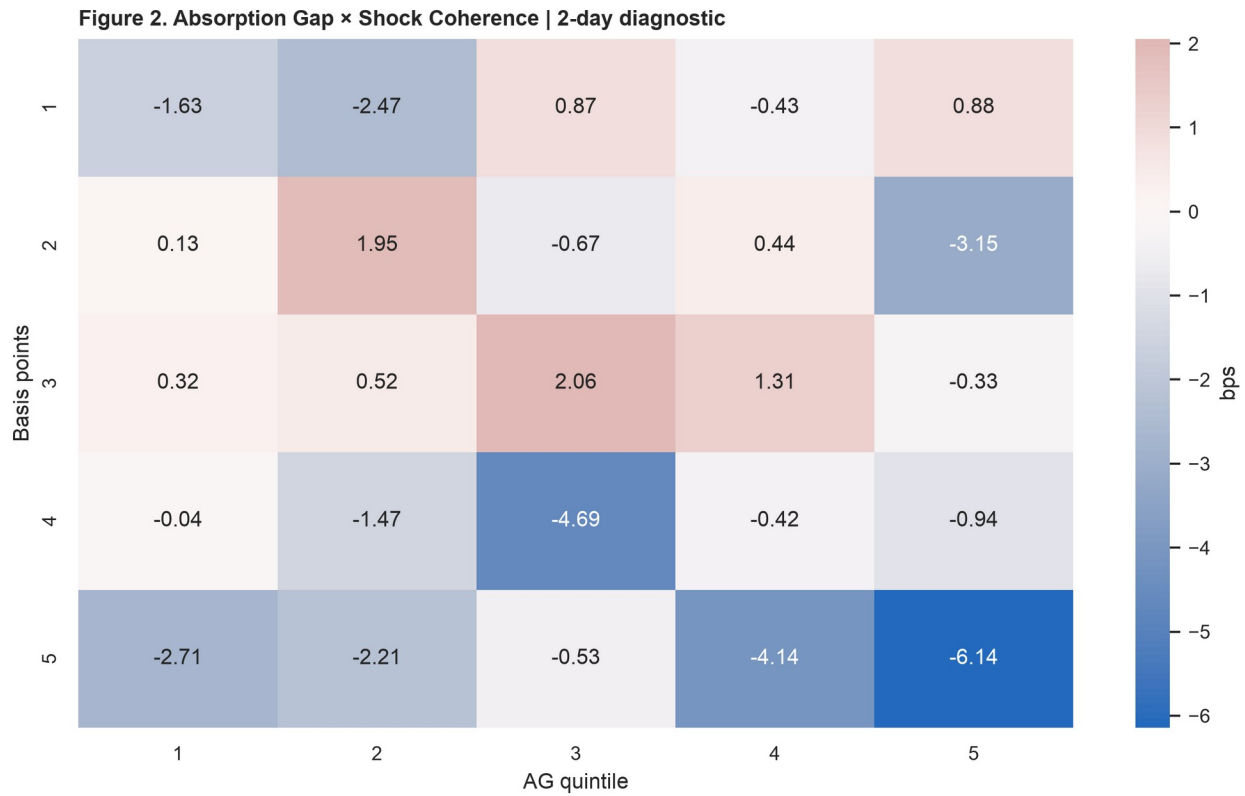


Figure 7. Two-day residual returns by Absorption Gap group and market-coherence state. The high-versus-low coherence difference again has the opposite sign to the proposed strengthening effect. Source: figure_data/ag_x_sc_h2.csv.

5.4 Portfolio economics and cost attribution

The corrected signal is approximately flat before costs: gross Sharpe is 0.006 and gross CAGR -0.07% . At 5 bps, net Sharpe falls to -2.367 and net CAGR to -9.83% . Arithmetic annualized cost drag is approximately 10.28%. The failure is therefore a weak gross signal destroyed by high turnover and realistic cost assumptions.

Figure 4. Frozen baseline gross versus net wealth



Figure 8. Growth of one currency unit before and after transaction costs. Gross wealth remains close to its starting value, while frequent rebalancing produces a persistent decline after costs. Source: figure_data/equity_curve_gross_vs_net.csv.

Figure 5. Frozen 5-bps baseline drawdown

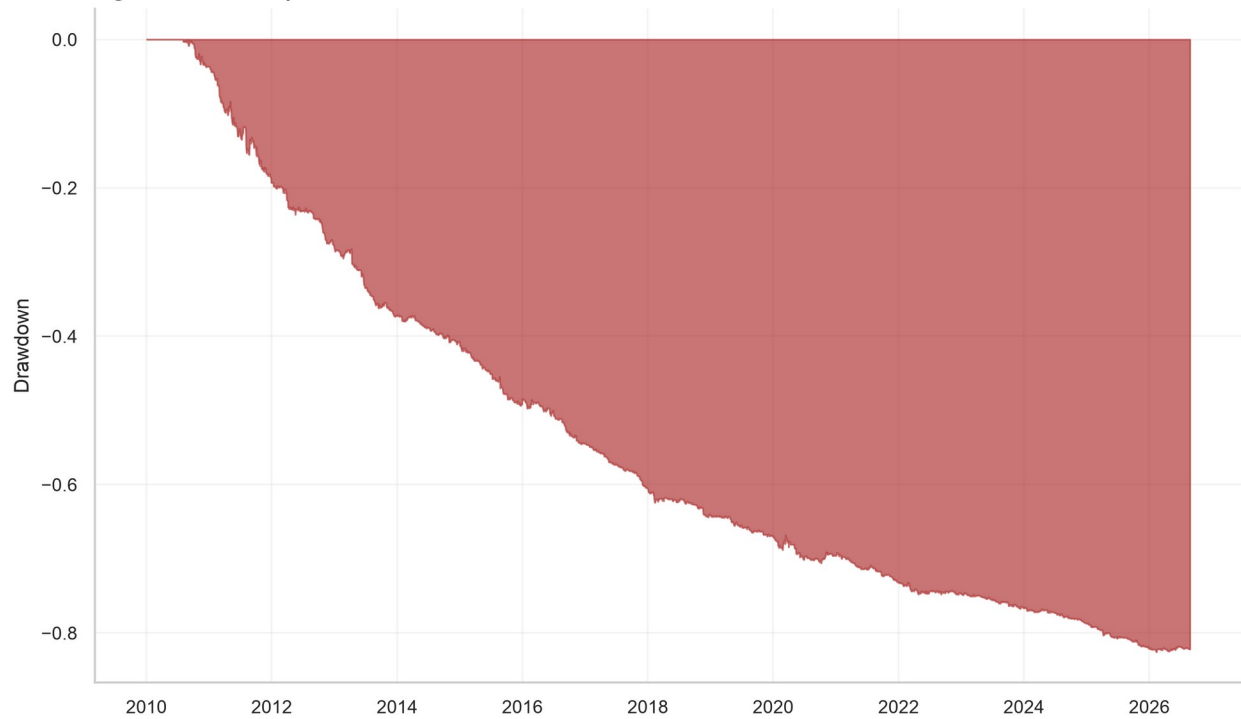


Figure 9. Net drawdown of the corrected baseline. Maximum drawdown reaches -82.56%. Source: figure_data/drawdown.csv.

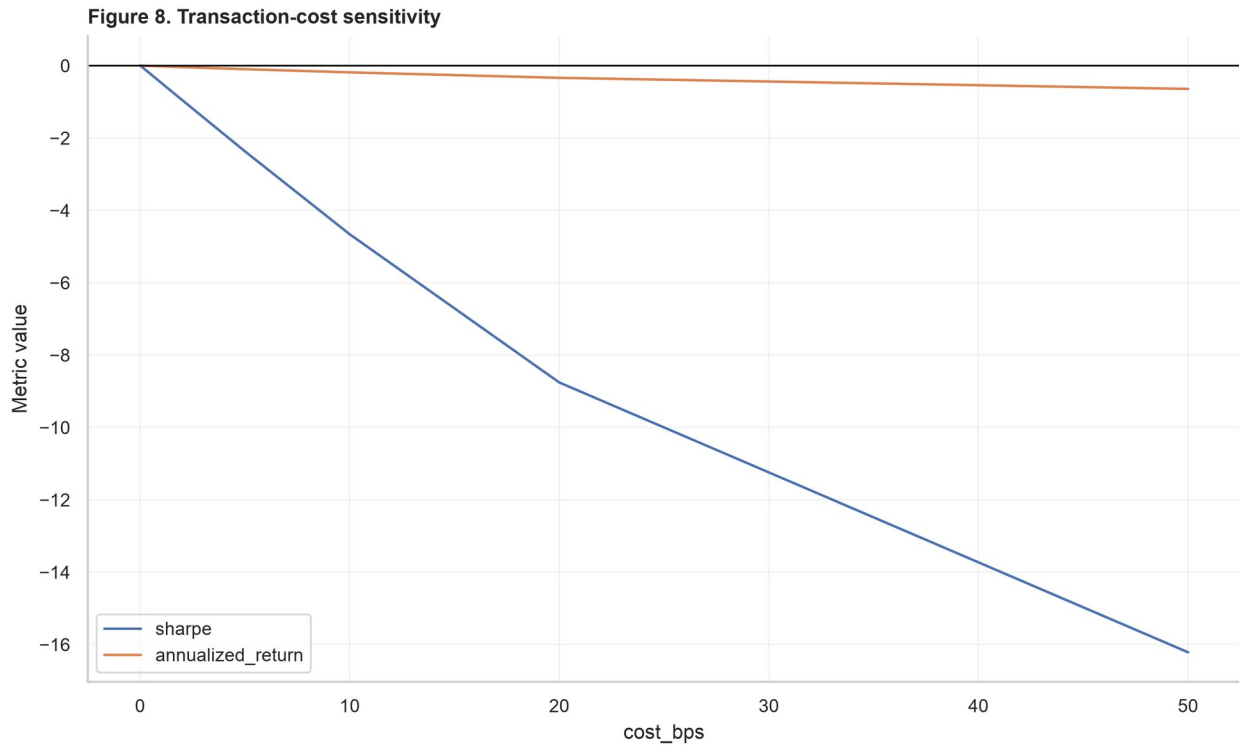


Figure 10. Annualized performance under alternative transaction-cost assumptions. Even small costs overwhelm the approximately flat gross result because average daily turnover exceeds 80%. Source: figure_data/transaction_cost_sensitivity.csv.

Cost (bps)	CAGR	Vol.	Sharpe	Max DD	Avg. turnover
0	-0.07%	4.30%	0.006	-19.4%	81.6%
2	-4.09%	4.31%	-0.948	-53.1%	81.6%
5	-9.83%	4.33%	-2.367	-82.6%	81.6%
10	-18.65%	4.41%	-4.660	-96.8%	81.6%
20	-33.79%	4.69%	-8.757	-99.9%	81.6%
50	-64.36%	6.34%	-16.214	-100.0%	81.6%

Table 1. Portfolio economics across linear transaction-cost assumptions. Turnover is unchanged by the accounting-only cost sweep.

5.5 Long-short and factor attribution

Side attribution is asymmetric. Long-side gross Sharpe is +0.537 with CAGR +2.68%; short-side gross Sharpe is -0.544 with CAGR -2.84%. This does not validate a long-only strategy; it records a future hypothesis. HAC factor regressions align both strategy and proxy returns to the same open-to-open interval. They estimate gross daily alpha of -0.02 bps ($t=-0.06$) and net daily alpha of -4.10 bps ($t=-9.98$). The factor-matrix condition number is approximately 426, so proxy coefficients require caution.

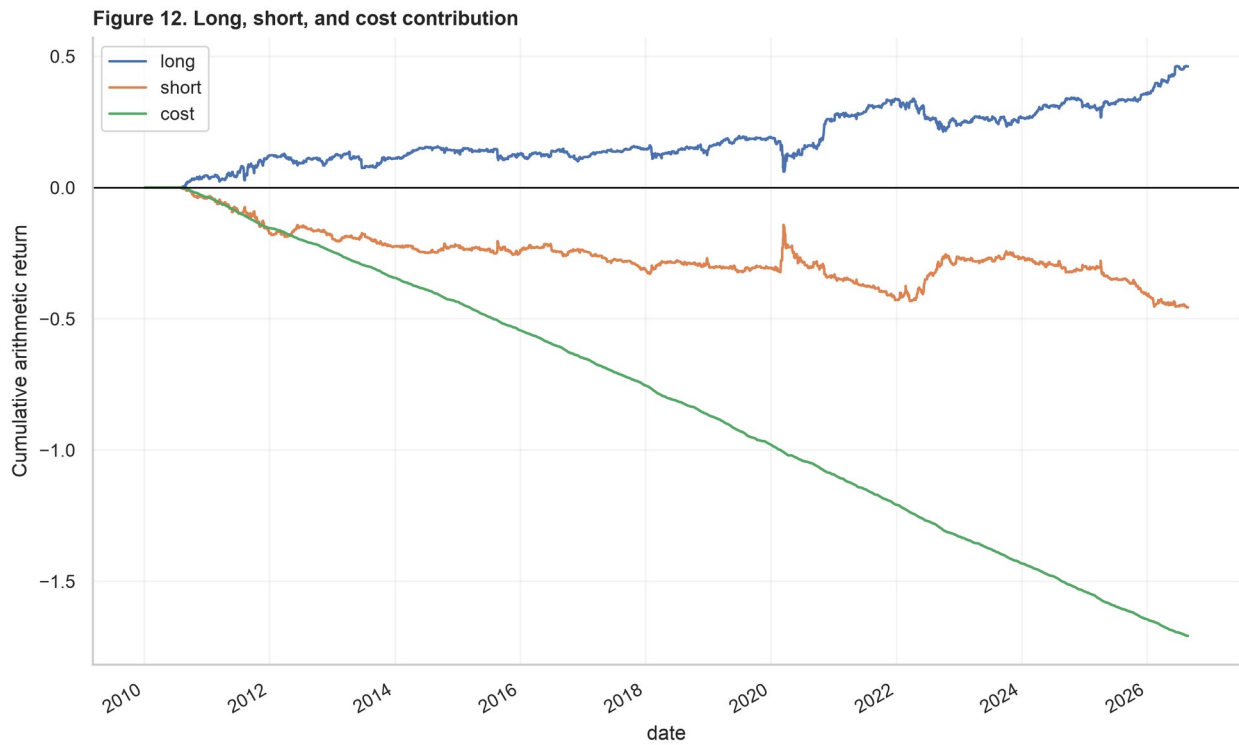


Figure 11. Cumulative contribution of the long and short books. Positive long-book performance is offset by the short book, leaving little gross return before costs. Source: `figure_data/long_short_contribution.csv`.

Portfolio	Alpha (bps/day)	HAC t	p-value	R-sq.	Cond. no.
gross	-0.02	-0.06	0.9535	0.002	426.4
net_5bps	-4.10	-9.98	0.0000	0.002	426.4
long	-0.14	-0.32	0.7483	0.395	426.4
short	0.11	0.26	0.7939	0.374	426.4

Table 2. Factor attribution using broad ETF proxies on the same open-to-open interval as strategy returns. Coefficient-level interpretation is limited by collinearity.

6. Diagnostics, Robustness, and Falsification

6.1 Persistence across forecast horizons

The close-to-close residual Rank IC diagnostic by horizon is -0.00265 , -0.00346 , -0.00262 , 0.00002 , -0.00863 at 1, 2, 3, 5, and 10 days. This is neither a positive propagation profile nor a stable decay signature. Separately, the rolling executable Rank IC changes sign through time, reinforcing the conclusion that the public score does not encode a persistent tradable ranking.

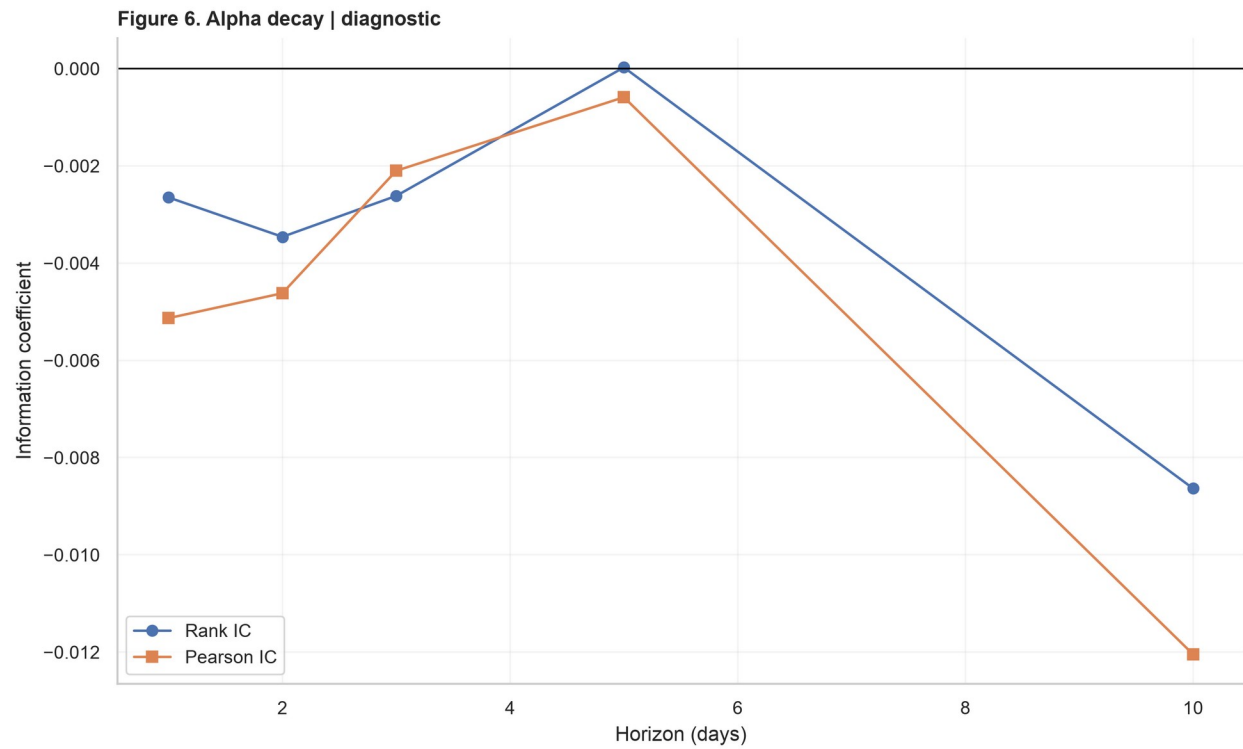


Figure 12. Rank correlation at one-, two-, three-, five-, and ten-day horizons. The relationship does not begin positive and then fade smoothly, as a persistent diffusion signal would. Source: figure_data/alpha_decay_ic.csv.

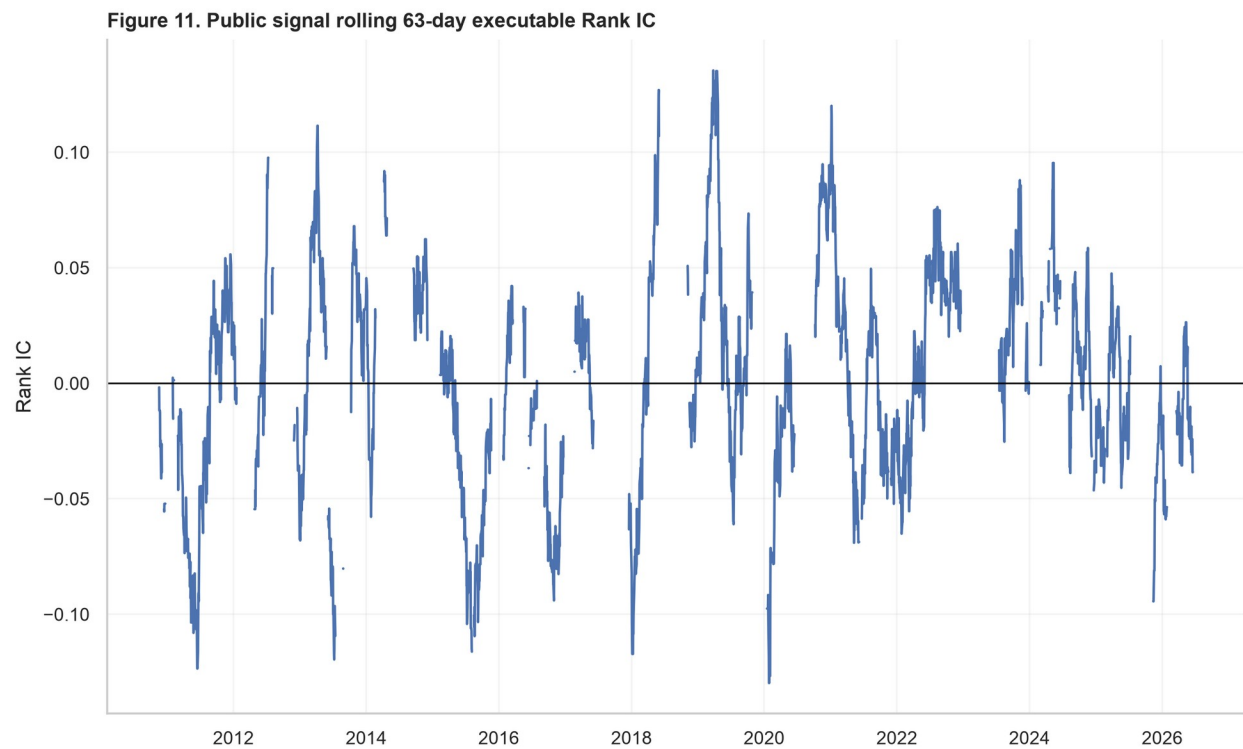


Figure 13. Rolling 63-day executable rank correlation. Frequent sign changes show that the weak average relationship is not stable through time. Source: figure_data/rolling_rank_ic.csv.

6.2 Alternative scaling, component removal, and portfolio breadth

For the close-to-close residual diagnostic, the conclusion is unchanged when AG is scaled by trailing median and MAD, rolling mean and standard deviation, or percentile rank. Their Rank IC estimates are -0.00265 , -0.00334 , and $+0.00067$, respectively. The component-removal test rebuilds the SC state filter after adding each input family in sequence; none of these versions produces positive net economics. The portfolio-breadth test changes the proportions assigned to the long and short tails and reports the resulting Sharpe ratio and turnover. Factor-lookback grids that were not executed are omitted rather than presented as evidence.

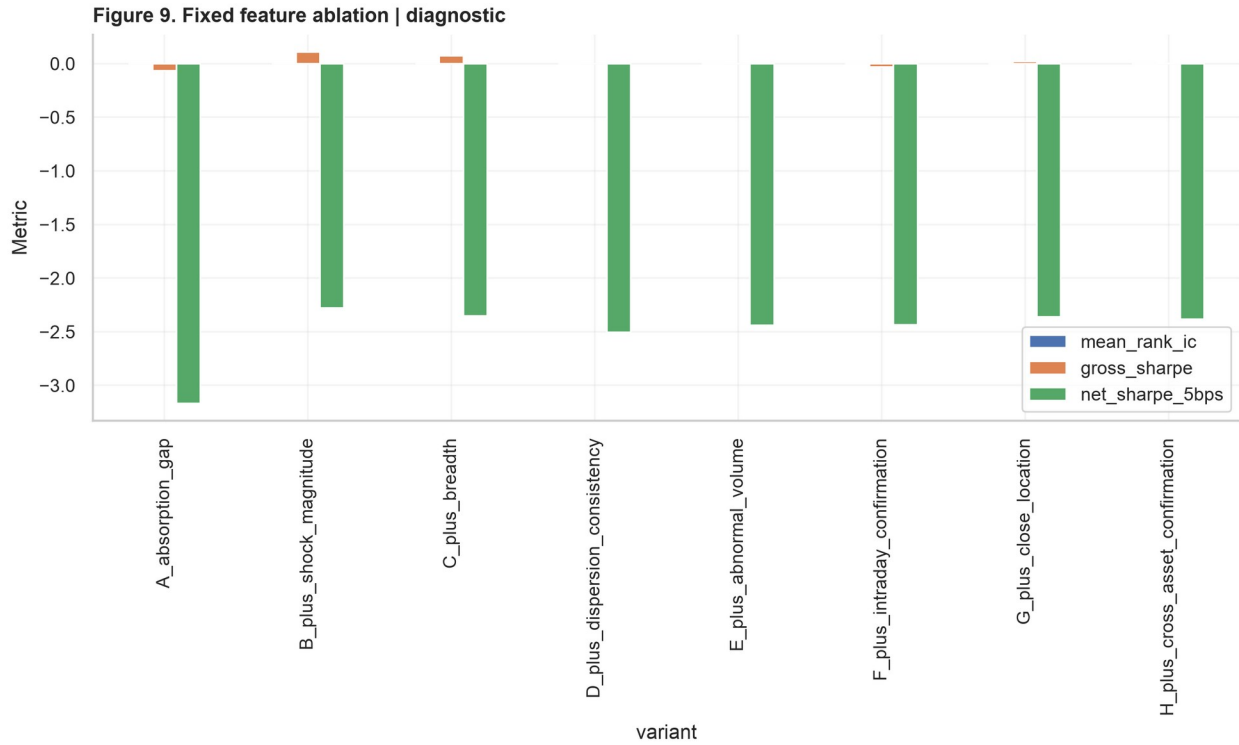


Figure 14. Net and gross performance as SC component families are added sequentially. No partial version of the public state filter produces positive net economics. Source: `figure_data/feature_ablation.csv`.

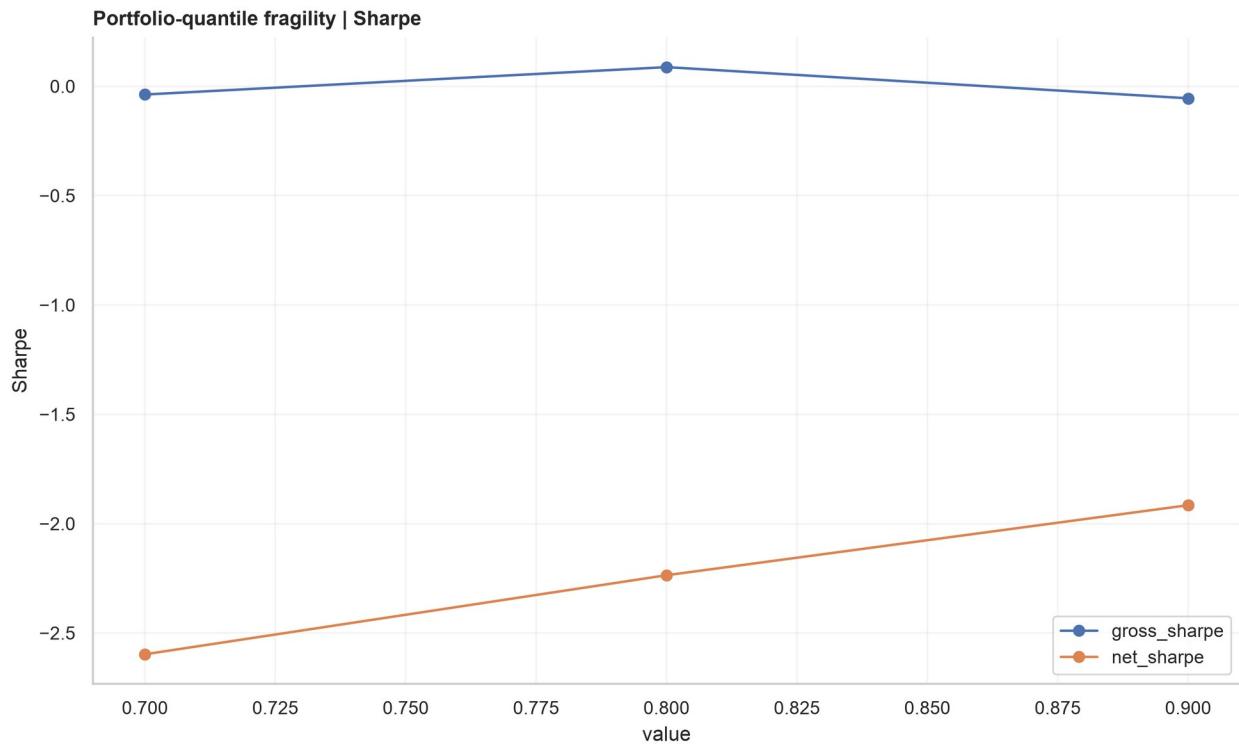


Figure 15. Sharpe ratio as the long and short portfolio tails are widened or narrowed. No tested breadth produces a broad positive net region. Source: figure_data/parameter_stability_sharpe.csv.

6.3 Performance through time

The sample is divided into six sequential evaluation periods, each separated from the preceding estimation period by a one-day embargo. Rank correlation and gross Sharpe change sign from one period to another, indicating that neither ranking quality nor pre-cost performance is stable. Every period has a negative net Sharpe ratio at 5-bps costs. Because the complete history has already been inspected, these segments measure consistency through time rather than untouched out-of-sample performance.

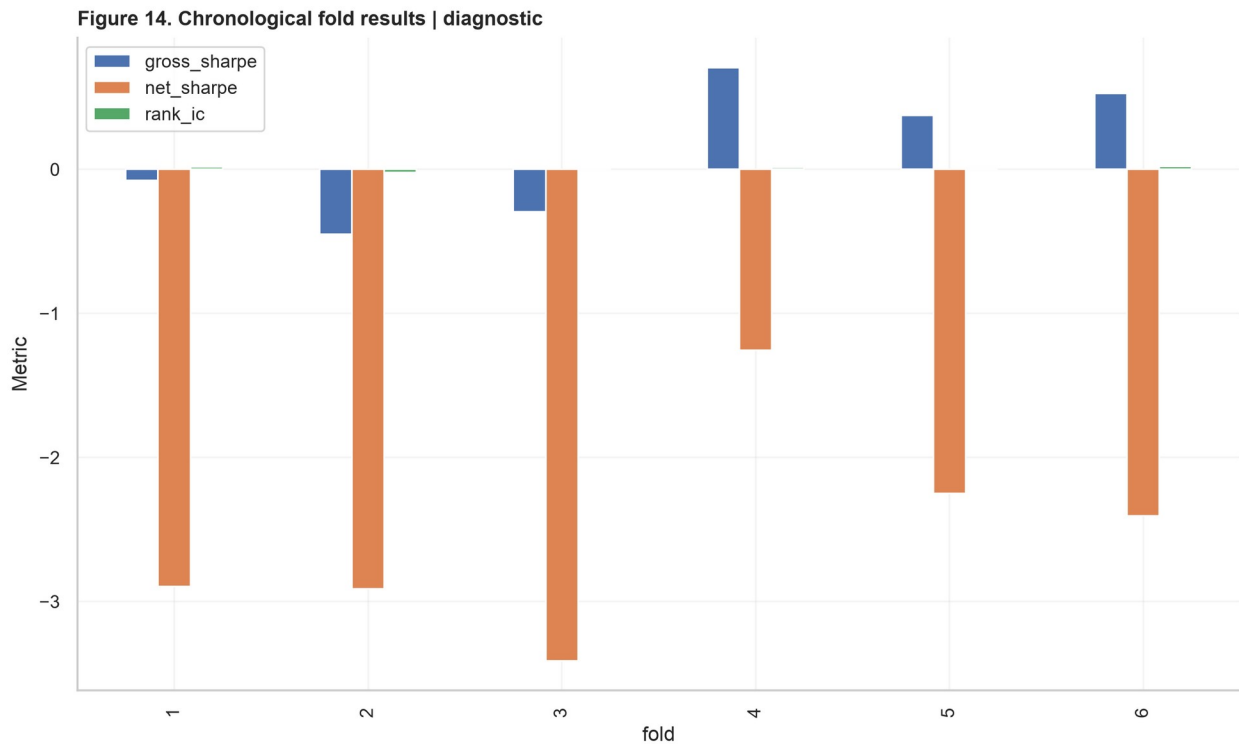


Figure 16. Signal correlation and portfolio Sharpe ratio across six chronological evaluation periods. Gross results change sign, while all six net results are negative. Source: figure_data/walk_forward_results.csv.

Fold	Evaluation window	Rank IC	Gross Sh.	Net Sh.	Max DD
1	2013-01-07 to 2015-01-06	0.0157	-0.08	-2.89	-18.3%
2	2015-01-07 to 2017-01-05	-0.0233	-0.45	-2.91	-22.6%
3	2017-01-06 to 2019-01-08	0.0072	-0.29	-3.41	-21.6%
4	2019-01-09 to 2021-01-07	0.0146	0.71	-1.26	-17.9%
5	2021-01-08 to 2023-01-09	-0.0000	0.37	-2.25	-18.4%
6	2023-01-10 to 2025-01-13	0.0208	0.53	-2.41	-16.9%

Table 3. Chronological out-of-fold diagnostics. These are not described as a genuinely untouched holdout.

6.4 Continuation and reversal after positive and negative shocks

The events are separated by shock direction, whether the asset initially moved less or more than the factor model implied, and whether market coherence was low, medium, or high. Several high-coherence groups subsequently move opposite to the proposed continuation direction, including underreaction after negative shocks and overshooting after positive shocks. The asymmetry may reflect reversal, an omitted downside factor, or error in the expected-response model. Because these groups were examined after the baseline result, they are descriptive rather than trading rules.

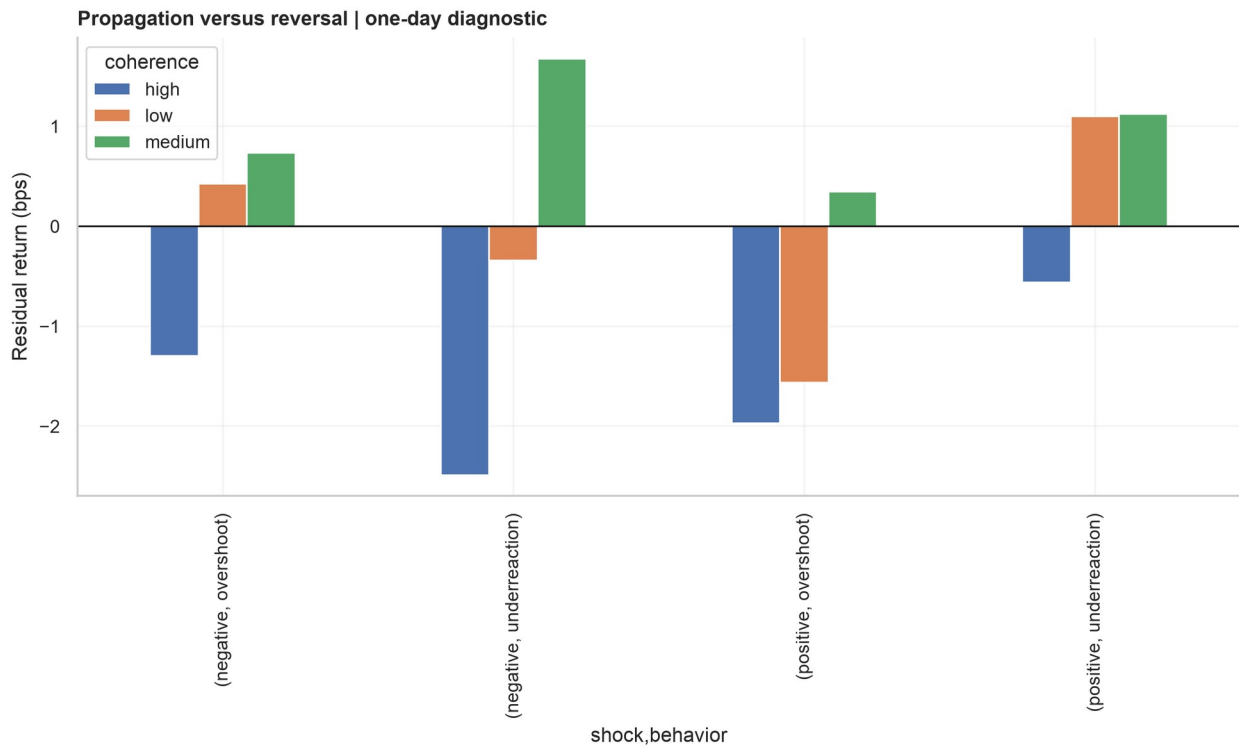


Figure 17. Average next-day residual return after apparent underreaction or overshooting, shown separately for positive and negative shocks and for three market-coherence states. The groups do not support a uniform continuation rule. Source: figure_data/propagation_vs_reversal.csv.

6.5 Machine-learning extension

Chronological logistic regression and gradient boosting produce ROC-AUC values of 0.5008 and 0.4939, where 0.50 represents random classification. Their Brier scores are approximately 0.250, also consistent with uninformative probabilities for a balanced binary outcome. Future residual-return Rank IC is 0.00436 for logistic predictions and 0.00168 for gradient boosting. These statistics are reported without portfolio returns because neither model establishes sufficient predictive information to justify an additional trading specification.

7. Discussion

7.1 What the evidence supports

The evidence supports a narrow but firm conclusion about the disclosed public model. Its executable next-open Rank IC is -0.00020 , the score groups are not monotonic, and the rank relationship changes sign across chronological periods. The within-date randomization test does not distinguish the observed ordering from chance. SC identifies a date-level market state, but because it is common to every asset it cannot enrich the cross-sectional ordering; under the implemented rule it only activates or deactivates AG.

The portfolio result reaches the same conclusion through a different endpoint. The long book contributes positively before costs, the short book offsets much of that gain, and the aggregate zero-cost Sharpe is 0.006. Average daily turnover is 81.6%, so a 5-bps linear cost reduces Sharpe to -2.367 and CAGR to -9.83% . Agreement between the executable rank test, portfolio accounting, chronological splits, and randomized-signal benchmark supports rejection of this public specification, not a claim that the broader diffusion mechanism is false.

7.2 What the evidence does not support

The results do not establish profitable public alpha or prove that gradual information diffusion causes the observed returns. They do not support monotonic public-signal ranks, a stronger AG relationship on high-coherence dates, stable persistence across horizons, or improved prediction from the tested machine-learning models. Adjusted daily ETF data also cannot establish opening-auction execution quality, borrowing availability, market impact, or institutional capacity. Conditional subgroup patterns remain hypotheses until they are fixed in advance and tested on new data.

7.3 Public baseline and private-model boundary

A separate private alpha layer exists outside the public repository and is deliberately withheld to preserve its proprietary value. It was developed after diagnosing the weaknesses of the public baseline and changes the model in three material respects: the ranking signal varies by asset rather than relying on a date-level scalar, factor relationships are tailored more closely to the traded instrument, and portfolio formation incorporates turnover-aware entry, exit, and holding logic. Exact features, coefficients, thresholds, and implementation parameters are not disclosed.

The author's private research records positive net performance for that proprietary implementation after transaction costs and with the same next-open accounting convention used here. That result has not been independently audited and is intentionally excluded from the paper's tables and figures because readers could not reproduce it without the withheld specification and data. Every empirical number reported here therefore belongs to the public baseline. The private result is disclosed only to explain why the public repository contains an extension interface and why the public model should not be interpreted as the complete alpha process.

7.4 Why the baseline may fail

Several mechanisms may explain the public result. Eight broad factor proxies are too coarse and correlated to estimate every instrument's expected response precisely, and the same factor set is applied to all assets rather than tailored by sector or instrument. With only 16 traded ETFs, the long and short tails contain few positions. SC also requires every normalized component, while the final score is not winsorized and the portfolio has no explicit beta or sector constraint. These limitations matter, but portfolio optimization alone cannot repair a ranking whose measured relation with future returns is close to zero.

7.5 Research implications

The public daily specification should not be deployed. Its value is diagnostic: it identifies which elements of the original mechanism survive measurement, where the public ranking fails, how the short book changes the result, and how turnover converts a flat gross return into a large net loss. Further public research should begin from a small number of written hypotheses and evaluate them on later observations or independent point-in-time data, rather than optimize repeatedly on the history already examined.

8. Limitations

Public-data quality

Adjusted Yahoo OHLCV is adequate for a reproducible public study but may contain vendor revisions, stale observations, and imperfect corporate-action adjustments. The download does not reconstruct the information set, index membership, or tradable quotes available to an investor on each historical date. Results may therefore differ when reproduced with institutional point-in-time data.

Daily sampling

Daily bars combine all intraday events into one open, high, low, and close. They cannot determine whether information arrived before or after a price move, and they omit bid-ask spreads, order-book depth, auction imbalances, and the sequence of transactions. The analysis can document association at daily frequency but cannot establish the intraday transmission mechanism.

Execution and capacity

The linear cost deduction is a sensitivity analysis rather than a transaction-level fill model. The study does not observe borrow availability, stock-loan fees, opening-auction participation, market impact, or competition for liquidity. The resulting net returns should be read as a friction screen, not as an institutional capacity estimate.

Factor correlation and model specification

The broad ETF factors are strongly correlated, and the open-to-open attribution matrix has a condition number near 426. Ridge regularization stabilizes the signal-model coefficients, but the residual may still contain omitted common exposures or estimation error. The small 16-asset trading cross-section and absence of explicit beta or sector constraints further limit the interpretation of AG as a pure response delay.

Cross-sectional and temporal dependence

Securities observed on the same date are not independent. The principal tests reduce each cross-section to a daily rank correlation or spread and then use HAC or block-bootstrap inference across dates, but residual dependence can remain. Longer horizons also overlap, reducing the effective number of independent observations.

Multiple testing

Forecast horizons, market regimes, interaction cells, event subgroups, normalization methods, and machine-learning models were examined after the public baseline. They are labeled as diagnostic or exploratory, but labels do not remove selection bias. None of these comparisons should be treated as confirmation without a specification fixed before new data are observed.

Incomplete diagnostics

A complete market-only residual-model backtest and bootstrap confidence intervals for portfolio Sharpe remain unexecuted. Factor-lookback grids are omitted rather than represented as completed robustness evidence. The paper reports these omissions so that the evidentiary boundary is clear.

9. Forward Research Agenda

The following proposals are recorded before any further test is run and are not part of the present evidence. Each requires at least 504 new trading days or an independent point-in-time dataset, unchanged next-open accounting, and transaction-cost assumptions fixed before evaluation. A failed test should remain in the record rather than trigger another round of same-sample parameter search.

Hypothesis	Pre-specified forward test	Falsification rule
Low-coherence reversal	Trade -signed AG only when $SC \leq 0$; one-day next-open holding	Future Rank IC ≤ 0 or 5-bps net spread ≤ 0
Three-day persistence	Trade three-day trailing mean of the public score	Turnover falls $< 25\%$ or future net Sharpe ≤ 0
Sector-specific response	Add a prospectively specified sector proxy to market-only ridge	No residual-variance reduction or monotonicity gain
Intraday confirmation	Interact signed AG with positive intraday-confirmation z-score	Future interaction spread not positive
Quote-aware execution	Preserve signal; use observed spread/auction fills	Realized net spread nonpositive

Table 4. Prospective extensions. These tests are not evidence for the current paper and would require dated preregistration before execution.

10. Conclusion

This study develops and audits a daily cross-asset shock-diffusion pipeline. The economic premise remains plausible: securities exposed to the same information need not adjust at the same speed, and the meaning of an unexplained response may depend on the surrounding market state. The disclosed implementation, however, does not separate that mechanism from factor-model error and ordinary return noise well enough to produce a reliable public ranking signal.

The public signal has a small negative Rank IC, no orderly return pattern across score groups, approximately flat gross portfolio performance, and large losses after transaction costs. Its performance changes sign across time, the high-coherence state does not strengthen the AG spread, and the tested nonlinear models remain close to random classification. The evidence therefore rejects the public baseline as a trading strategy. It does not establish that all versions of the underlying mechanism must fail.

The separate private alpha layer is intentionally not disclosed and is not used to support the public empirical tables. Its existence explains the repository's extension architecture and the distinction between a reproducible research baseline and a proprietary trading process. For the public research program, the next credible step is to register a small set of revised mechanisms before testing them on genuinely new observations, while retaining the same execution timing, cost accounting, and dependence-aware inference.

Declarations

Funding and conflicts

This research received no external funding. The author declares no conflict of interest.

Data and code availability

The public analysis uses historical market data downloaded programmatically from a publicly accessible vendor. Raw vendor files are not redistributed. The reproducibility package contains the public configuration, executed methodology, timing audit, machine-readable result tables, figure-level source data, and reproduction instructions. The proprietary alpha implementation, its parameters, and its private evaluation outputs are deliberately excluded. A repository URL should be added when the public code archive is released; until then, the public package is available from the author on request.

Disclaimer

This paper is for research and educational purposes. Simulated results are not actual trading performance and do not capture every friction faced by a live strategy.

References

- Bailey, D. H., and Lopez de Prado, M. (2014). The Deflated Sharpe Ratio: Correcting for selection bias, backtest overfitting, and non-normality. *Journal of Portfolio Management*, 40(5), 94-107. <https://doi.org/10.3905/jpm.2014.40.5.094>
- Cohen, L., and Frazzini, A. (2008). Economic links and predictable returns. *Journal of Finance*, 63(4), 1977-2011. <https://doi.org/10.1111/j.1540-6261.2008.01379.x>
- Daniel, K., Hirshleifer, D., and Subrahmanyam, A. (1998). Investor psychology and security market under- and overreactions. *Journal of Finance*, 53(6), 1839-1885. <https://doi.org/10.1111/0022-1082.00077>
- Fama, E. F., and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1), 3-56. [https://doi.org/10.1016/0304-405X\(93\)90023-5](https://doi.org/10.1016/0304-405X(93)90023-5)
- Hong, H., and Stein, J. C. (1999). A unified theory of underreaction, momentum trading, and overreaction in asset markets. *Journal of Finance*, 54(6), 2143-2184. <https://doi.org/10.1111/0022-1082.00184>
- Hou, K., and Moskowitz, T. J. (2005). Market frictions, price delay, and the cross-section of expected returns. *Review of Financial Studies*, 18(3), 981-1020. <https://doi.org/10.1093/rfs/hhi023>

- Jegadeesh, N., and Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance*, 48(1), 65-91. <https://doi.org/10.1111/j.1540-6261.1993.tb04702.x>
- Lee, C. M. C., and Swaminathan, B. (2000). Price momentum and trading volume. *Journal of Finance*, 55(5), 2017-2069. <https://doi.org/10.1111/0022-1082.00280>
- Lo, A. W. (2002). The statistics of Sharpe ratios. *Financial Analysts Journal*, 58(4), 36-52. <https://doi.org/10.2469/faj.v58.n4.2453>
- Lo, A. W., and MacKinlay, A. C. (1990). When are contrarian profits due to stock market overreaction? *Review of Financial Studies*, 3(2), 175-205. <https://doi.org/10.1093/rfs/3.2.175>
- Menzly, L., and Ozbas, O. (2010). Market segmentation and cross-predictability of returns. *Journal of Finance*, 65(4), 1555-1580. <https://doi.org/10.1111/j.1540-6261.2010.01578.x>
- Moskowitz, T. J., and Grinblatt, M. (1999). Do industries explain momentum? *Journal of Finance*, 54(4), 1249-1290. <https://doi.org/10.1111/0022-1082.00146>
- Newey, W. K., and West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3), 703-708. <https://doi.org/10.2307/1913610>

Appendix A. Exact Public Baseline

Component	Implemented specification
Universe	24 ETFs; eight proxies excluded; 16 traded assets
Sample	2010-01-04 through 2026-08-28
Factor model	Rolling ridge; lookback 126; minimum 60; penalty 0.001; exposures through t-1
AG	Expected common response minus observed return; trailing median/MAD normalization with historical standard-deviation fallback when MAD is zero
SC	Date-level mean of seven lag-normalized components; all required; clipping [-5,5]
Signal	signed_AG when SC > 0; otherwise zero
Portfolio	Top/bottom 20%; strength weights; gross 1.0; position cap 0.15
Execution	Signal after close t; effective at open t+1; earns open t+1 to open t+2
Cost	Turnover x 5 bps; no nonlinear impact
Neutrality	No explicit beta or sector constraint in baseline constructor

Table A1. Pre-specified public specification. Later diagnostics do not replace it.

Designed-versus-implemented differences

The executed public model applies the same broad factor set to every traded asset rather than using sector- or instrument-specific mappings. Factor proxies are excluded from the trading universe, every SC component must be available, and several SC inputs are aggregated across assets before normalization. The public score is not winsorized, and the portfolio has no explicit beta or sector constraint. These implementation choices determine the results reported in the paper; the undisclosed private layer is not substituted into any public table or figure.

Appendix B. Hypothesis Decisions

Hypothesis	Estimate	Statistical evidence	Economic evidence
H1: Absorption Gap predicts future residual returns	-0.00265	See absorption_gap_ic.csv	See AG deciles
H2: The positive-SC market state conditions the AG relationship	-0.00032	HAC and bootstrap in interaction_spreads.csv	Returns by AG group and SC state
H3: Combined signal has monotonic cross-sectional prediction	-0.00020	Corrected Rank IC is negative and insignificant	Approximately zero gross Sharpe; negative net Sharpe
H4: Propagation and reversal differ economically		Subgroup table	Exploratory subgroup analysis
H5: Signal changes predictably across forecast horizons	-0.00265	Rank correlations and spreads at several horizons	No horizon selected
H6: Predictability survives reasonable implementation costs	-2.36732	Negative corrected net Sharpe	Costs destroy an approximately flat gross result

Table B1. Hypothesis-level conclusions from the generated evidence package. Estimates are rounded to five decimal places.

Appendix C. Reproducibility and Integrity Checklist

Control	Outcome
Legacy reconciliation	Pre-audit metrics are retained as superseded values; corrected metrics intentionally differ
Automated tests	26 passed, including end-to-end next-open timing
Linting	Ruff passed for public source, tests, and scripts
Static typing	Mypy passed for all 60 public source files; private paper-generation utilities are excluded from the public package
Figure provenance	Every publication figure has a matching CSV
Timing audit	Lagged estimators; close-t signal; open-t+1 execution; open-t+1 to open-t+2 return
Research status	Audited baseline, diagnostic, exploratory, and future tests separated
Repository state	No commit invented; git_sha and working-tree status recorded in the manifest
Unexecuted work	Market-only backtest and Sharpe bootstrap disclosed

Table C1. Reproducibility controls. The manifest records the actual repository state rather than claiming a nonexistent revision.

Appendix D. Additional Diagnostics

The appendix retains diagnostics that are useful for specialist review but too detailed for the main narrative. They show feature correlation, the relationship between SC and future volatility, performance across market states, sensitivity of turnover to portfolio breadth, and calendar-year variation. None changes the conclusion for the disclosed public baseline.

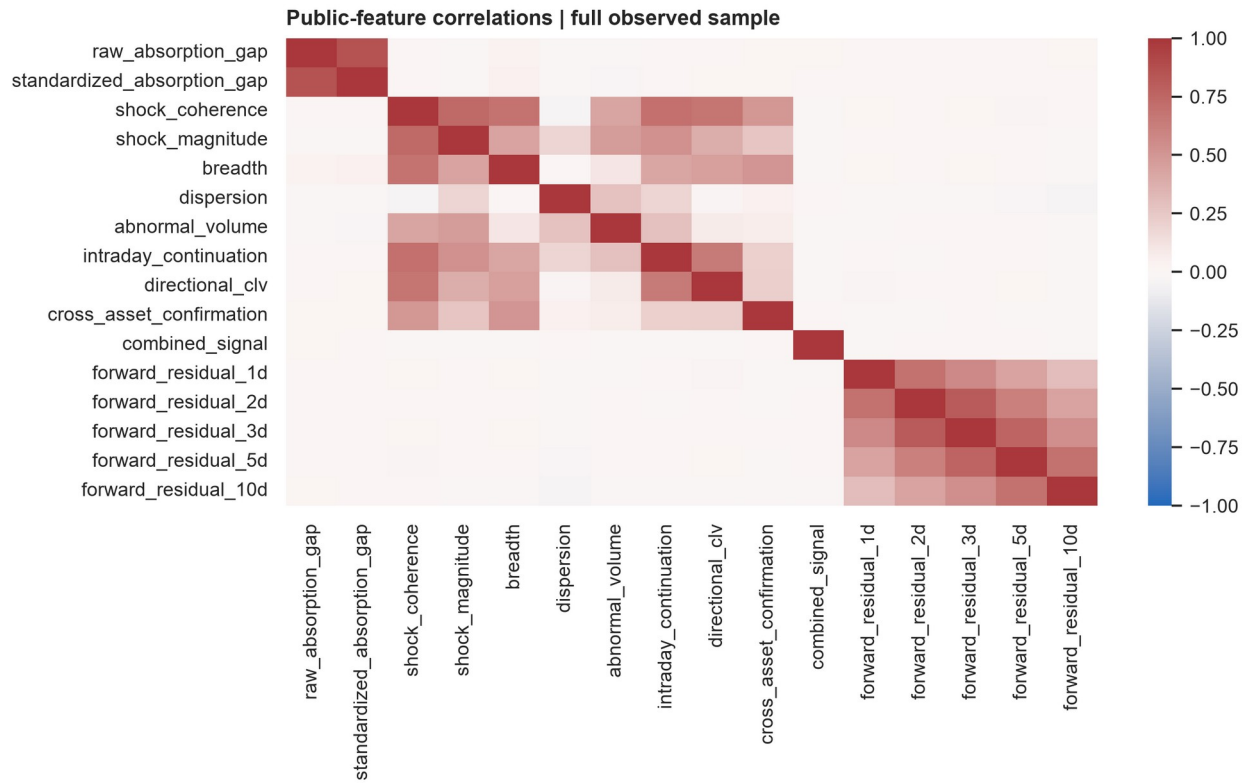


Figure D1. Correlation among public model inputs. Strong relationships among several variables show why individual coefficients and component effects should not be interpreted in isolation. Diagnostic/exploratory; source data are in the matching figure_data CSV.

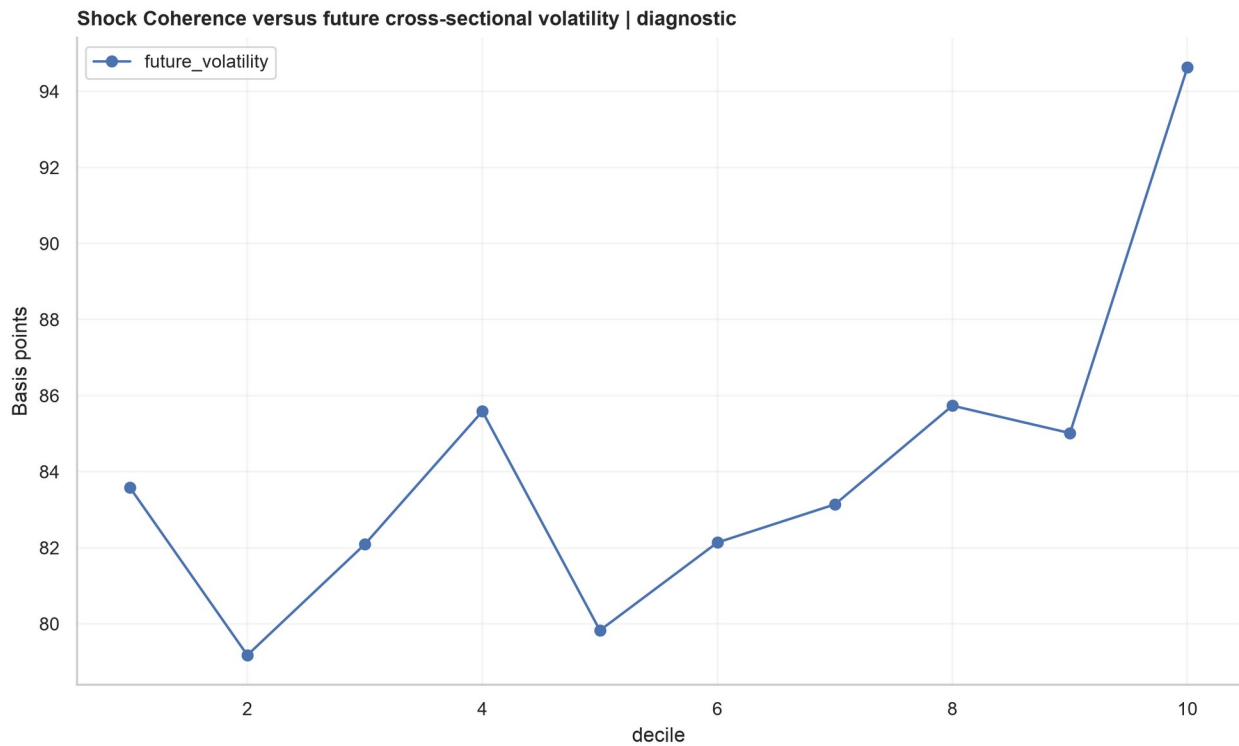


Figure D2. Average future absolute return across SC groups. Higher SC is associated with a different volatility state, not with an asset ranking. Diagnostic/exploratory; source data are in the matching figure_data CSV.

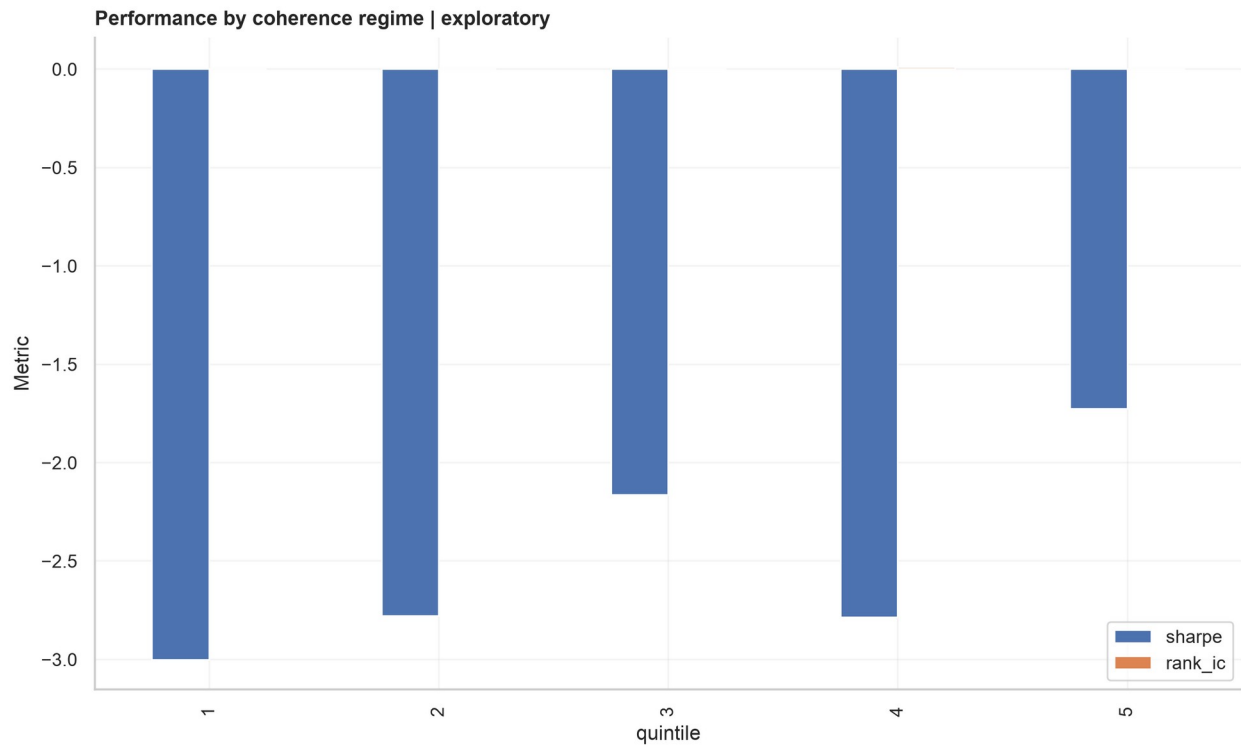


Figure D3. Public portfolio Sharpe ratio and rank correlation across five SC states. Diagnostic/exploratory; source data are in the matching figure_data CSV.

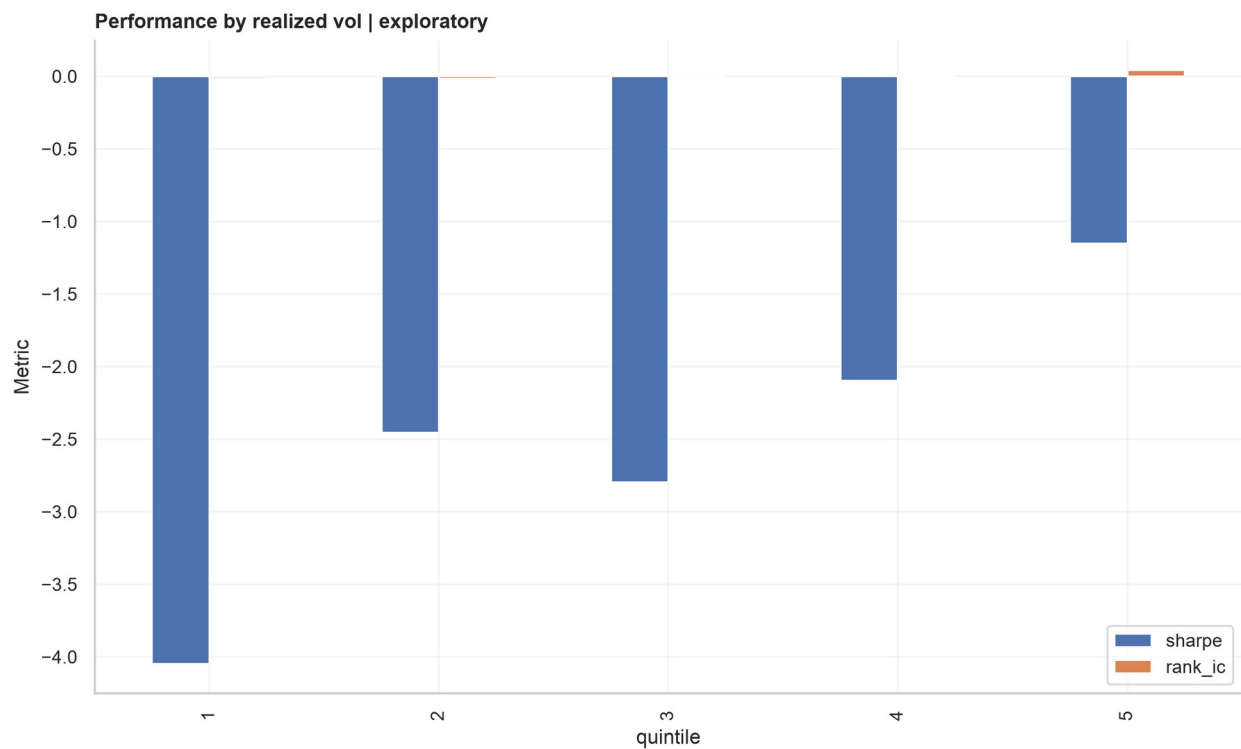


Figure D4. Public portfolio Sharpe ratio and rank correlation across five realized-volatility states. Diagnostic/exploratory; source data are in the matching figure_data CSV.

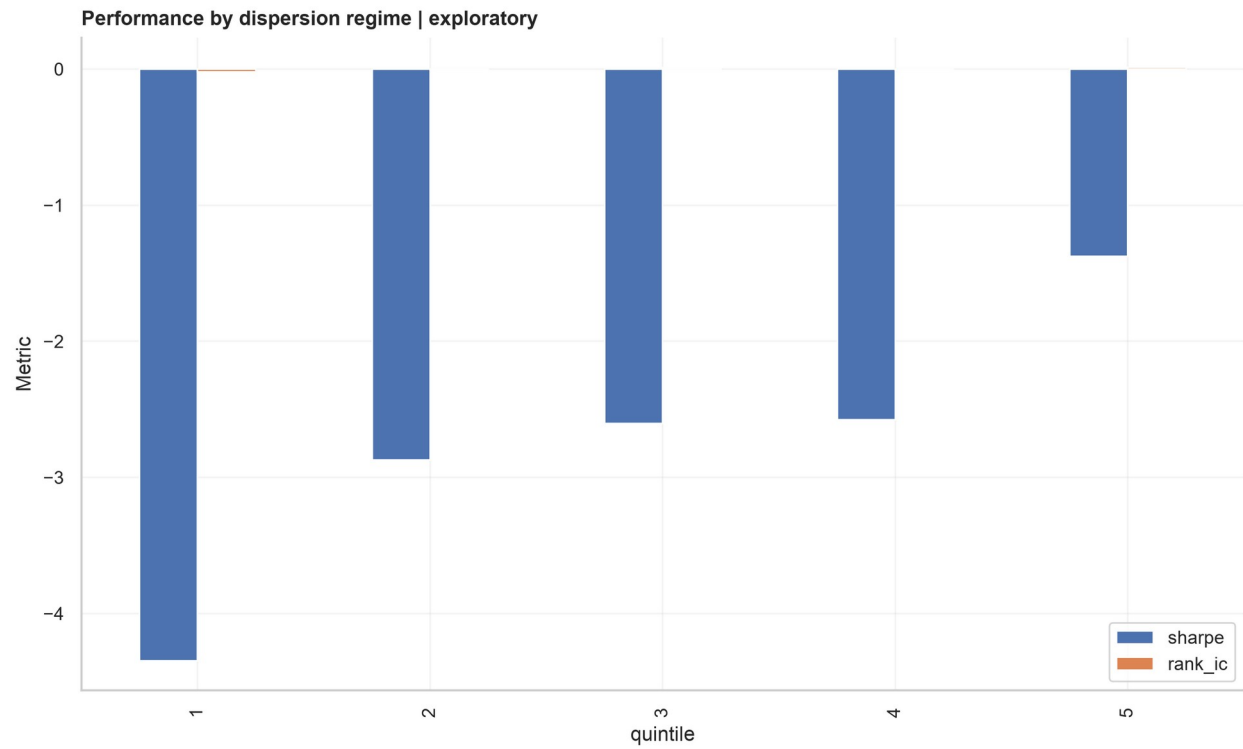


Figure D5. Public portfolio Sharpe ratio and rank correlation across five cross-sectional-dispersion states. Diagnostic/exploratory; source data are in the matching figure_data CSV.

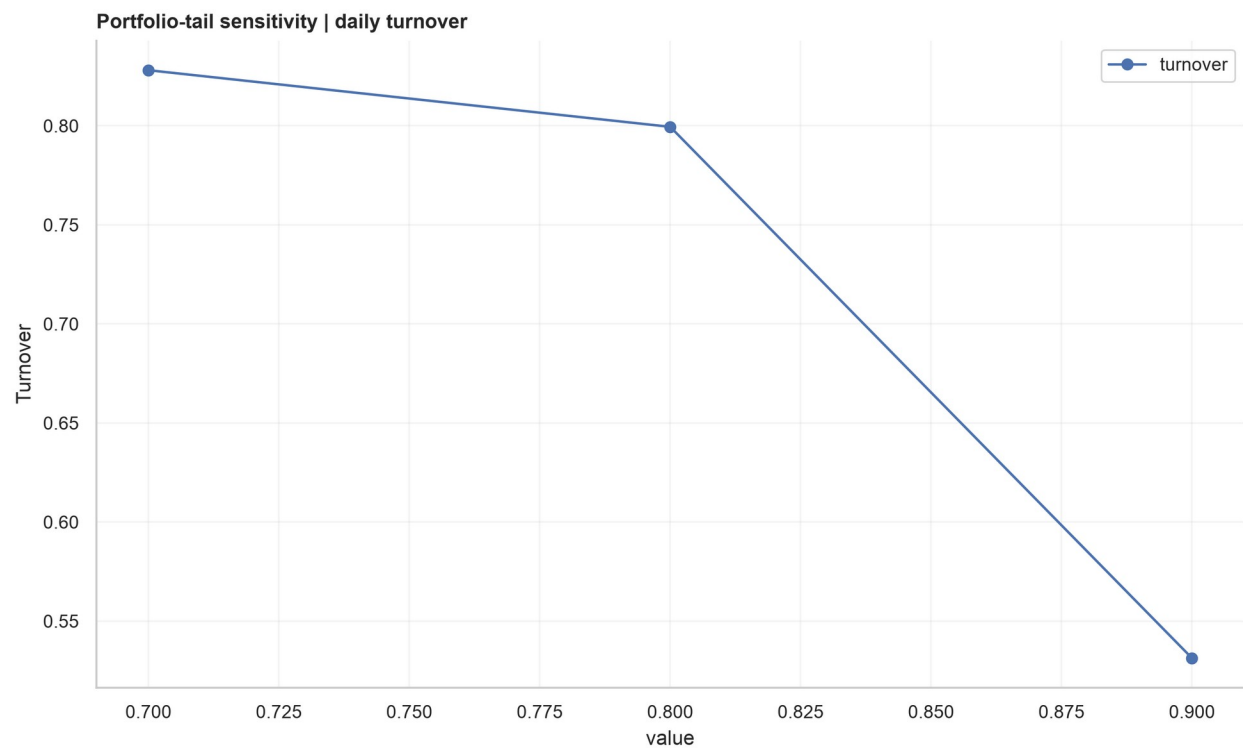


Figure D6. Average daily turnover as the long and short portfolio tails are widened or narrowed. Diagnostic/exploratory; source data are in the matching figure_data CSV.

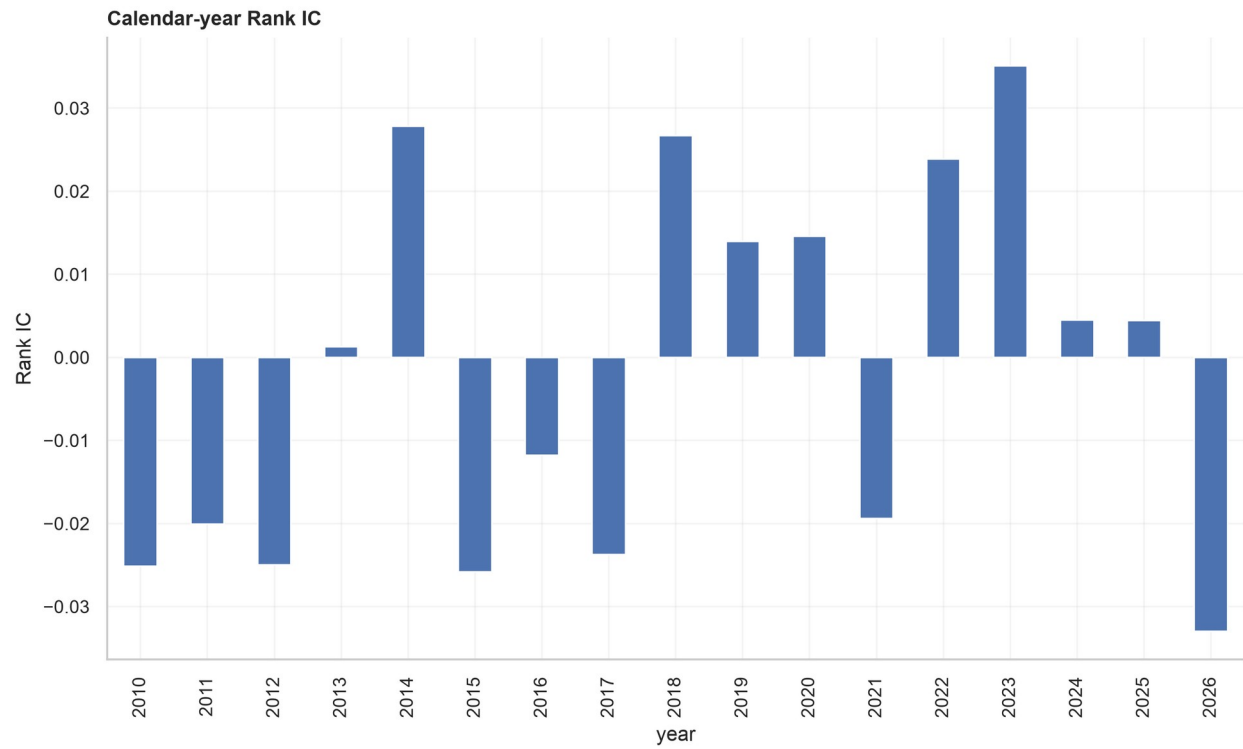


Figure D7. Mean signal-versus-return rank correlation by calendar year. Diagnostic/exploratory; source data are in the matching figure_data CSV.

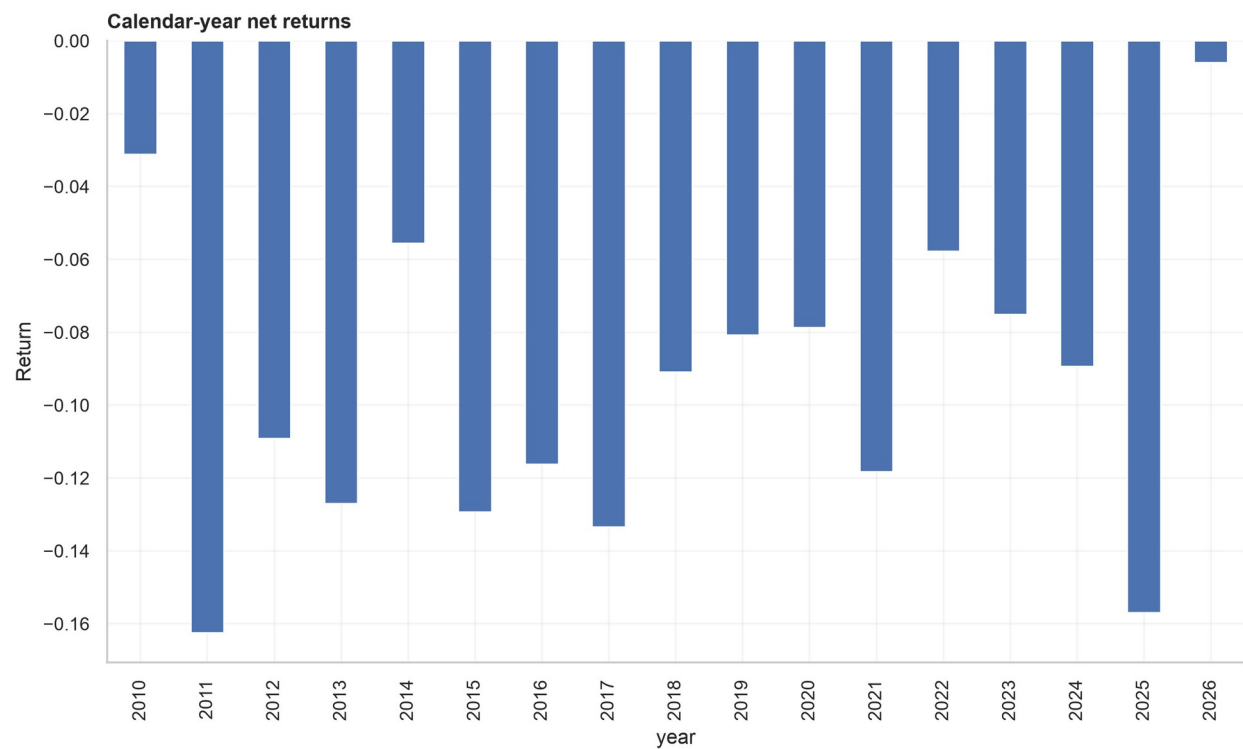


Figure D8. Net public-baseline return by calendar year after 5-bps transaction costs. Diagnostic/exploratory; source data are in the matching figure_data CSV.