Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

When Do Distributional Alarms Improve Downside Risk Control? Evidence from Four Equity Markets

Fuli Yang

dr@garch.ac

GARCH Institute

Tingting Qiu

GARCH Institute

Research Article

Keywords:

Posted Date: November 10th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11661086/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Research manuscript

When Do Distributional Alarms Improve Downside Risk Control? Evidence from Four Equity Markets

Abstract

A distributional alarm changes both when a portfolio cuts equity exposure and how much exposure it retains when only volatility signals stress. We replay allocation rules on the trading calendars of four equity markets and compare net returns at 10 basis points of one-way trading costs. Holding the other state weights fixed, cutting exposure by 25% on distributional-only dates raises mean net returns on common extreme dates by 0.061 percentage points relative to a fixed 22-session realized-volatility rule (95% interval 0.012 to 0.114). The corresponding certainty-equivalent difference is −0.123 percentage points, with an interval that spans zero. Against a volatility rule calibrated to the same historical exposure target, the overlay improves the mean of each policy's own worst days by 0.022 percentage points and trails by 0.072 on common extreme dates. Both intervals span zero. The largest positive contribution to certainty equivalent in that comparison comes from retaining more equity on volatility-only dates. Fixed volatility and flat exposure yield higher certainty-equivalent point estimates across the trading-cost grid. The value of the additional signal therefore depends on the tail dates used to assess protection and on how the exposure budget is allocated across alarm states.

Keywords Distributional alarms, downside risk, exposure timing, transaction costs, certainty equivalent

JEL classification G11, G15, C58

1 Introduction

An equity portfolio can use a distributional alarm to cut exposure when a volatility monitor is quiet. The same alarm also changes how much equity the portfolio can hold when only volatility signals stress. At a common historical exposure target, the two-signal rule can take a smaller cut on those volatility-only dates. We estimate how these two changes contribute to realized downside protection and investor utility after trading costs.

The experiment starts from the monitoring framework and historical sample of the related monitoring study [citation withheld for review]. That study examines which stress episodes the online Wasserstein monitor recognizes, when it responds and how much monitoring it requires. Here we use the same four equity indices, return histories and alarm design to ask a decision question. We construct exposure rules, match their targets using only earlier years and trace their realized return differences to the dates when the two signals agree or disagree. The four index histories, rolling Wasserstein monitor, threshold protocol, event definitions and historical detection counts are shared with the related monitoring study [citation withheld for review]. The severity weights are applied here to those inherited event records. The distinct analysis begins with the target-matched allocation rules and evaluates trading-session execution, net tail outcomes, CE and alarm-state contributions.

Volatility timing studies connect risk estimates to allocation and utility (Fleming, Kirby, and Ostdiek, 2001, 2003; Moreira and Muir, 2017; Nolte and Xu, 2015). The mixed real-time evidence on volatility-managed

portfolios makes the policy comparator consequential (Cederburg et al., 2020). Changes in turnover can alter the utility ranking of allocation rules, including rules that rebalance only after detected changes (Balduzzi and Lynch, 1999; Qiao et al., 2023). Our comparison keeps the calibration rule observable and tracks the difference in exposure across alarm states. The signal itself draws on sequential monitoring (Page, 1954; Ross, 2015) and the geometry of probability distributions (Panaretos and Zemel, 2019). Horvath, Issa, and Muguruza (2024) use Wasserstein distances to cluster financial regimes, while Zeng, Huang, and Chen (2026) study sequential detection in distribution-valued data. Our empirical question concerns how an alarm from this class of methods changes a specific allocation decision.

The analysis proceeds from event coverage and severity (Section 2) to the exposure rules and their calendar replay (Sections 3–4). Section 5 tests whether the results survive changes in benchmark, tail definition and trading cost. Sections 6–8 separate the alarm-state contributions, vary the additional cut and test delayed execution.

2 Signal provenance and historical benchmarks

We use the four index histories and two inherited signals described in Appendix D. DRP compares the distribution of the most recent 22 returns with low- and high-stress prototypes. RV22 uses the root mean square of the same returns. Both signals use pre-2007 thresholds and a five-observation persistence rule.

Across seven historical episodes, DRP detects 21 of 28 market–episode cases and RV22 detects 15. They share 13 detections. RV22 covers all four markets in the COVID episode versus three for DRP. The expanding-reference KS monitor ties DRP at 21 detections with fewer off-window alarms. Appendix D gives the paired lead-time and coverage checks.

Table 1. Historical detection and monitoring workload

Monitor	Detected	Off-window days	Alarm runs
Raw DRP	21/28	3.37%	112
RV22	15/28	3.11%	46
EWMA (0.94)	14/28	3.48%	41
GARCH(1,1)	13/28	3.40%	48
Expanding KS	21/28	2.59%	112

Historical detection counts and alarm workload are inherited from the related monitoring study [citation withheld for review]. Off-window shares are equally weighted market averages. Expanding-reference KS uses the aligned training alarm budget.

DRP generates more alarm runs than RV22 at similar off-window alarm-day shares, so monitoring workload depends on both the number and duration of alerts.

Table 2. Share of realized event severity covered

Weight	DRP %	RV22 %	Gap pp	Leave-one-episode-out gap pp
Maximum drawdown, simple returns	84.19	75.02	9.17	3.88–17.00
Worst five observations, simple returns	82.25	70.46	11.78	6.49–20.43
Cumulative negative log returns	88.43	76.02	12.41	6.90–18.73
Maximum drawdown, historical	84.56	75.48	9.07	3.85–16.61
Worst five observations, historical	82.40	70.77	11.63	6.40–20.29

Severity coverage is a new weighting of the inherited event and detection records. Weights use the full event window. The first two rows compound simple returns. The cumulative measure sums the magnitudes of negative log returns. The final two rows reproduce the archived convention.

DRP detects cases accounting for 84.19% of the drawdown severity and 82.25% of the worst five-observation loss severity, compared with 75.02% and 70.46% for RV22. The comparison also favors DRP when weighting by cumulative negative log returns, and each severity gap remains positive after omitting one episode at a time. A case can count toward coverage even when the alarm occurs after its deepest loss. The allocation test therefore uses subsequent returns and explicit policy timing.

3 Exposure rules and evaluation design

The combined rule starts from full equity exposure. On an RV22 alarm it holds 50% equity, including dates when both signals alarm. It holds 75% when only DRP alarms and 100% when neither does. We compare it with a fixed 50% RV22 cut, an RV22 cut capped at full exit and calibrated to prior alarm frequencies, an RV22 rule with an exactly matched prior-year exposure target, and a flat exposure rule set from the same training history.

For the exact match, let p be the prior frequency of RV22 alarms, and let T be the overlay's mean historical cut, equal to half the frequency of RV22 alarms plus a quarter of the frequency of DRP-only alarms. Each market and return year uses an expanding training sample from 2007 through the previous year. The RV22 control cuts min(1, T/p) on alarm dates, with a full-cut fallback if p is zero. If the cap binds, the control applies a constant cut of (T − p)/(1 − p) on non-alarm dates. The flat control applies T every day. Thus the historical targets match by construction. The realized mean exposures can differ as the alarm frequencies change.

Our main policy comparison reconstructs both alarms on archived exchange-session grids for London, Hong Kong, Tokyo and New York, including two Hong Kong closure corrections. We remove non-session observations, retain genuine zero returns and mask returns that may span a gap or anomalous price. Scores require 22 consecutive valid returns, and missing scores break alarm streaks. We reproduce the archived session-aware DRP series and recalibrate both thresholds in the pre-2007 history. We match each eligible signal at session t to the index return dated at the immediately following scheduled session. The delay sensitivity shifts the entire target path by one additional scheduled session for both policies, before omitting unscored returns.

Missing scores leave positions unchanged. Costs on masked return dates are charged on the next scored date. The replay starts from full equity exposure, with the first evaluation row in each market used for position initialization and omitted from scoring. Of 13,397 scheduled rows from 2013 through 24 June 2026, 11 returns are masked and four initialization rows are set aside. Both policies and both timing rules therefore share 13,382 scored returns. We also retain the frozen observation replay and rescore its existing policy paths on the exchange sample dates. The frozen sequence contains 313 evaluation return dates flagged as non-sessions in the calendar audit. A one-observation lag on that grid can differ from one exchange session.

We convert log index returns to simple returns before applying the equity weight and charge 10 basis points per unit of one-way weight change. The residual earns zero. We compute annualized mean, volatility and mean–variance certainty equivalent (CE) within each market at 252 observations per year and average the market results equally. For risk aversion γ, CE is 252 times the daily mean net return minus $252\gamma/2$ times the daily sample variance. Tail performance is the average of net returns at or below the empirical 2.5% quantile. We calculate maximum drawdown on each compounded net return path. Uncertainty uses 4,000 paired calendar-year resamples, sharing the year draws across markets and retaining the calibrated paths.

The base replay associates a signal computed with end-of-session data with the next close-to-close index return. Realizing the full interval would require execution at its starting close. The one-session delay measures sensitivity to shifting that interval. The available close series does not identify a next-open fill.

For each market and policy, let w_t denote the weight applied to the simple index return r_t, τ_t the charged one-way turnover, and c the cost per unit of turnover (0.001 at 10 bp). Let μ and s^2 denote the sample mean and sample variance of net daily returns. CE is expressed in percent per year and averaged across markets after calculation.

$$r_t^{net} = w_t r_t - c\tau_t$$

$$CE_\gamma = 100 \times 252 \times (\mu - \frac{\gamma}{2} s^2)$$

4 Downside risk under matched exposure targets

Table 3. Overlay minus equal-target RV22 by observation grid

Grid and timing	Rows	Mean return	Volatility	Tail mean	Max. DD	CE γ = 5	95% CE interval
Frozen grid, next observation	13,703	+0.094	−0.103	+0.028	−1.286	+0.181	[−0.172, +0.677]
Frozen grid, extra observation lag	13,703	+0.173	−0.091	+0.024	−1.282	+0.250	[−0.138, +0.753]
Frozen grid, common dates	13,382	+0.097	−0.104	+0.029	−1.290	+0.185	[−0.173, +0.689]
Frozen grid, common dates and lagged	13,382	+0.177	−0.091	+0.025	−1.281	+0.255	[−0.139, +0.763]
Exchange grid, next session	13,382	+0.071	−0.081	+0.022	−1.303	+0.147	[−0.254, +0.670]
Exchange grid, extra session lag	13,382	+0.226	−0.070	+0.022	−1.025	+0.290	[−0.123, +0.809]

Max. DD denotes maximum drawdown. Results are equal-market averages. Differences and intervals are in percentage points. Common-date rows score the original frozen policy paths on the exchange sample dates. The delayed rows shift both paths by one unit on the corresponding grid.

With the next-session convention, the overlay has lower volatility, a better lower-tail mean and a smaller drawdown than the equal-target RV22 rule on the exchange calendar. The respective differences are −0.081, +0.022 and −1.303 percentage points. The CE gap at γ = 5 is +0.147 percentage points, with a 95% interval of [−0.254, +0.670]. The frozen-grid gap is +0.181 percentage points, and it is +0.185 on the same return dates. The difference between the two replays therefore persists after aligning the scoring dates. The signals and positions were reconstructed on their respective grids.

The next-session equity exposure of the overlay averages 0.283 percentage points less than the matched RV22 control. Its annualized net mean return is 0.071 percentage points higher. Without the S&P 500 the CE gap falls to +0.018 percentage points, with a 95% interval of [−0.228, +0.276]. Requiring available signals at both lags leaves 13,249 observations and a CE gap of +0.147 percentage points. The next-session overlay trails fixed RV22 by 0.123 percentage points and the flat control by 1.050 percentage points in CE at γ = 5. The historical matching exercise evaluates the timing and extent of exposure cuts against a volatility-triggered rule. The flat policy gives a different return-risk trade-off.

5 Same-sample comparisons and trading costs

The equal-target control addresses a specific question about exposure timing. Fixed RV22 and flat exposure provide separate comparisons with simpler decisions. Tables 4 and 5 put all five policies on the same 13,382 rows. They also introduce a shared-date tail measure. Within each market we select the dates at or below the 2.5% quantile of the unhedged index return and apply that identical date set to each policy's net returns.

Table 4. Policy performance on the same exchange-calendar sample

Policy	Ann. return	Ann. vol.	Own tail	Shared tail	Max. DD	CE	Turnover
RV22 + DRP	7.572	17.340	−3.196	−3.136	38.218	−0.148	0.909
Equal-target RV22	7.500	17.421	−3.218	−3.064	39.521	−0.294	0.563
Fixed RV22	7.962	17.658	−3.249	−3.197	38.557	−0.024	0.438
Flat exposure	8.810	17.608	−3.244	−3.244	37.903	0.902	0.004
Capped RV22	7.505	17.440	−3.222	−3.068	39.581	−0.304	0.563

All performance measures except turnover are in percent. CE is an annualized level at $\gamma = 5$, calculated from each policy's net mean and variance without subtracting a benchmark. Appendix C uses the same CE definition on a different observation grid. Negative levels occur when the variance penalty exceeds the mean. Turnover is annual one-way notional per unit of portfolio value. Own tails use each policy's worst 2.5% of net returns. Shared tails use the dates of the unhedged index's worst 2.5% of simple returns within each market. The sample, timing, and one-way cost of 10 bp are identical across rows.

Table 5. Paired intervals for overlay minus each control

Control	CE gap	95% CE interval	Own tail	95% interval	Shared tail	95% interval
Equal-target RV22	+0.147	[−0.254, +0.670]	+0.022	[−0.059, +0.097]	−0.072	[−0.314, +0.081]
Fixed RV22	−0.123	[−0.366, +0.120]	+0.053	[+0.008, +0.103]	+0.061	[+0.012, +0.114]
Flat exposure	−1.050	[−1.744, −0.386]	+0.048	[−0.102, +0.213]	+0.107	[−0.072, +0.329]
Capped RV22	+0.156	[−0.241, +0.678]	+0.026	[−0.054, +0.099]	−0.069	[−0.309, +0.084]

Differences and intervals are in percentage points. We resample the 14 calendar-year blocks, including the partial 2026 block, 4,000 times with shared draws across markets and policies. In each draw we recompute both the own-policy and the unhedged-index quantiles, keeping the fitted policy paths fixed.

Against equal-target RV22, the policy-specific tail gap is +0.022 percentage points with a 95% interval of [−0.059, +0.097]. On the shared extreme dates the gap is −0.072, with an interval of [−0.314, +0.081]. The same bootstrap gives CE differences of −0.123 percentage points against fixed RV22 and −1.050 percentage points against flat exposure. The matched-target tail estimates change sign when the date set is held fixed, and both intervals span zero. Against flat exposure the CE interval lies entirely below zero.

The own-tail measure describes the severity of each strategy's worst outcomes and is related to expected shortfall expressed on the return scale (Acerbi and Tasche, 2002). Since the strategies can change which days enter their worst 2.5%, its gap reflects both the payoffs and the composition of the tail. The shared-tail measure holds the market dates fixed and asks which rule fares better during the same index selloffs. The sign reversal against equal-target RV22 shows why a less severe own tail does not establish stronger protection in common selloffs. Both intervals cross zero, and the sign reversal alone cannot quantify how much of the own-tail gap comes from date composition. Against fixed RV22, both tail definitions favor the overlay and both intervals exclude zero.

With the fitted paths held fixed, the cost grid tests how much the ordering depends on the assumed execution cost. The overlay turns over 0.909 times its value annually versus 0.563 for equal-target RV22. The matched-target CE gap declines from +0.181 percentage points at zero cost to +0.078 percentage points at 30 bp. Fixed RV22 and flat exposure have higher CE point estimates throughout the cost grid.

Table 6. CE differences under alternative one-way costs

Cost bp	Equal-target RV22	Fixed RV22	Flat exposure	Capped RV22
0	+0.181	−0.077	−0.960	+0.190
10	+0.147	−0.123	−1.050	+0.156
20	+0.112	−0.170	−1.140	+0.122
30	+0.078	−0.217	−1.231	+0.088

Overlay minus comparator at γ = 5 in annualized percentage points. Cost changes are applied to the same realized weight paths and turnover. Exposure parameters stay fixed.

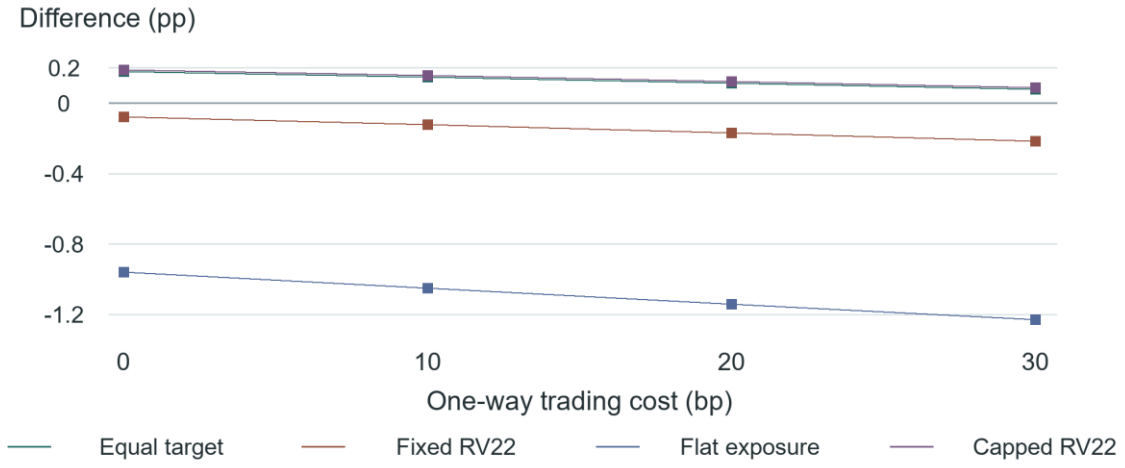


Figure 1. Annualized CE gaps, overlay minus each comparator, across the one-way cost grid. The policy paths are held fixed. See Table 6.

6 Which alarm states account for the gains

We group scored returns by the alarm state used to set the position. Each state's net return contribution is its sum of daily policy return differences divided by the full market sample size, then annualized. For the lower-tail mean, each policy first selects its own worst returns. State contributions allocate that policy's tail sum across states using its total tail count as the denominator. For CE, we add the state's net return contribution and its share of the difference in squared deviations from the two policies' full-sample means. The state contributions therefore sum to the corresponding full-period differences.

Table 7. Alarm-state contributions on the exchange calendar

Alarm state	Rows	Tail mean pp	Net P&L pp	CE γ = 5 pp
DRP only	530	+0.127	−0.362	−0.096
RV22 only	417	−0.099	+0.384	+0.214
Both	47	+0.000	+0.069	+0.060
Neither	12260	−0.005	−0.018	−0.031
No score	128	−0.000	−0.001	−0.001

Next-session policies at a 10 bp one-way cost. Counts sum across markets. All contributions are equally weighted across markets. A missing score retains the previous position.

The 530 DRP-only observations contribute +0.127 percentage points to the lower-tail mean gap. Their annualized net return contribution is −0.362 percentage points and their CE contribution is −0.096 percentage points. The overlay protects against some losses when RV22 is quiet and also gives up returns on dates when that additional cut proves costly. The lower-tail improvement thus has a different source from the positive total CE estimate.

The 417 RV22-only observations contribute +0.384 percentage points to annualized net return and +0.214 percentage points to CE. On those dates the combined rule generally holds more equity than the matched RV22 control, which has to concentrate more of its historical cut on volatility alarms. This state contributes −0.099 percentage points to the lower-tail mean difference. The accounting identifies where each result arises under the specified rules. The overall CE estimate also depends on market composition and the control used.

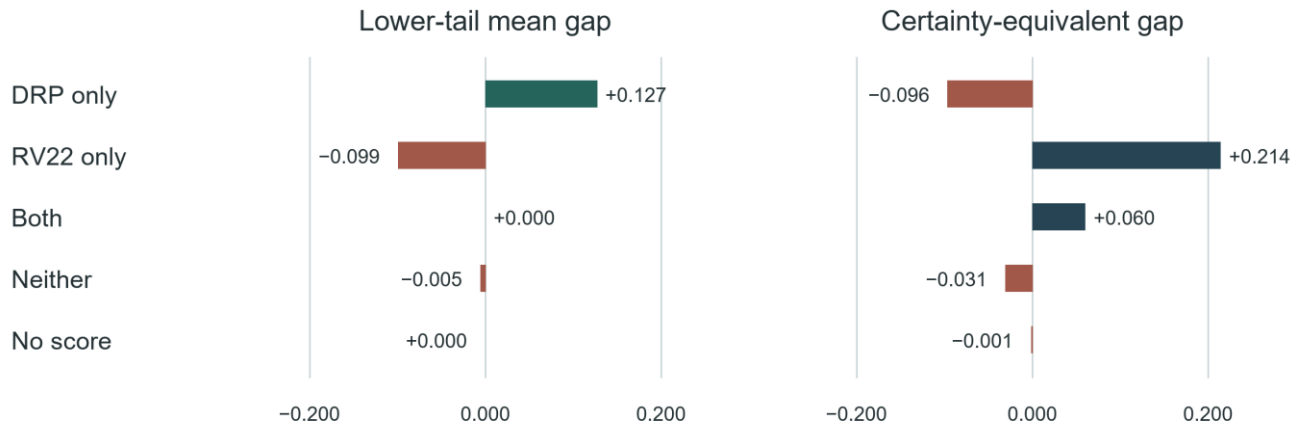


Figure 2. Contributions to the overlay-minus-equal-target RV22 gaps, in percentage points, by alarm state. The exchange-calendar next-session sample and 10 bp cost are the same in both panels. The tail panel uses each policy's own tail dates. CE uses $\gamma = 5$. Values sum to the overall differences, subject to rounding.

A separate decomposition of frozen-grid gross return attributes −0.024 percentage points to the mean exposure gap and +0.149 percentage points to the covariance of relative exposure with subsequent returns. Incremental costs absorb 0.031 percentage points. The covariance describes timing across all states, including relative exposure on RV22-only dates. Appendix B retains the frozen-grid state contributions and their turnover.

7 Isolating the additional DRP-only cut

The state accounting compares complete policies. To isolate the extra cut within the overlay, we keep the 50% equity cut on RV22 alarm dates and full exposure on no-alarm dates. Only the cut triggered by DRP alone varies. We replay each choice on the full exchange grid with the same implementation rules and recompute costs from its own path. At zero additional cut the series matches fixed RV22 exactly. At 25% it matches the main overlay exactly.

Table 8. Varying the DRP-only cut with other target weights fixed

DRP-only cut	Ann. return	Own tail	95% interval	Shared tail	95% interval	CE gap	95% CE interval
12.5%	−0.195	+0.029	[+0.004, +0.056]	+0.030	[+0.006, +0.057]	−0.052	[−0.176, +0.073]
25%	−0.390	+0.053	[+0.008, +0.103]	+0.061	[+0.012, +0.114]	−0.123	[−0.366, +0.120]
37.5%	−0.585	+0.071	[+0.010, +0.139]	+0.091	[+0.018, +0.171]	−0.214	[−0.576, +0.144]
50%	−0.780	+0.080	[+0.010, +0.158]	+0.121	[+0.024, +0.228]	−0.323	[−0.799, +0.140]

Each row minus the zero-cut arm. All differences are in percentage points. RV22-alarm and no-alarm target cuts are fixed at 50% and zero. Each arm replays its own realized turnover at 10 bp and holds its last weight when a score is unavailable. These arms vary average equity exposure as well as DRP-only decisions. This is a sensitivity grid.

Switching from zero to the 25% DRP-only cut changes the policy-specific tail mean by +0.053 percentage points and the shared-date tail mean by +0.061 percentage points. The latter interval is [+0.012, +0.114]. Annualized net return changes by −0.390 percentage points and CE by −0.123 percentage points, with a CE interval of [−0.366, +0.120]. Across the four positive cuts, both tail point estimates improve as the cut deepens, while annualized net return and CE point estimates decline. The comparison identifies the consequences of changing the rule within this replay. The 25% cut lowers realized average equity exposure from 98.249% to 97.262%. Sections 3–4 match the historical exposure target to compare timing at the same ex ante exposure budget. Realized exposure can still differ as alarm frequencies change. The dose experiment leaves the exposure reduction in the measured effect, so its tail improvement includes the benefit of holding less equity and cannot by itself measure the information value at a common exposure budget.

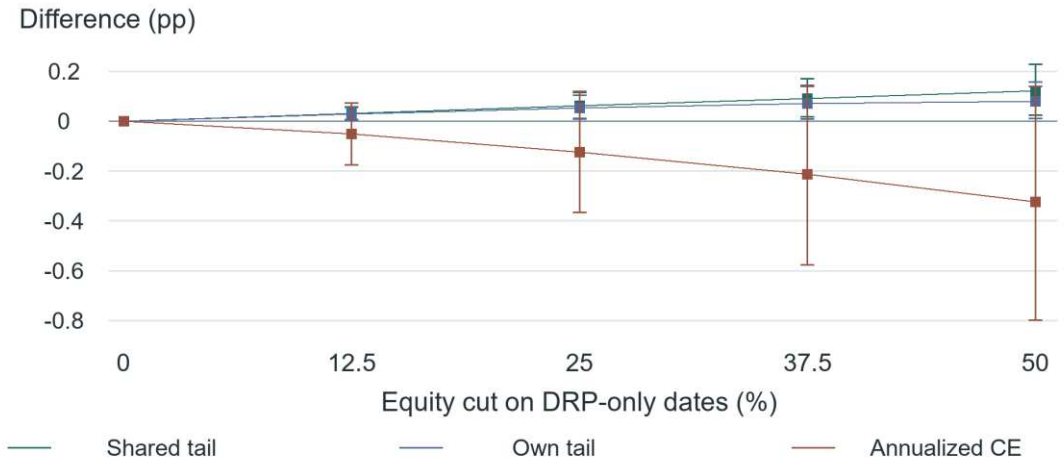


Figure 3. Tail and CE differences relative to zero DRP-only cut, with paired 95% intervals. The two tail series use daily net returns. CE is annualized. See Table 8.

8 Execution sensitivity and scope

With one additional scheduled session of lag, both strategies are scored on the same exchange-grid dates. The overlay retains a better lower-tail mean and a smaller drawdown than the equally delayed RV22 control. The relative CE difference rises to +0.290 percentage points, with a 95% interval of [−0.123, +0.809]. The overlay's own annualized volatility increases from 17.340% to 17.392%, the tail mean falls from −3.196% to −3.217%, and its drawdown rises from 38.218% to 38.867%. These absolute changes and the relative CE estimate answer different questions about delay.

The results concern a fixed overlay in four equity markets, conditional on the two signal definitions and historical calibration. The estimated CE gain depends heavily on the S&P 500. The fixed and flat controls also yield higher CE than the overlay in the base exchange replay. Appendix D reports the inherited forecasting checks separately from the allocation outcomes.

9 Conclusion

A distributional alarm changes both where a volatility rule adds protection and how much exposure it retains on volatility-only dates. On the exchange calendar, the additional cuts on DRP-only dates account for the main positive lower-tail contribution. Those dates contribute negatively to net return and CE. The largest positive CE contribution occurs on RV22-only dates, where the overlay cuts less aggressively than the equal-target control. The relative improvement on each policy's own tail dates persists with an extra execution lag. The CE estimate remains uncertain and sensitive to the market and comparator. On common extreme dates the tail gap versus equal-target RV22 changes sign, and the paired interval spans zero. Keeping the other target weights fixed, the additional 25% cut yields a shared-tail difference of +0.061 percentage points against fixed RV22 and a CE difference of −0.123. The additional alarm expands coverage of severe episodes and gives the rule a distinct way to reduce tail losses. Whether that protection survives an equal exposure budget depends on which dates define the tail. The CE ranking depends on where the full rule retains exposure, the comparison policy and trading costs.

Statements and Declarations

Data and code availability

The index series are based on Yahoo Finance adjusted closes for ^GSPC, ^FTSE, ^N225 and ^HSI. The archived monitoring code is licensed for editorial and peer-review verification. The authors retain frozen inputs and separate allocation scripts. Requests for follow-on materials can be directed to the corresponding author. Access to frozen price-derived inputs and related panels is subject to verification of source terms. The current scripts require local dependencies.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

OpenAI ChatGPT assisted with manuscript editing, reference checks, and analysis and formatting code. The authors take responsibility for the final manuscript.

References

Acerbi, C., & Tasche, D. (2002). On the coherence of expected shortfall. *Journal of Banking & Finance, 26*(7), 1487–1503. DOI

Balduzzi, P., & Lynch, A. W. (1999). Transaction costs and predictability: Some utility cost calculations. *Journal of Financial Economics, 52*(1), 47–78. DOI

Cederburg, S., O'Doherty, M. S., Wang, F., & Yan, X. (2020). On the performance of volatility-managed portfolios. *Journal of Financial Economics, 138*(1), 95–117. DOI

Fleming, J., Kirby, C., & Ostdiek, B. (2001). The economic value of volatility timing. *Journal of Finance, 56*(1), 329–352. DOI

Fleming, J., Kirby, C., & Ostdiek, B. (2003). The economic value of volatility timing using "realized" volatility. *Journal of Financial Economics, 67*(3), 473–509. DOI

Horvath, B., Issa, Z., & Muguruza, A. (2024). Clustering market regimes using the Wasserstein distance. *Journal of Computational Finance, 28*(1), 1–39. DOI

Moreira, A., & Muir, T. (2017). Volatility-managed portfolios. *Journal of Finance, 72*(4), 1611–1644. DOI

Nolte, I., & Xu, Q. (2015). The economic value of volatility timing with realized jumps. *Journal of Empirical Finance, 34*, 45–59. DOI

Page, E. S. (1954). Continuous inspection schemes. *Biometrika, 41*(1–2), 100–115. DOI

Panaretos, V. M., & Zemel, Y. (2019). Statistical aspects of Wasserstein distances. *Annual Review of Statistics and Its Application, 6*, 405–431. DOI

Qiao, W., Bu, D., Gibberd, A., Liao, Y., Wen, T., & Li, E. (2023). When "time varying" volatility meets "transaction cost" in portfolio selection. *Journal of Empirical Finance, 73*, 220–237. DOI

Ross, G. J. (2015). Parametric and nonparametric sequential change detection in R: The cpm package. *Journal of Statistical Software, 66*(3), 1–20. DOI

[Related forecasting study withheld for double-anonymized review.]

[Related monitoring study withheld for double-anonymized review.]

Zeng, Y., Huang, Z., & Chen, X. (2026). Beyond Euclidean summaries: Online change point detection for distribution-valued data. In *Proceedings of the 43rd International Conference on Machine Learning*, Proceedings of Machine Learning Research, vol. 306, pp. 153311–153346. Proceedings

Appendix A Historical episode definitions

Table A1. Event boundaries and market counts

Episode	Inclusive window	Timing anchor	DRP	RV22
Global Financial Crisis	2008-09-01 to 2009-06-30	2008-09-15	4	4
Euro-area debt stress	2011-07-01 to 2011-12-30	2011-08-05	4	2
China-centered sell-off	2015-06-01 to 2016-02-29	2015-08-24	3	3
Volmageddon	2018-01-15 to 2018-04-30	2018-02-05	2	0
COVID	2020-02-01 to 2020-06-30	2020-03-11	3	4
Inflation	2022-01-01 to 2022-10-31	2022-01-05	4	2
Banking stress	2023-03-01 to 2023-05-31	2023-03-10	1	0

Each episode contributes four market–episode cases.

Appendix B Frozen-grid policy results

Table B1. Overlay minus equal-target RV22 on the frozen grid

Execution and markets	Exposure	Mean return	Volatility	Tail mean	Max. DD	CE γ = 5	95% CE interval
Next observation, four markets	−0.333	+0.094	−0.103	+0.028	−1.286	+0.181	[−0.172, +0.677]
Next observation, excluding S&P 500	−0.199	−0.060	−0.064	+0.019	−0.532	+0.002	[−0.203, +0.216]
One extra observation later, four markets	−0.334	+0.173	−0.091	+0.024	−1.282	+0.250	[−0.138, +0.753]
One extra observation later, excluding S&P 500	−0.200	+0.078	−0.057	+0.021	−1.554	+0.133	[−0.178, +0.499]

All differences are in percentage points. The 13,703 common timing observations precede any market exclusion. An extra observation lag applies to both policies.

Table B2. Frozen-grid alarm-state contributions and turnover

Execution and alarm	Rows	Gross P&L pp	Net P&L pp	Overlay turnover	Control turnover	Tail mean pp	CE pp
Next · DRP only	519	−0.244	−0.263	0.282	0.088	+0.142	+0.001
Next · RV22 only	428	+0.291	+0.295	0.185	0.226	−0.105	+0.138
Next · Both	92	+0.077	+0.082	0.046	0.100	+0.000	+0.068
Next · Neither	12664	+0.000	−0.021	0.439	0.227	−0.010	−0.026
Delayed · DRP only	519	−0.268	−0.287	0.282	0.088	+0.131	−0.039
Delayed · RV22 only	427	+0.355	+0.360	0.185	0.226	−0.084	+0.209
Delayed · Both	92	+0.116	+0.121	0.046	0.100	−0.006	+0.105
Delayed · Neither	12665	+0.001	−0.021	0.439	0.227	−0.017	−0.026

Counts sum across markets. Annualized return, turnover, tail and CE contributions are computed within each market and then averaged equally.

On the frozen grid, the S&P 500 contributes +0.718 percentage points to its own full-period next-observation CE difference. The largest positive year contributions are 2025 (+0.627) and 2022 (+0.345), while 2020 contributes −0.218. On RV22-only dates, the net daily difference is +3.743 percentage points on 24 March 2020 and −4.780 percentage points on 16 March 2020. Year contributions allocate full-sample CE to dates.

Appendix C Alternative frozen-grid benchmarks

Table C1. Frozen-grid performance against alternate controls

Outcome	Fixed RV22	RV22 + DRP	Calibrated RV22	Flat exposure
Annualized mean	7.983%	7.693%	7.601%	8.744%
Annualized volatility	17.411%	17.086%	17.197%	17.377%
Mean return on worst 2.5% of return observations	−3.214%	−3.157%	−3.187%	−3.215%
Maximum drawdown	38.528%	37.863%	39.179%	37.800%
Average equity exposure	98.097%	97.150%	97.538%	94.108%
Annual one-way turnover	0.508	0.951	0.642	0.011
Annualized CE at γ = 5	0.226%	0.202%	0.016%	1.053%

Uses all 13,707 frozen evaluation rows and simple returns at a 10 bp one-way cost. Calibrated RV22 is the capped alarm-only rule. The exact-match rule is evaluated separately above.

Table C2. Frozen-grid CE under two return conventions

Risk aversion	Archived CE gap pp	95% interval pp	Simple CE gap pp	95% interval pp
3	+0.124	[−0.206, 0.577]	+0.148	[−0.189, 0.599]
5	+0.161	[−0.214, 0.643]	+0.185	[−0.165, 0.648]
10	+0.253	[−0.315, 0.917]	+0.278	[−0.226, 0.917]

Overlay minus the capped RV22 control in percentage points. The archived version uses the stored log returns directly in its policy calculations. Intervals use paired calendar-year draws.

Table C3. Frozen-grid overlay minus capped RV22 by market

Market	Volatility pp	Daily tail mean pp	Drawdown pp	CE at γ = 5 pp
FTSE	−0.049	+0.033	−0.879	−0.086
HSI	−0.198	+0.036	−0.503	+0.087
N225	+0.024	−0.006	−0.335	+0.023
SPX	−0.220	+0.055	−3.549	+0.718

Simple returns, 10 bp cost. The market rows underlie the equal-market averages.

Appendix D Inherited monitor and forecast checks

The inherited data cover adjusted-close log returns on the S&P 500, FTSE 100, Nikkei 225 and Hang Seng Index. The monitoring history runs from January 2000 through June 2026. The inherited distributional regime monitor, denoted DRP, compares a rolling empirical distribution of 22 returns with low- and high-stress prototypes. Their distances are the average absolute differences between corresponding ordered returns, the one-dimensional W_1 distance for equally weighted samples. The first 126 eligible windows initialize a deterministic two-cluster W_1 k-medians partition. The cluster with the higher root-mean-square prototype is labeled high stress. Each new window is scored before updating the coordinate-wise medians of its assigned cluster.

$$DRP_t = \frac{d_{L,t} - d_{H,t}}{d_{L,t} + d_{H,t}}$$

The volatility benchmark, RV22, is the root mean square of the same 22 returns. For each market and signal, calibration selects the lowest threshold on the pre-2007 quantile grid (90%, 92.5%, 95%, 97.5%, and 99%) for which the alarm-day share does not exceed 5%. A signal must exceed its threshold for five consecutive eligible observations to become an alarm. The threshold is held fixed thereafter.

Across the seven historical episodes in Appendix A, DRP identifies 21 of 28 market–episode cases and RV22 identifies 15. They share 13 detections. Eight cases are unique to DRP and two to RV22. DRP covers three of four markets in the COVID episode, compared with all four for RV22. Across jointly detected cases, DRP's median lead over RV22 is eight observations on the frozen grid. Under the exchange-session sensitivity the coverage comparison is 21/28 versus 14/28 and the paired median lead is 2.5 sessions. The frozen-grid event-level interval for the DRP–RV22 coverage gap is [0.0, 42.9] percentage points. The two-sided seven-episode sign-flip test gives p = 0.1875 under symmetry.

The archived exchange-session forecasting check reports a 0.084% reduction in log MSE for soft DRP routing (HAC p = 0.850) and a 2.659% increase for hard switching (HAC p = 0.0784). Related work on adaptive volatility forecasts also examines the tail behavior of forecast errors [related forecasting citation withheld for review]. These forecasting results are separate from the exposure-policy replay.