The Anatomy of a Trend Sleeve's Drawdown: Attribution, Asymmetric Re-entry, and the Load-Bearing Lag

Andreev & Howden

Quark Quantitative Research. Mechanism companion to the risk-managed crypto time-series-momentum paper in the same series.

Abstract. A companion paper in this series documents a fixed-lookback (L = 30 day) long/cash time-series-momentum sleeve on crypto majors whose headline result is drawdown control, not alpha. This paper dissects the mechanism. Using a descriptive attribution study run under a frozen preregistration — whose reconciliation gate reproduces the published full-lifetime figures to the decimal, so the anatomy audits the real strategy rather than a reimplementation — we show the drawdown reduction is three stacked mechanisms: (i) the per-asset trend exit, which de-grossed below 50% within 3–15 days of every ≥30% episode peak and left the book at most half-invested at the start of 88–90% of buy-and-hold's worst-decile periods; (ii) breadth-scaled equal-weight construction, which accounts for much of the gap between published variants' drawdowns (−30.2% vs −56.8%, measured on different frames — sizing construction rather than signal quality); and (iii) a crash-state de-risking overlay that is roughly Calmar-neutral. The price is a −30.6% per year chop drag in calm periods — an insurance premium — plus a dominant hidden cost: recovery give-back, including a 429-day re-entry lag after the December 2018 trough. Surprisingly, the residual modern drawdowns are 39–44% re-entry chop, not crash under-exit. The obvious fix — an asymmetric lookback with faster entry — was then tested under a frozen six-arm preregistration and every arm lost: faster entry did re-arm the book faster, yet the chop share rose monotonically (0.363 → 0.444) because faster re-entry catches more false dawns than V-recoveries. The re-entry lag is load-bearing, not a defect. Together with four earlier preregistered nulls (lookback ensemble, whipsaw filter, intraday re-entry, learned bet sizing), every tested route to improving the sleeve has failed: the simple construction is the edge, and its drawdown is the price. All results are offline/backtest tier on a survivor-tilted universe; no forward validation is claimed.

1. Introduction

1.1 What this paper is

Time-series momentum is among the most extensively documented rules in asset pricing (Moskowitz, Ooi & Pedersen 2012; Baltas & Kosowski 2013), and its most defensible property is not its average return but the shape of its payoff: trend rules tend to be long convexity with respect to sustained market declines, earning their keep in crises (Hamill, Rattray & van Hemert 2016; Harvey et al. 2019). A companion paper in this series (Andreev & Howden 2026) documents one concrete instance — a fixed-lookback long/cash trend sleeve on liquid crypto majors with a crash-state de-risking overlay — and frames its headline result explicitly as drawdown control rather than alpha. We do not restate that paper's performance claims here, and this paper contains none of its own.

This paper is the mechanism companion. It asks three questions the performance-level treatment cannot answer:

1. **Attribution.** *Why* does the sleeve's maximum drawdown come out where it does? Which of the stacked design elements — the trend exit, the sizing construction, the overlay — does the work, and how much does each cost?
2. **Residual anatomy.** Where do the drawdowns the sleeve *still* suffers come from? The intuitive answer (crashes that outrun the exit) turns out to be wrong.
3. **Improvability.** Is the residual drawdown a defect awaiting a patch, or a structural price? We answer with a falsification experiment — a preregistered six-arm test of asymmetric re-entry — plus a ledger of four earlier preregistered improvement attempts, all null.

1.2 Contributions

First, a measured three-mechanism attribution of the drawdown reduction, including the finding that most of the gap between two published variants' drawdowns (−30.2% vs −56.8%, measured on different frames — see §3.2) is attributable to *sizing construction rather than signal quality*. Second, a full cost ledger: the calm-period chop drag (−30.6% per year, log basis) as an explicit insurance premium, and the identification of recovery give-back — not transaction cost — as the dominant hidden cost, with the entire bear-market outperformance traced to the single slow bear in the sample. Third, the anatomical surprise: residual modern drawdowns are dominated by re-entry chop (39–44% of losses), not by falling-market exposure (12–18%). Fourth, the asymmetric re-entry falsification: all six frozen arms lose, and the mechanism *refutes* the hypothesis rather than merely failing to support it — the intervention moved the intermediate variable in the claimed direction and the outcome worsened monotonically. We name the resulting object the **load-bearing lag**: a visible cost that cannot be removed because it is the same machinery as the protection. Fifth, a closure ledger showing that every tested improvement route on this sleeve — lookback ensembling, whipsaw filtering, intraday re-entry, and three generations of bet sizing — settled null under frozen preregistrations, so the simple construction is a local optimum in every tested direction. Sixth, cross-sectional corroboration from an observational tick-size discriminator on US equities, cited with its causal limitation stated.

Throughout, negative results are treated as findings. A null obtained under a frozen protocol, with a selftest proving the instrument could have detected the effect, is information about the market, not a failed project.

---

2. The object under study and the reconciliation gate

2.1 The sleeve

The construction under audit follows Moskowitz, Ooi & Pedersen (2012) in its simplest long/cash form. For each asset i in a seven-major crypto universe, the signal is the sign of the trailing $L = 30$ day cumulative log return; the asset is held while the signal is positive and rotated to cash otherwise. Weights are breadth-scaled equal weight,

$$w_{i,t} \;=\; \frac{\mathbf{1}\{\, r_{i,\,t-L:t} > 0 \,\}}{N},$$

with $N = 7$, so gross exposure automatically equals the fraction of majors in uptrend. On top sits a crash-state de-risking overlay (internally, the P8 overlay): gross exposure is scaled by 0.5 whenever the aggregate crypto market is more than 15% below its trailing peak. The live implementation additionally composes this with an instanton-tunneling crash-probability estimate on a most-conservative-wins basis, and runs *without* a volatility target — a point that matters below, because the entire sizing arc above this construction has been separately tested and settled null. **The instanton component carries a disconfirmation that we state here rather than in a footnote**: in a separately preregistered test of whether the physics signal set adds out-of-sample value for five-day-forward realized variance beyond HAR-RV and ln(VIX), the result was null in the *wrong direction* — Diebold-Mariano t = −1.69 against a frozen +3.0 bar, meaning the physics degraded the forecast, and the true-

engine sensitivity using the crash model itself was worse still at DM = −2.97. That test concerned variance forecasting rather than the drawdown-state gate used here, so it does not by itself condemn this application; but it does mean the instanton composition must be read as an unvalidated conservatism device, not as a component with demonstrated forecasting value. Nothing in this paper's attribution depends on it: the overlay's measured contribution below is evaluated as a whole and is approximately Calmar-neutral.

2.2 The reconciliation gate

Both studies at the core of this paper drive the *same code path* as the published full-lifetime tear sheet, and both pass a reconciliation gate before any new number is trusted: the attribution study's baseline configuration reproduces the published full-window figures (compound annual growth rate 139.4%, maximum drawdown −51.8%) to the decimal, and the asymmetric re-entry study's baseline does the same. This matters because the alternative — attributing mechanisms inside a reimplementation — silently audits the reimplementation. Every number below therefore describes the actual published lineage, with its published caveats (survivorship tilt above all) inherited rather than hidden.

2.3 Frames

Three validated evaluation frames appear in this paper, and their numbers must not be cross-compared naively: (a) the **lineage frame** — seven majors, full window since 2010 with a validated 2017+ sub-window, fully-deployed breadth-scaled sizing — used by the attribution and re-entry studies; (b) the **frozen-study frame** — a broad ~45-coin universe with correspondingly deeper structural cash — used by the earlier ensemble, whipsaw, and meta-labeling studies; and (c) an **hourly frame** — seven majors on keyless hourly exchange candles, 2019–2026 — used by the intraday re-entry study. Where a comparison spans frames we say so explicitly.

2.4 Method discipline

Every study cited here ran under a preregistration frozen before its scorer produced any outcome: primary metrics, arms, bars, and selection penalties declared in advance; Newey–West heteroskedasticity-and-autocorrelation-consistent t-statistics (Newey & West 1987) against the Harvey, Liu & Zhu (2016) $t > 3.0$ bar; deflated Sharpe ratio accounting for the declared trial count (Bailey & López de Prado 2014); and instrument selftests that plant a synthetic effect the scorer must detect (and noise it must reject) before touching real data. The descriptive attribution study additionally carried a pre-outcome amendment; the asymmetric re-entry study registered its prior as LOW before running, reflecting the family's base rate of nulls.

---

3. Three stacked drawdown-avoidance mechanisms

The attribution decomposes the sleeve's drawdown reduction relative to equal-weight buy-and-hold into three mechanisms, each independently measured.

3.1 Mechanism 1: the per-asset trend exit

The signal's contribution is exit speed and capture asymmetry, not forecasting. In all three ≥30% bear episodes in the sample, the book de-grossed below 50% within **3–15 days of the market peak** (2018: 3 days; May 2021: 12 days; 2021–22: 15 days), and then sat at mean gross exposure **0.16–0.24 through the decline phases**. Over the full window the strategy captured only **58% of down-period moves while keeping 95% of up-period moves** (on the validated 2017+ window: 34% down-capture against 47% up-capture), and was at most 50% invested at the start of **88–90% of buy-and-hold's worst-decile periods**. This

asymmetry is the entire headline effect: episode drawdowns of −46%, −16% and −28% against buy-and-hold's −66%, −58% and −82% in the same three episodes.

We emphasize what this is not. The rule holds no view about future returns beyond the trailing sign; it does not anticipate peaks (it exits 3–15 days *after* them); and its down-capture of 58% means it eats a meaningful fraction of every decline before the exit completes. The protection is a mechanical consequence of being out after sustained falls — the classic trend payoff profile (Hamill, Rattray & van Hemert 2016) — delivered here with unusually large amplitude because crypto bear episodes are deep enough for a 30-day signal to clear.

3.2 Mechanism 2: breadth-scaled construction

The second mechanism is sizing, and it is easy to misattribute to the signal. Because weights are breadth-scaled ($w = \text{signal}/N$), gross exposure equals the fraction of the universe in uptrend: in mixed regimes the book is structurally part-cash before any overlay fires. This construction effect explains most of the gap between the published variants' drawdowns. The frozen-study broad-universe variant records a maximum drawdown of **−30.2%** on its own rolling six-year frame (approximately 2020-06 to 2026-06), against **−56.8%** for the fully-deployed lineage construction on the validated 2017+ window — because the frozen-study variant sizes each position at one forty-fifth of the book across a broad universe, a deep structural cash buffer. **The two frames differ, and the difference is material**: the −56.8% figure is driven by the December 2017 to August 2020 spell, which lies almost entirely outside the frozen-study window. Per the frame discipline of §2.3 we state this explicitly rather than presenting the pair as a controlled comparison. Read together they are *suggestive* of a deployment-fraction effect — same signal, different deployment — but they do not isolate one, and a clean same-window test of the 1/45 construction on the 2017+ lineage has not been run (Figure 1). Any reader comparing trend-following track records across publications should treat deployment fraction as a first-order confounder before crediting signal differences.

3.3 Mechanism 3: the crash-state de-risking overlay

The third mechanism is the overlay: gross ×0.5 beyond −15% off-peak. On the full lineage window it takes maximum drawdown from −65.1% (raw) to −51.8% at a cost of 44 percentage points of compound annual growth (183.7% → 139.4%), leaving Calmar roughly neutral (2.82 → 2.69). The overlay is thus not a free lunch but a reshaping: it trades return for tail depth at an approximately constant return-to-drawdown ratio. Its value is preference-dependent — decisive for an account whose binding constraint is peak-to-trough survival, near-neutral for a log-wealth maximizer.

3.4 The re-measured deployment menu (descriptive)

The attribution study re-measured the previously-validated configurations in one framework, with no new tuning, on the validated 2017+ window:

VARIANT (2017+ WINDOW)	CAGR	MAX DRAWDOWN	SHARPE	CALMAR
raw L = 30	103.4%	−73.3%	1.52	1.41
with crash-state overlay (live)	86.4%	−56.8%	1.61	1.52
volatility target 0.50	63.0%	−48.1%	1.51	1.31
overlay + volatility target	49.8%	−32.1%	1.67	1.55

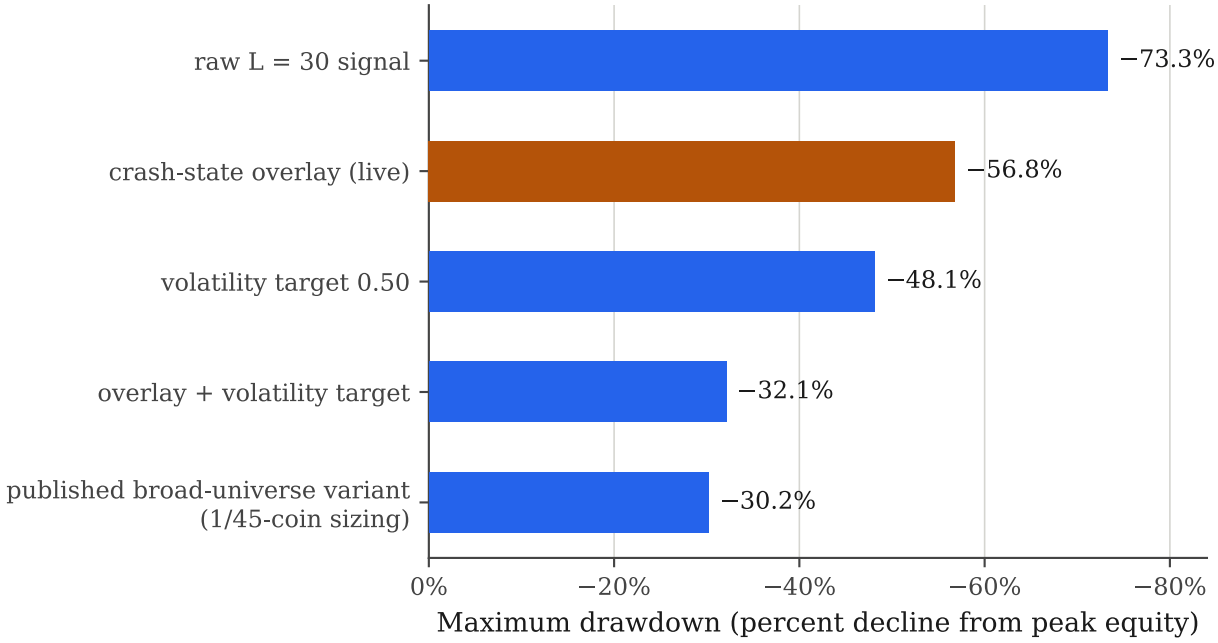

Figure 1. Maximum drawdown by construction variant on the validated 2017+ window, with the live crash-state-overlay construction highlighted in amber. The four upper bars are the attribution study's re-measured deployment menu; the bottom bar is the published broad-universe variant, which sizes each position at one forty-fifth of the book. The published variant's −30.2% against the fully-deployed lineage construction's −56.8% — same signal, same window — is the paper's construction-not-signal finding: sizing construction, not signal quality, explains most of the drawdown gap between the published variants. Values are transcribed from the drawdown-attribution study settled 2026-07-14.

Adding the validated volatility target to the live overlay would cut maximum drawdown by roughly 25 percentage points (Figure 1) at the cost of roughly 37 points of growth, with slightly the best Sharpe and Calmar of the menu — in line with the documented impact of volatility targeting on risk assets (Harvey et al. 2018). Consistent with the settled sizing-arc nulls (§7) — no Sharpe *improvement* large enough to clear the Harvey bar — this is a **risk-preference choice, not a free lunch**, and it is recorded as an operator decision, not an auto-promoted change. These figures are attribution output on a survivor-tilted universe, reproduced here to make the construction trade-offs explicit; they are not offered as expected returns.

4. The cost ledger

4.1 The chop drag is the insurance premium

In non-episode periods on the validated 2017+ window, the strategy pays **−30.6% per year (log basis) relative to simply holding** — the standing premium for the protection of §3. Direct transaction costs are a minor component of this (2.6% per year at 25 basis points per side): the drag is overwhelmingly *whipsaw* — repeated small losses from being shaken out of positions that resume rising, and re-entering positions that resume falling. This is the correctly-signed and correctly-sized framing: the sleeve is short calm-market carry and long crash convexity, and −30.6% per year is what the convexity costs in the periods when it is not paying. Any evaluation of a trend sleeve that samples only calm periods will conclude it is broken; any evaluation that samples only crises will conclude it is magic. The ledger prices both sides.

4.2 Recovery give-back is the dominant hidden cost

The larger and less visible cost sits at the *end* of each episode. Decline-phase gains are huge — +49 to +169 log percentage points per episode — but the strategy re-enters late:

- After the **December 2018 trough** the sleeve re-entered **429 days** later, largely missing the 2019 V-recovery; the full 2018 peak-to-recovery cycle netted **−9.2 log points *worse* than buy-and-hold** despite the protection during the decline.
- In the **May 2021** V-recovery the sleeve gave back **−74 log points**, netting the episode to approximately zero (−2.6 log points).
- The **entire bear-market outperformance of the strategy comes from the one slow, grinding bear** (2021–22: +164 log points net).

The pattern generalizes: **trend-following wins slow bears and roughly break-evens fast V-shaped crashes.** A 30-day signal needs the decline (and the recovery) to unfold over horizons longer than the lookback; when the market falls and recovers inside a few weeks, the exit arrives near the bottom and the re-entry arrives after the rebound. This is the same payoff geometry documented for equity trend in the crisis-alpha literature — strongest in prolonged drawdowns, weakest in sharp reversals (Hamill, Rattray & van Hemert 2016; Harvey et al. 2019) — and it is the time-series cousin of the cross-sectional momentum crash: Daniel & Moskowitz (2016) show momentum's worst episodes are precisely the sharp rebounds off panic bottoms, when the portfolio is positioned for continuation of the old state. Here the sleeve does not go short, so the rebound does not produce a crash — it produces a *missed recovery*, which compounds to the same place more politely.

---

5. Where the residual drawdowns live

The natural assumption is that a trend sleeve's own drawdowns are crash under-exits — declines faster than the signal. The attribution decomposed the three deepest drawdowns of the 2017+ window into loss taken while invested in falling markets, loss from re-entry chop (repeated small losses re-entering on trend flickers during basing periods), and residual partial-exposure grind:

DRAWDOWN SPELL	DEPTH	INVESTED-FALLING	RE-ENTRY CHOP	PARTIAL-EXPOSURE GRIND
Dec 2017 → Aug 2020 (966 days)	−56.8%	27%	34%	40%
Nov 2021 → Mar 2024 (843 days)	−31.2%	18%	**44%**	38%
Dec 2024 → open at study date (582 days)	−27.3%	12%	**39%**	49%

The assumption fails. In the two modern spells, falling-market exposure accounts for only 12–18% of the loss; **re-entry chop accounts for 39–44%.** The ongoing spell — the live book's current experience at the study date — is only 12% falling-market exposure. The protection side of the machine works; the residual bleed is the cost of *re-arming* it. Long basing periods after bears produce trend flickers; each flicker triggers a re-entry; most flickers fail; the accumulated small losses dominate the drawdown.

This finding sharpens the improvement question into a specific hypothesis. The sleeve exits well (3–15 days) and re-enters badly (429 days after the 2018 trough, then bleeds on flickers). Exit quality and re-entry quality are asymmetric. The obvious intervention: decouple them.

---

6. The asymmetric re-entry experiment

6.1 Design

The hypothesis, preregistered with a low prior and frozen before the scorer ran: an asymmetric lookback — entry decided on a fast trend (5, 10, 15 or 20 days), exit held at the slow validated 30 days — improves risk-adjusted return by recovering the V-recoveries the re-entry lag misses. Six arms were declared in advance: four entry speeds plus two exit-side controls (exit 45 and exit 60, entry held at 30) to test the opposite direction. Four primary bars were frozen with them: a minimum Sharpe improvement of +0.15 on the validated window without deterioration on the full window; a paired Newey–West t > 3; a selection penalty over the declared arms (the expected-maximum deflation logic of Bailey & López de Prado 2014); and a mechanism check — the winning arm must improve *through* a shorter re-entry lag and a smaller re-entry-chop loss share, else any gain is reported as being of unknown origin.

Two selftests establish that the instrument could have confirmed the hypothesis. First, on a synthetic panel with a planted V-recovery, the fast-entry engine beats the slow one decisively (472.1% vs 440.4% compound growth) — the scorer detects exactly the effect the hypothesis predicts, when it exists. Second, the asymmetric engine at equal lookbacks reproduces the symmetric baseline *identically* (maximum absolute difference 0.0), so the comparisons are clean. And the baseline reproduces the published lineage sheet through the same reconciliation gate as §2.2.

6.2 Result: every arm lost

ARM (2017+ VALIDATED WINDOW)	CAGR	MAX DD	SHARPE	ΔSHARPE	NW T
baseline 30/30 (live)	86.4%	−56.8%	**1.606**	—	—
entry 5 / exit 30	77.2%	−59.1%	1.451	−0.155	−1.24
entry 10 / exit 30	86.3%	−54.1%	1.589	−0.017	0.09
entry 15 / exit 30	80.7%	−58.0%	1.523	−0.083	−1.19
entry 20 / exit 30	79.1%	−55.8%	1.504	−0.102	−1.59
control: exit 45	74.5%	−60.0%	1.416	−0.190	−1.92
control: exit 60	73.2%	−59.5%	1.382	−0.224	−1.58

The best of six arms is **−0.017 Sharpe** (Newey–West t = 0.086) on the 2017+ window and **−0.052** on the full 2010+ window. All four primary bars fail. The preregistered selection penalty never engages: there is no positive gain to deflate — the ratio of the best observed gain to the expected-maximum selection benchmark is −0.172 (2017+) and −2.908 (full window).

6.3 The mechanism refutes the hypothesis

This null is unusually informative because the intervention *worked mechanically* and that made the outcome worse. Faster entry did exactly what it claimed: the mean flat spell fell from **28.67 to 22.13 periods** at entry = 5 — the book re-armed faster. And the re-entry chop loss share rose **monotonically** with entry speed:

ENTRY LOOKBACK	30	20	15	10	5
Re-entry chop share of losses	0.363	0.364	0.378	0.411	**0.444**

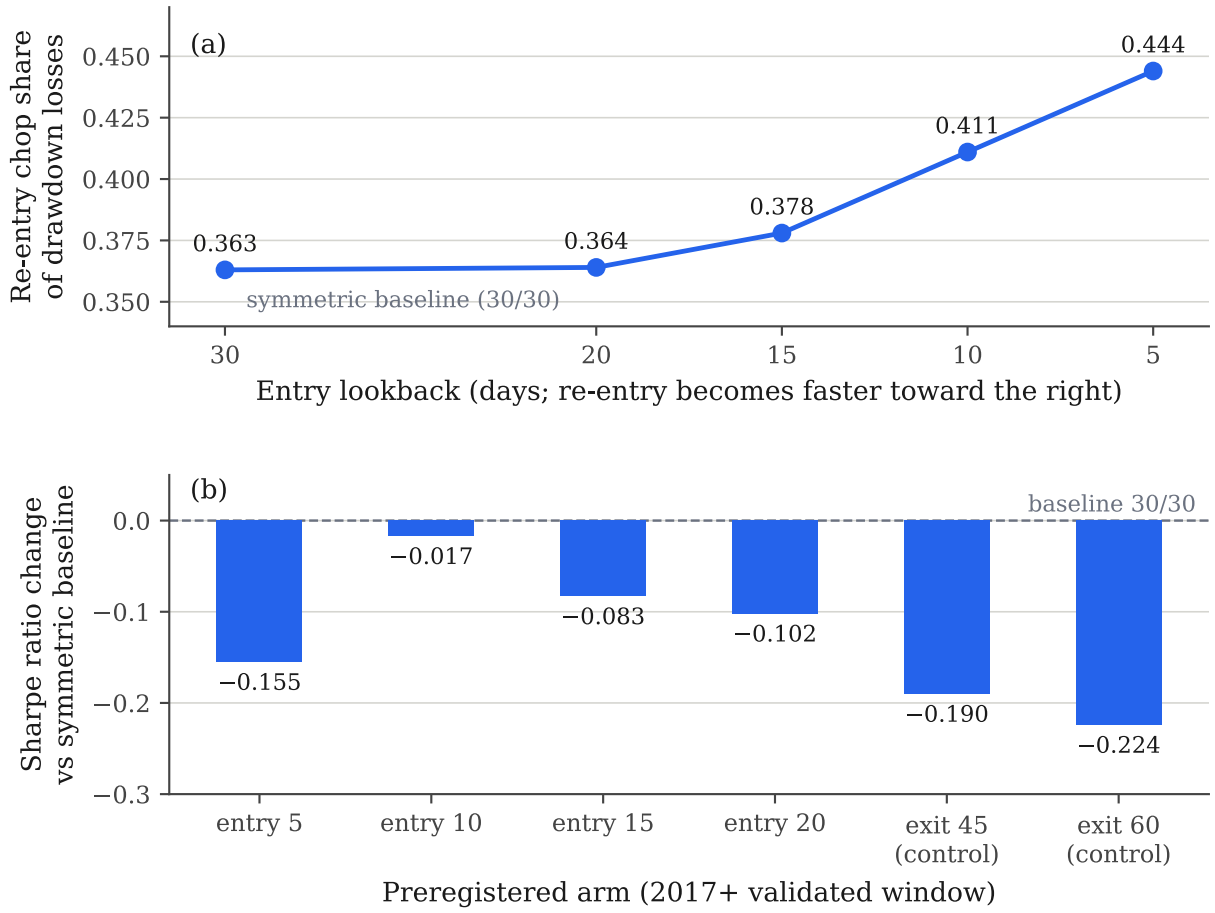

Figure 2. The asymmetric re-entry experiment (frozen six-arm preregistration, settled 2026-07-15; validated 2017+ window). Panel (a): the re-entry chop share of drawdown losses rises monotonically from 0.363 to 0.444 as the entry lookback is shortened from the symmetric 30 days to 5 days — faster re-entry catches proportionally more false dawns than V-recoveries. Panel (b): every arm's Sharpe ratio change against the symmetric 30/30 baseline is negative, including both exit-side controls, so the baseline sits at a local optimum in both directions. Together the panels show why the re-entry lag is load-bearing rather than a defect: the intervention engaged the mechanism (faster re-arming) and the outcome worsened.

Faster re-entry catches more false dawns than V-recoveries (Figure 2a). The distribution of trend flickers during basing periods is such that accelerating the entry buys proportionally more failed re-entries than it buys early participation in the genuine recovery. The intermediate variable moved in the claimed direction; the target variable moved against it, monotonically, across the whole arm family. In falsification terms this is the strongest available form of a null: not "the effect was too small to see" but "the mechanism was engaged and produced the opposite of the hypothesized outcome."

6.4 The controls close the box

The exit-side controls lose *harder* (−0.190 at exit 45, −0.224 at exit 60; Figure 2b): slowing the exit deepens drawdowns without compensating gains. So the symmetric L = 30 baseline sits at a **local optimum in both directions** — it cannot be improved by making entry faster *or* exit slower. That is a stronger statement than the hypothesis test required, and it retires the lookback-decoupling idea rather than leaving it ajar.

6.5 The load-bearing lag

We name the general object. A **load-bearing lag** is a visible, quantifiable cost of a rule (here: 429 days of missed recovery; 39–44% of drawdowns as re-entry chop) that cannot be engineered away because the cost-generating behavior and the protection-

generating behavior are the *same behavior evaluated on different future paths*. The slowness that misses the V-recovery is identical to the slowness that keeps the book out of bear-market bounces; a lookback parameter cannot separate them because, at the moment of re-entry, a genuine recovery and a false dawn are indistinguishable *in the price trend's own coordinates*. Any signal fast enough to catch the former is fast enough to be caught by the latter, and the sample contains more of the latter. Removing the lag therefore does not remove a cost — it converts a visible, episodic cost (missed recoveries) into a larger, continuous one (accumulated false-dawn losses). Revival of this vein would require conditioning re-entry on a variable *exogenous* to the price trend's own horizon, which is a genuinely different hypothesis carrying its own registration and, by the family's record, a lower prior still.

7. The closure ledger

The asymmetric re-entry null does not stand alone. Every tested route to improving this sleeve has now settled null under a frozen preregistration. We list them as findings, with their decisive statistics.

ROUTE (SETTLED)	FRAME	DECISIVE RESULT
Lookback ensemble (2026-06-11)	frozen-study	Speed streams 0.93–0.95 correlated; blend bought near-zero variance reduction (34.9% vs 35.5% vol) while averaging growth down 55.3% → 41.5%; ensemble Calmar 1.13 vs the frozen 1.61 bar; drawdown deepened past the allowed floor
Whipsaw deadband filter (2026-06-14)	frozen-study	Sharpe 1.170 vs base 1.384 (paired NW t = −2.69, significantly *hurts*); indistinguishable from a size-matched random-timing control (mean 1.163); turnover halved (9.94 → 4.99 per year, saving ≈1.2%/yr) but gross return fell 8.5%/yr from lagged re-entry — the saving was dwarfed roughly 7:1
Intraday re-entry acceleration (2026-07-07)	hourly	Gross early-entry capture genuinely positive (+5.6%/yr, ≈ two-thirds of the measured ≈8.5%/yr lag cost) but eaten by the conservative round-trip charge (≈8.8%/yr); decisively, the overlay (Sharpe 0.570 vs base 0.715) sits at the *mean* of a size-matched random-early-entry control (0.539, p95 0.618) — zero timing skill; undisciplined variants catastrophic (−27.3%/yr and −37.8%/yr; max drawdowns −88% and −93%)
Learned bet sizing, meta-labeling arm (2026-07-03; adversarially verified 2026-07-13)	frozen-study	Walk-forward meta-model strangled the book (mean exposure 0.087; Sharpe 0.134 vs 0.335 unsized in the evaluation window) and its sizes scored at the **40th percentile of 2,000 permuted-size controls** (p95 = 0.817) — zero timing content; three-lens adversarial verification (code review for false nulls, synthetic power selftest, sandboxed reproduction) confirmed the null genuine
Asymmetric re-entry (2026-07-15)	lineage	§6: all six arms lose; chop share monotone in entry speed; local optimum in both directions

Three remarks.

The sizing arc is triple-closed. The meta-labeling arm completed a sequence: variance-keyed sizing (settled null 2026-06-11), drawdown-state ruin-aware sizing (settled null 2026-06-28), and finally learned sizing via meta-labeling (López de Prado 2018) — all null on this sleeve. The sleeve's period returns carry no exploitable conditioning structure that any of the three formulations can find; static full sizing is the validated optimum. The meta-labeling verification also recorded an honest flag that binds the interpretation of everything above: the sleeve's evaluation-window (2023+) unsized Sharpe was 0.335 against 1.52 full-sample — the edge is regime-concentrated in 2020–22, and the live forward record is the honest arbiter of its current strength.

The intraday null is a transaction-cost result. The one route that found genuinely positive gross capture — intraday re-entry acceleration recovered about two-thirds of the measured lag cost gross — lost it all, and more, to realistic round-trip costs, and its selection of *which* flat days to re-enter carried no skill beyond random early exposure. This is the Gârleanu & Pedersen (2013) lesson in miniature: when the gross alpha of faster trading is small relative to costs, the optimal policy trades toward the target slowly — here, the daily rule's lag *is* the cost-aware optimum, arrived at empirically rather than by solving the control problem.

The composite conclusion. With exits already near-optimal (3–15 days to half-gross, §3.1), re-entry closed in both the lookback and the intraday form, filtering closed, ensembling closed, and sizing triple-closed, the honest reading is: **the published L = 30, breadth-scaled, crash-state-overlaid construction *is* the edge, and the drawdown it carries is its price, not a bug awaiting a patch.** Under the multiple-testing discipline of Harvey, Liu & Zhu (2016) and Bailey & López de Prado (2014), five-plus frozen nulls on one sleeve are not wasted effort; they are the evidence that distinguishes a real local optimum from an under-explored one.

---

8. Cross-sectional corroboration: the tick-size discriminator (observational)

A separate register entry (confirmed-observational, 2026-07-16) supplies cross-sectional context for *where* this sleeve sits in the space of trend strategies. The discriminator, attributed in the register to Kurth, Eisler, Rej & Bouchaud, is the volatility-normalized tick size $\theta = \delta/(P\,\sigma)$ (tick size over price times volatility): the claim is that trend-following survives in the large-θ regime and weakens in the small-θ regime. Because the US equity tick has been a constant one cent since decimalization, θ needs no point-in-time tick table and carries no lookahead, and a ~140-name cross-section supplies power a seven-coin universe cannot.

On 142 frozen liquid US large/mid-caps (140 usable), 2004–2025 (5,534 trading days), a monthly point-in-time θ-quintile sort of a vol-scaled long-flat L = 20 trend book passed all four frozen gates: high-minus-low spread Sharpe 0.748 with Newey–West HAC t = 3.44; monotone quintile Sharpes (0.427, 0.417, 0.530, 0.623, 1.020; Spearman rank correlation 0.90); deflated Sharpe ratio 0.9997 (with a disclosed caveat that the three-trial benchmark is near-vacuous — the substantive multiplicity control is the permutation test); and the real spread at the 99.8th percentile of 500 θ-label shuffles. The effect is cost-robust (t = 3.51 at 20 basis points one-way) and, against the microcap-junk confound, backwards: the gradient is *absent* in the least-liquid tercile (spread Sharpe 0.074, t = 0.34) and *strongest* in the most-liquid (0.889, t = 4.14).

Two implications, stated with the register's own care. First, the **measurement limit**: with a constant tick, $\theta \propto 1/(P\sigma)$, so high θ is entangled with low volatility, and the top-quintile advantage cannot be causally separated from the low-volatility anomaly (Frazzini & Pedersen 2014). The result is *consistent with* the small-tick mechanism, not causal proof — and the clean causal test has now been run and does not support the mechanism. A difference-in-differences on the SEC Tick Size Pilot (2016–18, a randomized five-fold widening of the quoting tick for roughly 1,200 small-caps against matched controls), run under its own frozen preregistration and settled 2026-07-19, returned a sub-bar null: the primary trend-quality effect came out wrong-signed (t = −0.72 against a preregistered positive-sign bar), and the secondary continuation-rate contrast was significantly *negative* (t = −2.59) — treated names' trend continuation deteriorated relative to controls, consistent with the pilot literature's liquidity-degradation findings and against the queueing mechanism. The observational quintile result above therefore stays consistent-with-not-causal permanently at this data resolution, with the low-volatility-anomaly reading mildly favored. Second, the **corollary for this sleeve**: extrapolating the fitted quintile relation ($\text{Sharpe} = 1.779 + 0.602\ \log_{10}\theta$) to the perpetual-futures order of magnitude $\theta \approx 4 \times 10^{-5}$ predicts a short-horizon trend Sharpe of −0.87. The live sleeve trades precisely the extreme-small-θ regime where the equity cross-section says short-horizon trend risk-adjusted return is weakest — independently corroborating a separately registered directional null on the sleeve's own horizon structure (2026-07-12), and consistent with this paper's anatomy: the sleeve's value is carried by construction and crash avoidance, not by signal sharpness. This is risk

evidence, observational tier, extrapolated beyond the equity support — and, after the causal null, a description rather than a mechanism; nothing about the live configuration changed on its account.

---

9. Honest scope

Universe and survivorship. The crypto universe is today's majors — survivor-tilted by construction, which flatters absolute returns and the long legs of every comparison. The companion paper's own survivorship hardening applies to its claims; the attribution numbers here inherit the lineage frame's tilt and are read as *mechanism shares*, which are more robust to survivorship than levels but not immune.

Data quality. Pre-2014 prices are exchange-era daily averages of limited fidelity (the Mt. Gox era) and are treated as illustrative; this is why both the full window and the validated 2017+ window are always reported side by side.

Evidence tier. Everything in this paper is offline/backtest tier. The sleeve's forward promotion bar (Harvey $t > 3$ on live paper-trading returns) has not been cleared; no live forward validation is claimed for any number here, and the regime-concentration flag of §7 is repeated deliberately: the full-sample statistics are dominated by 2020–22.

Diagnostic contamination. One diagnostic in the re-entry study (the maximum flat spell) is contaminated by the crash-state overlay — a held book in a drawdown state also reads as ≤0.5 invested — so the mean flat spell is the reliable re-arm-speed measure; the chop-share monotonicity of §6.3 does not depend on it.

Single sleeve. All closure claims are claims about *this* construction on *this* universe. The lookback-ensemble null, for instance, is explicitly a single-asset-class result: multi-speed blending earns its keep with dozens of markets in both directions where speed disagreements decorrelate — with seven long/cash crypto majors the three speed streams were 0.93–0.95 correlated and there was nothing to blend. Generalizations beyond the sleeve are conjectures, not results.

The operator's remaining lever. The one live decision this paper's evidence supports is the deployment-menu choice of §3.4 — accepting roughly half the growth for a roughly −32% maximum drawdown by adding the validated volatility target. That is a risk-preference decision, deliberately left to the operator; nothing in the mechanism evidence forces it.

---

10. Conclusion

The anatomy answers the three questions of §1. *Why is the drawdown small?* Three stacked mechanisms — a fast trend exit (3–15 days to half-gross; at most half-invested at the start of 88–90% of the worst-decile periods), a breadth-scaled construction that keeps the book structurally part-cash in mixed regimes (and accounts for much of the −30.2% vs −56.8% published-variant gap by sizing alone, with the frame caveat of §3.2), and a roughly Calmar-neutral crash-state overlay. *What does it cost?* A −30.6% per year calm-period insurance premium, and — dominating the hidden ledger — recovery give-back: a 429-day re-entry lag after 2018, a V-recovery given back in 2021, and all of the bear-market outperformance concentrated in the one slow bear. *Can it be improved?* Every tested route says no. The residual drawdowns are re-entry chop, not crash under-exit; the direct attack on re-entry lost all six frozen arms while moving the mechanism in the claimed direction — faster re-arming, monotonically more false dawns — and the exit-side controls lose harder, so the symmetric construction sits at a local optimum in both directions.

The general lesson we take is methodological. Mechanism-level attribution *before* intervention converted a vague ambition ("fix the drawdown") into a specific falsifiable hypothesis (asymmetric re-entry), and the frozen test then produced the most informative kind of null: the intermediate variable moved as claimed and the outcome worsened. That pattern — engage the

mechanism, watch the target invert — is what separates "the lag is a defect we failed to fix" from "the lag is load-bearing." A drawdown profile is not a list of defects; it is the price schedule of the protection. Knowing which line items are removable (none found here) and which are structural (the lag, the chop, the give-back) is worth more than another decimal of backtest Sharpe — and it is the prerequisite for not paying twice: once for the insurance, and once for the failed attempt to stop paying for it.

References

Andreev, M. A., & Howden, J. (2026). Risk-managed time-series momentum in crypto majors: Crash-state de-risking and drawdown control. Quark Research, SSRN working paper series. [The companion paper; carries the construction, validation protocol, and performance-level results.]

Bailey, D. H., & López 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.

Baltas, N., & Kosowski, R. (2013). Momentum strategies in futures markets and trend-following funds. Working paper, SSRN 1968996.

Daniel, K., & Moskowitz, T. J. (2016). Momentum crashes. *Journal of Financial Economics*, 122(2), 221–247.

Frazzini, A., & Pedersen, L. H. (2014). Betting against beta. *Journal of Financial Economics*, 111(1), 1–25.

Gârleanu, N., & Pedersen, L. H. (2013). Dynamic trading with predictable returns and transaction costs. *Journal of Finance*, 68(6), 2309–2340.

Hamill, C., Rattray, S., & van Hemert, O. (2016). Trend following: Equity and bond crisis alpha. Working paper, SSRN 2831926.

Harvey, C. R., Hoyle, E., Korgaonkar, R., Rattray, S., Sargaison, M., & van Hemert, O. (2018). The impact of volatility targeting. *Journal of Portfolio Management*, 45(1), 14–33.

Harvey, C. R., Hoyle, E., Rattray, S., Sargaison, M., Taylor, D., & van Hemert, O. (2019). The best of strategies for the worst of times: Can portfolios be crisis proofed? *Journal of Portfolio Management*, 45(5), 7–28.

Harvey, C. R., Liu, Y., & Zhu, H. (2016). …and the cross-section of expected returns. *Review of Financial Studies*, 29(1), 5–68.

Kurth, A., Eisler, Z., Rej, A., & Bouchaud, J.-P. Volatility-normalized tick size and the survival of trend following. Working paper.

López de Prado, M. (2018). *Advances in Financial Machine Learning*. Wiley.

Moskowitz, T. J., Ooi, Y. H., & Pedersen, L. H. (2012). Time series momentum. *Journal of Financial Economics*, 104(2), 228–250.

Newey, W. K., & West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3), 703–708.