

Strategy Due Diligence:

combined_sleeve_v1

Instrument: long-only US equity and ETF sleeve (three-edge blend) **Status:** ARCHIVED. Live entries halted. Not offered for allocation. **Document date:** 2026-07-25 **Prepared for:** allocator and operational due-diligence review **Register:** this is a factual due-diligence record. It documents a strategy that was validated, deployed, then failed its own pre-registered acceptance tests and was retired. It makes no forward performance claim.

1. STRATEGY DESCRIPTION

combined_sleeve_v1 is a long-only sleeve that allocates across US large-cap equities and a small set of liquid ETFs. It holds no shorts, no options, and no leverage, and was designed to run in a cash brokerage account with zero commission. It is an equal-weight blend of three sub-strategies, one third of capital to each:

1. **Trend filter overlay.** Holds only names trading above their 200-day simple moving average, then volatility-targets the resulting book. A trend-participation filter with a risk-control layer.
2. **Sector and cross-asset rotation.** Ranks a set of sector and cross-asset ETFs on trailing momentum and rotates into the leaders, parking in short-duration bonds (SHY) when momentum is absent. A defensive rotation leg.
3. **Crash-protected momentum.** A 12-month-minus-1-month cross-sectional momentum book that moves entirely to cash when the S&P 500 is below its 200-day moving average. A bear-market shield that works by de-risking rather than by hedging.

Signals are computed on daily closes and lagged one bar before execution, so no bar's signal uses that bar's own close. The blend weight is fixed at one third per leg. This equal weighting was a deliberate anti-overfitting choice: an inverse-volatility weighting produces near-identical weights (approximately 0.33 / 0.36 / 0.31), so no allocation tuning is embedded in the blend.

The intended risk profile was explicitly not “beat the S&P on raw return.” It was to deliver a comparable or better risk-adjusted return with materially lower drawdown, suited to a small account operating under a 5 percent daily-loss circuit breaker. Position exits used a 3 percent hard stop and an 8 percent trailing stop.

The strategy was implemented in two code paths: a research reference and a live execution port, with a fidelity test asserting the two produced identical weights to within $1e-12$. This document concerns the strategy as validated and as deployed. Where the two diverged in risk, that divergence is documented in Section 3.

2. VALIDATION METHODOLOGY

The strategy was assessed against a fixed acceptance battery, referred to internally as the gauntlet. The thresholds were pre-registered, meaning they were fixed in writing before any result on the point-in-time universe existed, so the bar could not be moved after seeing a number. Every figure cited in this document was recomputed from the project environment rather than quoted from a prior report; the two independent reproductions agreed to three decimal places on the headline Sharpe ratios.

The gauntlet has six components:

2.1 Point-in-time universe. The original backtest ran on a 48-name universe that was frozen in mid-2026 and contained only companies that were alive and liquid at that date. This is a textbook survivorship-selection defect: any company that was delisted, went bankrupt, or was acquired between 2019 and 2026 cannot appear, so the tradeable opportunity set is reconstructed only from its winners. The validation therefore rebuilt a point-in-time constituent universe from a historical membership source (1,206 symbols across 2,015 membership-change events, spanning 2019 to 2026), so that each name contributes only over the window in which it was actually an index member.

2.2 Delisting-to-loss handling. A survivorship-free universe is only honest if the names that left are booked at their real outcome. The frame books a delisted name to a terminal loss rather than carrying its last price forward. Forward-filling a dead name would convert a bankruptcy into a flat, riskless hold and silently flatter returns; that failure mode is explicitly guarded, and the terminal loss survives the return computation instead of being masked to zero.

2.3 Deflated Sharpe ratio with an honest trial count. The deflated Sharpe ratio (DSR) discounts a strategy's Sharpe for the number of configurations that were tried before it was selected, which is the correct defence against multiple-testing. The critical methodological point is the trial count. The strategy's earlier "passing" DSR of 0.9709 was computed against a trial count of 13. The honest count from the full research trials ledger is 462 registered trials (64 with a non-null Sharpe). The DSR was re-computed against the full 462-trial count.

2.4 Fixed placebo test. The placebo compares the strategy against random baskets drawn from the same universe with matched turnover. If a random-selection basket reproduces the strategy's result, the "edge" is exposure, not selection skill. The placebo returns a percentile; the pre-registered rule requires a percentile of at least 0.95 to clear, and a null or errored placebo counts as a fail with a stated reason rather than a pass.

2.5 Coverage gate (fail-closed data quality). Before any Sharpe is computed, a data-quality gate checks that the price cache is complete enough to trust: at least 95 percent symbol coverage, at least 80 percent of in-window leavers priced, and zero suspect-truncated series. This gate is fail-closed. If it does not pass, the battery refuses to compute a Sharpe at all, on the reasoning that a Sharpe on a thin, half-missing-leavers cache would simply mint a fresh survivorship-biased number, which is the exact defect under audit.

2.6 Pre-registered acceptance thresholds and decision rule. Five criteria were fixed before the run: full-period Sharpe at least equal to the S&P over the same window; conservative DSR at least 0.95; the Monte-Carlo median gate passed; fixed-placebo percentile at least 0.95; and regime-overlay drag no greater than 15 percent. The decision rule was mechanical and set in advance: all gates pass implies arm; any gate fails implies archive; a partial pass implies validate-only with entries kept halted.

Every methodological choice in the rebuild was set to bias the measured edge against the strategy, not toward it: membership masking can only remove a name's contribution, never add it; delisting-to-loss adds real losses the survivor universe never saw; and deflating by the full 462-trial count applies the largest honest haircut. A strategy that fails under this stack fails despite a conservative deck, not because of a rigged one.

3. RESULT: THE STRATEGY FAILED

This section documents a kill. The strategy did not clear its pre-registered bar and was retired. That outcome is the validation process working as designed, and it is reported here in full rather than suppressed.

3.1 The headline number did not survive its own robustness check. The figure that originally justified the sleeve was an out-of-sample (2023 onward) Sharpe of 1.58 on the frozen 48-name universe. Two independent defects hollowed it out. First, 54.8 percent of the out-of-sample profit came from ten hindsight winners in the universe; removing eight of them dropped the out-of-sample Sharpe from 1.58 to 1.16, below the S&P's out-of-sample Sharpe of roughly 1.43. Second, the in-sample-to-out-of-sample jump was a regime-placement artifact, not evidence of generalisation: the strategy's single losing year, 2022, sat entirely inside the in-sample window by construction, and the same in-sample-to-out-of-sample jump appears in a raw S&P buy-and-hold that was never fit to anything. The apparent out-of-sample strength was universe selection meeting a bull market.

3.2 On the honest universe the edge was not present. When the strategy was re-measured (without any retuning) on the point-in-time universe with delisted names booked to loss, it scored a full-period Sharpe of 0.64 against the S&P's 0.93 over the same window. Total return was 16.9 percent against the S&P's 234 percent. Removing the eight hindsight winners barely moved the point-in-time result (a Sharpe delta of 0.004), because the point-in-time frame had already removed the concentration. What remained was a diluted long-only market tilt, not alpha.

3.3 The deflated Sharpe was zero on an honest count. Under the full 462-trial deflation the conservative DSR was 0.000, against a pre-registered threshold of 0.95. The strategy fails its own multiple-testing gate. The earlier 0.9709 "pass" was an artifact of an understated trial count and is void.

3.4 The placebo failed. The fixed-placebo percentile was 0.913 against a pre-registered threshold of 0.95, with the "beats random" flag false. The strategy did not clearly outperform matched random selection from the same universe.

3.5 The final verdict. Three of five pre-registered gates failed: full-period Sharpe (0.64 versus a 0.93 threshold), honest DSR (0.000), and fixed placebo (0.913). Under the pre-registered decision rule, any fail means archive. The strategy was archived on 2026-07-19.

3.6 A material honesty note on the verdict's basis. Free data sources could not fully satisfy the coverage gate; leaver coverage capped at 57.8 percent because an external price provider rate-limited the remaining names. Rather than pay for data to resolve a strategy already in doubt, the battery was run as a labeled upper bound. Missing names are disproportionately delisted losers, so every figure in Section 3.2 through 3.5 is flattered in the strategy's favour. A strategy that fails at a flattered upper bound is a decision-safe kill; only a pass would have required paid data to confirm. The archive decision is therefore conservative in the correct direction.

3.7 Deployed risk exceeded validated risk. Independent of the edge question, the deployed system did not run at the validated risk profile. The live allocator renormalised away the strategy's cash buffer, running the book at roughly 1.5 times validated risk: realised volatility ran toward 17 percent against a validated 12 percent, and maximum drawdown toward 18 percent against a validated 11 percent. A live regime-sizing multiplier further meant the deployed sleeve was, strictly, a different strategy from the one validated. This is recorded as a governance and change-control finding in its own right.

3.8 Live outcome and cost. The sleeve traded live for a short window on a real cash account of approximately 1,000 US dollars, reaching around 20 holdings. The live record is not statistically informative: the position sat flat on a majority of days owing to a signal-mapping defect, so the as-lived Sharpe was near zero and constitutes evidence neither for nor against the edge. Total realised cost of the live episode was 4.35 US dollars. Entries are halted and remain halted. The residual book is held pending an explicit manual instruction to liquidate; downside protection remains armed until the book is flat.

3.9 What the process produced. The durable output of this exercise is not a strategy but an admission stack: a point-in-time and delisting-aware universe, an honest deflated-Sharpe count, a fixed placebo, a Monte-Carlo median gate, and a fail-closed coverage gate. Every future candidate is required to pass through the same gauntlet before any capital decision.

4. RESIDUAL-RISK REGISTER

The validation above covers the strategy's statistical soundness and its provenance. It does not cover the operational controls an allocator weights most heavily. The following

controls are not yet in place. They are listed here in full rather than omitted, because naming what is deferred is itself a diligence signal. All items sit with the principal, who currently operates as a single individual with no segregation of duties.

#	Control	Status	Why not yet in place	Owner
1	Governance and segregation of duties (independent oversight; four-eyes approval on any configuration, universe, or risk-limit change; a person who checks the principal)	DEFERRED	Single-person operation; the same individual researches, deploys, and controls risk limits. This is the single largest operational gap and the first question an allocator should ask. The change-control failure in Section 3.7 (a live allocator silently altering the validated risk profile) is a direct consequence.	Principal
2	Independent valuation and pricing	DEFERRED	Position values are read from the executing broker and from internally computed marks with no independent second source.	Principal

#	Control	Status	Why not yet in place	Owner
3	Custody and cash controls	DEFERRED	No third party reprices the book. Assets sit in the principal's own brokerage and exchange accounts. There is no independent custodian and no segregation of client from principal assets.	Principal
4	Tested disaster recovery and business continuity	DEFERRED	No documented, exercised recovery procedure for the trading host, its state, or its keys. A restore drill that recovers state from backup and re-verifies it is planned but not yet routine.	Principal
5	Key storage and credential handling	DEFERRED	The exchange private key is held in an environment-secrets file on the trading host and the	Principal

#	Control	Status	Why not yet in place	Owner
			corresponding public address is recorded in plaintext in project documentation. There is no hardware security module, no multi-party signing, and no key-custody separation. An operational reviewer will flag this.	

Each row is a known gap with a named owner, not an oversight. None was built as part of the validation work documented above, and none should be assumed present.

5. WHAT IS AND IS NOT CLAIMED

Claimed: - The strategy was validated against a fixed, pre-registered acceptance battery. - On an honest point-in-time universe with delisted names booked to loss, the strategy failed three of five pre-registered gates and was archived. - Every figure in this document was recomputed from the project environment; the two headline reproductions agreed to three decimal places. - The archive verdict was reached at a data bound that flatters the strategy, so the kill is conservative in the correct direction. - The deployed system ran at approximately 1.5 times its validated risk, which is recorded as a change-control finding. - The operational controls in Section 4 are deferred and are not in place.

Not claimed: - No forward return, Sharpe, or performance figure is offered. The strategy is retired. - The short live episode (total cost 4.35 US dollars) is not offered as a track record; it is statistically uninformative. - The archived point-in-time figures (full

Sharpe 0.64, DSR 0.000, placebo 0.913) are not offered as the strategy's "true" performance. They are an upper bound on a failed candidate, computed to reach a decision, not to measure an edge. - No governance, custody, valuation, disaster-recovery, or key-management control is claimed to exist beyond what Section 4 states. - No representation is made that this is an institution-grade operation today. The single-person structure and the deferred controls in Section 4 are stated plainly.

The defensible claim from this record is narrow and specific: a functioning honest-validation process exists, and it killed the principal's own strategy at pre-registered thresholds rather than rationalising it into deployment. The operational controls that would make the surrounding activity institution-grade are named, deferred, and tracked.