

A Falsification-First Framework for Systematic Strategy Validation Under Proprietary-Capital Constraints

Sixty-five pre-registered hypothesis tests, dual-engine model validation, and Monte Carlo capital design for evaluation-gated futures accounts

Ali Asfar

Bayn Financial (independent research)

August 2026

Methodology disclosed in full. Strategy identity, parameters, and signal logic redacted by design. Research documentation, not investment advice.

A Falsification-First Framework for Systematic Strategy Validation Under Proprietary-Capital Constraints

Sixty-five pre-registered hypothesis tests, dual-engine model validation, and Monte Carlo capital design for evaluation-gated futures accounts

Ali Asfar | Bayn Financial (independent research) | August 2026

Abstract. This paper documents a complete research program for developing, validating, and capitalizing a single systematic intraday futures strategy under proprietary-firm (prop) account constraints. Across one week of concentrated research, 65 pre-registered hypothesis cells were tested against a falsification-first protocol: 54 alternative-hypothesis cells produced zero promotions, 11 control cells calibrated the base rates that deflated them, and the incumbent strategy survived every challenger on risk-adjusted efficiency. The validated strategy was then embedded in a three-layer Monte Carlo stack native to prop-account geometry: an absorbing-barrier model of the evaluation phase (43.3% pass probability, median 2-day verdict, \$267 expected cost to fund), a funded-phase policy optimization identifying an interior optimum in the cushion-and-sweep grid (\$7,638 expected extraction per account-year), and a multi-firm fleet simulation under perfect copy-trade correlation. The resulting instrument is a capped-loss structure whose 36-month success probability is an explicit engineering choice, rising from 58% to 92.5% as the risk budget and eval staggering are tuned, with maximum loss contractually bounded at the fee budget. Strategy identity, parameters, and signal logic are redacted throughout; every statistic, simulation, and governance mechanism is disclosed in full.

Note on redaction: this document discloses methodology, statistics, and capital architecture. Instrument identities, session definitions, filter thresholds, and entry/exit rules are withheld by design. Nothing here constitutes investment advice.

1. Introduction: optimizing for fewer false discoveries

Quantitative strategy research has a well-documented reproducibility problem: the large majority of backtested strategies fail on contact with live capital, with in-sample overfitting, multiple testing, and implementation gaps as the canonical causes. Retail research culture makes this worse by optimizing for discoveries. This program inverts the objective: the system is designed to minimize false discoveries, accepting that the modal outcome of any honest test is a rejection. The week documented here produced 54 consecutive rejections of alternative hypotheses and zero exceptions to pre-registered kill criteria, including for the most attractive results the program generated. The central claim is that this discipline, not any individual backtest, is the asset.

Four contributions follow. First, a governance protocol (pre-registration, family-level trial accounting, deflated performance bars, control cells) executed at retail scale. Second, dual-engine model validation, with an independent Python engine reconciled against the original platform implementation to within 1% of net P&L. Third, a simulation stack that treats prop-account rules as first-class objects: trailing floors as absorbing barriers, evaluations as options, fee budgets as hard loss caps. Fourth, a probability-engineering result: under copy-trade correlation of one, fleet success probability is governed by risk budget and time-axis staggering, and can be selected rather than hoped for.

2. Research governance

2.1 Pre-registration and kill criteria

Every hypothesis enters as a written cell before any result exists: thesis, exact configuration, evaluation windows, and the criteria that kill it. No parameter grids are permitted at intake; a cell is one configuration. Post-hoc rescue (re-filtering, sign-flipping, or window shopping after seeing a failure) is prohibited categorically. Failed ideas may be re-filed only as forward-tested paper hypotheses judged exclusively on data that did not exist at filing time.

2.2 Multiple-testing control

Trials are counted at family level and the admission bar rises with the count. Following the deflated Sharpe ratio methodology of Bailey and López de Prado, the expected maximum Sharpe under N independent trials, $E[\max \text{SR}] \approx$

$\sigma_{SR}[(1-\gamma)\Phi^{-1}(1-1/N) + \gamma\Phi^{-1}(1-1/(Ne))]$, is treated as the null a candidate must clear. At this program's terminal count (>60 family trials), a 29-trade cell printing a profit factor near 3 is expected by chance somewhere in the ledger and is therefore ineligible for direct promotion regardless of appearance; the program's single strongest cell was routed to forward paper testing on exactly this arithmetic.

2.3 The admission gate

A surviving candidate faces a three-hurdle portfolio gate against the incumbent book: (i) absolute daily correlation below 0.30, (ii) positive standalone expectancy, and (iii) an improvement in blended daily Sharpe. It must additionally remain accretive in a pre-specified lean window (the sample's weakest regime years) to strip out regime beta arriving in a strategy costume. Two cells passed the mechanical gate during the campaign; both were subsequently executed by the lean-window rule or by their own pre-registered control, and the gate's mechanical output was never permitted to outrank the written protocol.

Worked example of the bar in action: at N=15 family trials and ~5,400 daily observations, a stored deflated Sharpe of 0.974 recomputed to 0.957, and at the honest N=36 to 0.910, moving a legacy candidate from admitted to marginal before its evidence rows were even examined. Governance in numbers for the present campaign: 65 cells, 2 mechanical gate admissions, 0 promotions, 5 forward-log entries with dated future observations, and 1 pending pre-registered candidate holding a fixed verdict date. The asymmetry is the design: the ledger makes it easy to reject and expensive to admit.

2.4 Case study: the phantom survivor

Governance was stress-tested by auditing a legacy research ledger (174 runs) whose sole surviving strategy card showed a stage-3 pass. The audit found an identical-metrics kill row nine minutes earlier, an annotation freezing the card pending re-approval that never occurred, and a family-trial denominator understated by 3x. The family was closed at zero survivors and the verdict annotated immutably. The episode motivates the program's standing rule: a result selected after a search carries the search's inflation, and the ledger, not the researcher's memory, is the authority.

3. Data and engine validation

All research runs on exchange-grade continuous futures data (Databento, CME Globex MDP 3.0, volume-rolled continuous contracts, 2023-2026), resampled to the strategy's native bar size, with costs modeled at \$1.25 per side plus one tick of adverse slippage per fill. Because implementation spread across backtesting engines is a documented and material source of error, the strategy was re-implemented from scratch in an independent Python engine replicating the original platform's execution semantics (order-on-close processing, gap-through stop fills, identical indicator conventions). Reconciliation across 3.6 years:

Metric	Platform engine	Independent engine	Divergence
Trades (instrument A)	606	610	+0.7%
Net P&L; (instrument A)	\$137.3K	\$138.6K	+0.9%
Profit factor (A)	1.295	1.308	+0.013
Trades (instrument B)	500	491	-1.8%
Net P&L; (instrument B)	\$72.9K	\$68.9K	-5.5%

Figure/table: dual-engine reconciliation. Residual divergence traces to contract-roll handling and is carried as model risk rather than hidden.

4. The falsification campaign

The campaign asked one question at increasing generality: does any deployable edge exist for this engine outside the incumbent's home configuration? Sixty-five cells later the answer is a documented no, and the value is the map. Families tested: foreign-session breakout structures at two bar resolutions; the validated template transferred to instruments whose home sessions match those hours; cross-asset lead-lag (one instrument's session move trading another, four mechanisms including the best-documented equity spillover in the literature); prior-day-conditioned

session behavior; a flow-proxy test of the conditional overnight drift mechanism, with the unconditional control correctly reproducing the published near-zero result and thereby validating the pipeline; a twenty-strategy calendar and event catalog (turn-of-month, expiration structure, holiday, weekend, day-of-week, and macro-announcement cells); and a full exit-engineering program on the incumbent itself.

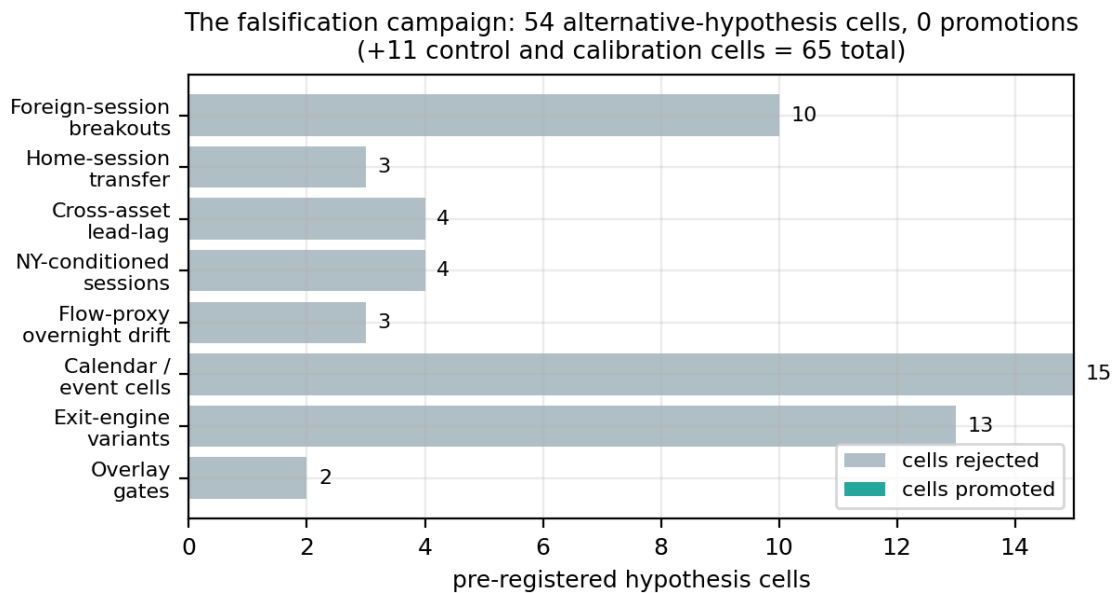


Figure 1. Cells by family. Every alternative-hypothesis cell was rejected on pre-registered criteria; 11 further cells were controls run to calibrate base rates.

Family	Cells	Strongest cell	Verdict and reason
Foreign-session breakouts (2 resolutions)	10	PF 0.88	Rejected: negative expectancy, cost floor
Classic overnight-range London bracket	2	PF 0.95	Rejected: matches published null on FX
Home-session template transfer (3 venues)	3	PF 0.82	Rejected: uniform non-transfer, PF 0.81-0.82
Cross-asset lead-lag (4 mechanisms)	4	PF 1.22	1 gate-admit, executed by lean-window rule
Prior-day-conditioned sessions	4	PF 1.20	Rejected: lean-window negative
Flow-proxy conditional overnight drift	3	PF 0.99	Null; control reproduced published near-zero
Session-open and fixing-flow clocks	7	PF 1.05	Rejected: expectancy, incl. last untested peer-reviewed mechanism
Calendar and event cells	12	PF 3.00	Best cell to forward log (n=29 at 60+ family trials); rest rejected by controls
Exit-engine variants on incumbent	13	ratio 5.6	All below incumbent efficiency 6.5; one durable finding (Sec. 6)
Overlay gates on overnight exposure	2	ratio 3.7	Both below pre-registered bars
Controls and calibration	5+	-	Instruments, not candidates (Sec. 4.2)

Table: the campaign ledger at family level. Full per-cell records (configuration, windows, printed kill reason, trade logs) are retained in the research ledger.

4.1 The kill taxonomy

Rejections clustered into four reusable diagnoses. **Beta in costume:** seven cells were a bull-market factor wearing a strategy's clothes, four smuggling a metals supercycle through a tilt, a currency signal, a fixing-flow story, and a calendar condition, and three repackaging overnight index drift; the lean-window rule executed each. **The friction floor:** short-horizon signals died wherever gross edge sat below the ~2-point round-trip cost floor, confirmed

independently on four instrument classes, replicating published falsification work. **Template non-transfer:** the validated structure exported to three foreign home sessions failed with near-identical profit factors (0.81-0.82), evidence the edge is venue-specific rather than structural. **Small-sample shine:** every glittering low-N cell was deflated by a matched control (an unconditional same-window base rate) or by the monthly/quarterly decomposition of its own condition.

4.2 Control cells as instruments

The program's strongest single result, a macro-announcement drift cell at profit factor 2.997 across 29 trades with an 8x per-trade excess over its 22-hour unconditional base rate, illustrates the machinery: its base rate was measured (not assumed), its family count made direct promotion statistically inadmissible, and it was routed to a dated forward paper log with its benchmark attached. The discipline is symmetric: controls also killed the program's most seductive calendar cell by decomposing its effect into a subset its own mechanism could not explain.

5. The incumbent strategy (redacted characterization)

The surviving strategy is a session-bounded intraday breakout system on two liquid CME futures (an equity index and a metals contract), holding minutes to hours, always flat by session end, with entries gated by a volatility-regime filter, a formation-range distributional ceiling, a participation (volume) confirmation, and a hard risk-to-range skip. Roughly 40% of sessions produce no trade by design.

Statistic	Instrument A	Instrument B	Combined book
Trades (2023-2026)	610	491	1,101
Net P&L; (1-lot, after costs)	\$138.6K	\$68.9K	\$207.5K
Profit factor	1.308	1.273	-
Expectancy (R)	0.137	0.13*	-
Max drawdown	27.1%	~25%	\$31.4K (daily)
Positive years	4 of 4	4 of 4	4 of 4
Daily mean / sd	-	-	\$292 / \$2,712

Table: incumbent performance, independent engine, \$1.25/side + 1 tick costs. *book-level R varies with the sizing convention; per-instrument figures at fixed risk.

Filter attribution quantified what the gates buy: removing them adds trades worth +\$15.3K over the full sample on the index leg, at the price of doubling max drawdown (to 40%) and, decisively, a -\$69.6K contribution from the skipped set in the current regime year. The filters are regime protection purchased with old-regime profits, and per the allocator literature on regime-dependent backtests, the program applies that discount to itself rather than reporting the unfiltered curve.

6. Exit engineering: a negative result with one durable exception

Thirteen exit-variant cells (bracket targets, breakeven ratchets, chandelier trails at two widths, hard stops at two widths, and gated overlays) produced a clean substitution theorem: entry filters and exit trails compete for the same job, damage control on hostile days, and the filter wins strictly, because a skipped trade pays no commission, no slippage, and no whipsaw. No trail variant beat the incumbent's return-to-drawdown efficiency (6.5) from either the filtered (best 5.6) or unfiltered (best 3.3) side. The exception is genuinely useful: on multi-hour overnight exposure, a wide disaster stop (4x ATR) preserved 97% of returns while cutting drawdown 19%, whereas a conventional 2x stop destroyed 29% of returns for an 11% drawdown improvement. Tight stops on overnight holds are a nightly insurance premium against a risk that streak-drawdowns never trigger; wide stops are nearly free tail protection. This finding is now a standing specification for any future overnight candidate.

Exit configuration	Net P&L;	Max DD	Return/DD	Verdict
Incumbent: stop-only, session-end flat (filtered)	\$138.6K	\$21.3K	6.5	champion, unchanged

Exit configuration	Net P&L;	Max DD	Return/DD	Verdict
Filtered + 2x ATR chandelier trail	\$133.5K	\$24.7K	5.4	higher PF, less money, more DD
Filtered + 3x ATR chandelier trail	\$128.5K	\$23.0K	5.6	rejected
Unfiltered, stop-only	\$153.9K	\$84.7K	1.8	undeployable
Unfiltered + 2x trail	\$130.5K	\$39.6K	3.3	trail tames DD, taxes edge
Unfiltered + 3x trail	\$140.1K	\$71.6K	2.0	rejected
22h overnight exposure, no stop (control)	\$353.7K	\$107.0K	3.3	beta yardstick
22h + 2x ATR hard stop	\$252.0K	\$96.3K	2.6	whipsaw tax quantified
22h + 4x ATR disaster stop	\$343.9K	\$86.7K	4.0	the durable finding

Table: exit-engineering battery (index leg shown). The efficiency column, net return per unit of maximum drawdown, is the promotion metric; profit factor is reported internally but never used for promotion because it rewards truncating the right tail.

7. Prop geometry I: the evaluation as an absorbing barrier

Prop evaluations invert normal strategy assessment: strategy drawdown does not exist, only the probability of touching a trailing floor exists. The account's end-of-day trailing maximum-loss rule is modeled as a ratcheting absorbing barrier, $\text{floor}_t = \max(\text{floor}_{t-1}, \min(\text{balance}_t - L, B_0))$, with L the trail width and the ratchet locking at the starting balance B_0 . A 10,000-path bootstrap of the strategy's joined daily P&L (IID daily sampling; autocorrelation risk carried in the assumptions register) yields the evaluation's full geometry:

Configuration	Pass	Median verdict	E[attempts]=1/p	E[cost to fund]
Incumbent book, campaign sizing	43.3%	2 days	2.31	\$267
Overnight sleeve (retired)	48.4%	31 days	2.07	\$239 + ~6 months
Both stacked	44.3%	4 days	2.26	\$261, worse tail

Table: evaluation Monte Carlo. Two independently built simulators agreed within one percentage point on the headline pass rate. The sleeve's higher pass rate is a trap: evaluations are options priced per attempt, and verdict speed, not pass rate, is the currency; the stacked configuration adds tail depth without adding pass probability.

Two geometric facts complete the evaluation picture. First, single-day absorption: the book's worst historical day is 3.8x the trail width, so the barrier can be crossed from anywhere in one step, and the pass probability is irreducibly a race between the drift reaching the target and a tail day reaching the floor, which is why verdicts resolve in a median of two days. Second, sizing dominance: an evaluation-sizing sweep found smaller sizing strictly worse at the fleet level, because slower verdicts burn calendar and per-month fees without proportionally reducing absorption risk. Fast, full-sized verdicts are optimal precisely because the evaluation is an option: the holder wants resolution, not longevity.

8. Prop geometry II: funded-phase policy optimization

A funded account under the same trailing rule is a positive-drift process trapped a cushion's width above a kill line. The design variables are the cushion (profit retained above the floor before withdrawals begin) and the sweep fraction (share of excess withdrawn each cycle). The objective is expected extraction per account-year, explicitly not survival: money parked inside a mortal account dies with it. Simulation of the policy grid (10,000 paths per cell) locates an interior optimum:

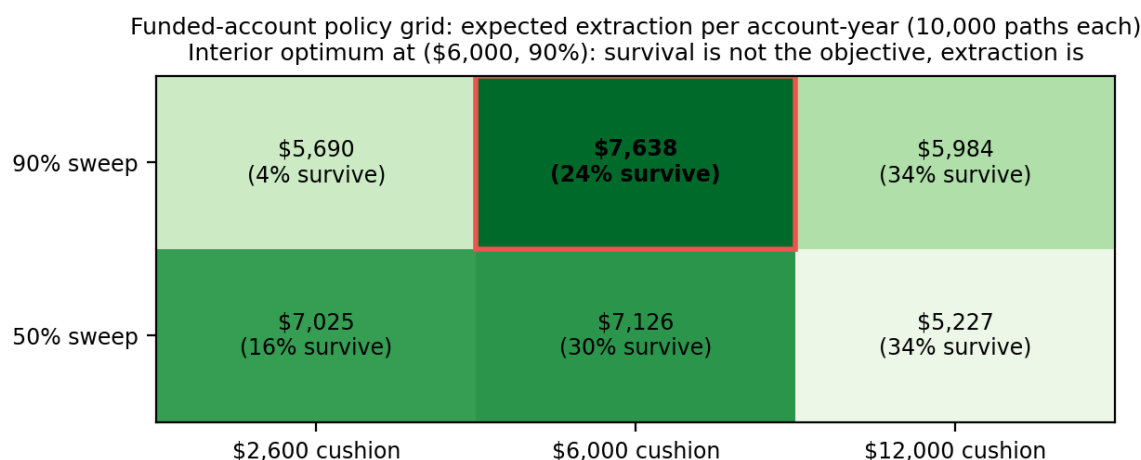


Figure 2. The fortress corner survives most and extracts least; the greedy corner dies before extracting. The optimum funds survival only to the point it services extraction. A second result: campaign (evaluation) sizing carried into the funded phase kills 100% of accounts within a week, quantifying why phase-scaled sizing is structural, not conservative.

8.5 Interlude: the same object outside prop constraints

The overnight-beta family rejected for prop deployment (its worst historical day alone equals a full trailing allowance at minimum size) is a coherent object in an unconstrained personal account, and the correct question there is allocation, not selection. Blending the incumbent book with k micro-sized units of the disaster-stopped overnight sleeve, using their empirical joint daily history (correlation +0.06), produces a concave efficiency curve with an interior optimum:

Sleeve units (micro)	Net (3.6y)	Max DD	Return/DD	Worst day
0 (incumbent alone)	\$207.5K	\$31.4K	6.61	-\$7.6K
1	\$241.9K	\$33.4K	7.24	-\$8.3K
2	\$276.3K	\$37.4K	7.39	-\$8.9K
3 (optimum)	\$310.6K	\$41.4K	7.50	-\$9.5K
4	\$345.0K	\$46.7K	7.39	-\$10.2K
5	\$379.4K	\$52.4K	7.25	-\$10.8K

Table: portfolio efficiency versus sleeve size. Near-zero daily correlation buys a 13.5% efficiency gain at $k=3$ before the sleeve's own drawdown dominates. The worst-day column is the honest footnote: tail correlation between two books on the same underlying is one, and grows linearly with k . Diversification improved the portfolio here in a way no exit modification improved either component, consistent with Section 6's substitution theorem: risk engineering belongs at the portfolio and entry layers, not the exit layer. This allocation is filed for unconstrained capital only and never enters the prop structure.

9. Prop geometry III: fleet architecture and probability engineering

9.1 Correlation one changes everything

Copy-traded accounts share every market day: with N accounts, portfolio variance scales as $N^2\sigma^2$, not $N\sigma^2$. Account count is therefore a leverage decision, never a diversification decision, and the fleet's bad worlds do not average away. The simulator enforces this by drawing one market day per step and applying it to every slot at its phase-appropriate scale.

9.2 The budget is an option premium

Total loss is structurally capped at the fee budget: evaluation fees, resets, and activations are the only capital ever at risk, and the simulator halts new spend at the cap. Small budgets produce a starkly bimodal outcome distribution: each path either escapes into a self-funding regime, where withdrawals finance unlimited retries, or exhausts the budget during early churn. Success probability is therefore a purchasable quantity, and two purchases dominate: budget size, and time-axis staggering of evaluation starts, which de-correlates fee burn across the only axis correlation-one permits.

Widening the stagger alone moved 36-month P(profit) from 58.1% to 73.8% at constant budget. A firm-level result also emerged: firms enforcing profit-consistency rules interact destructively with lumpy right-tail distributions (single days exceeding the evaluation target), doubling fee burn for a \$19K mean gain and an \$83K median loss in simulation, and were excluded from the final stack on that arithmetic.

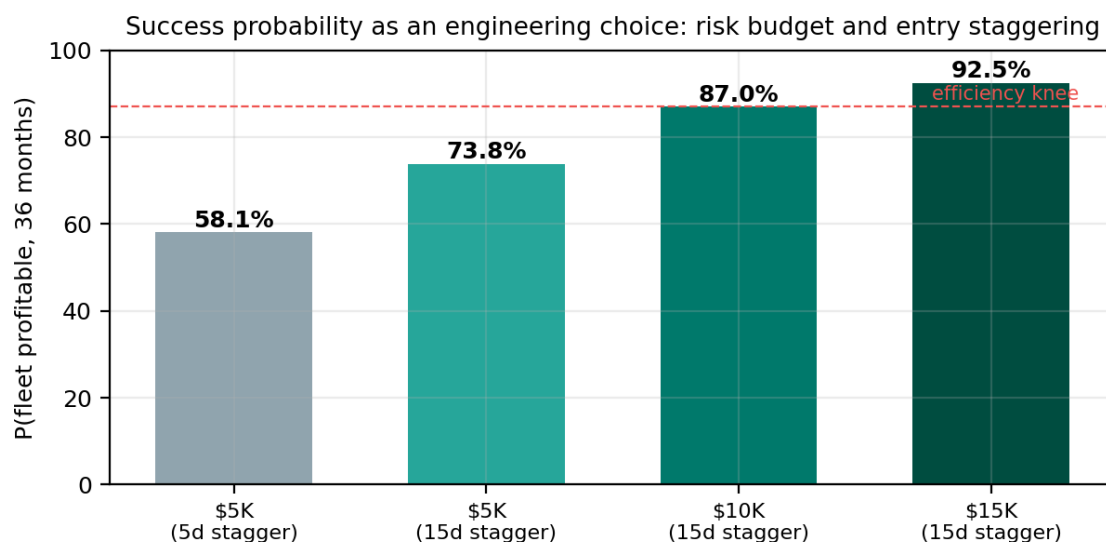


Figure 3. Engineered success probability across budget and stagger settings, 3,000-10,000 paths per configuration, 15-slot three-firm all-EOD-trailing stack, maximum loss identically equal to the budget in every configuration.

9.3 Headline projection under the final specification

Final specification: 15 copy-traded 50K slots across three EOD-trailing firms, full-size evaluations, 15-trading-day stagger, funded policy (\$6,000, 90%), budget released in tranches.

36-month projection (3,000 paths): mean net \$675K | median \$673K | p90 \$1.59M | steady-state mean \$18.1K/month | first payout median month 2.9 | P(profit) 74% at \$5K budget, 87% at \$10K, 92.5% at \$15K | **maximum possible loss = the budget, exactly.**

Conditioning assumptions, inseparable from the numbers above: live distribution matches the 2023-2026 backtest (the single untestable-in-simulation assumption; a live forward ledger with pre-registered acceptance gates exists to test it); IID daily sampling (real streaks are harsher); EOD trailing modeled throughout; per-firm terms-of-service on copy trading verified in writing before scale; copier and data fees excluded (~\$100-200/month at scale); the sample era is a favorable regime for the underlying instruments and the projection inherits that bias.

A projection stated without its conditioning assumptions in the same frame is marketing; with them, it is a forecast.

10. Risk decomposition and the live validation protocol

The residual failure probability decomposes into three components with disjoint levers. **Dead edge** (live distribution worse than backtest): untouchable by fleet engineering; the lever is sequencing, a single \$115 evaluation and a 20-signal forward ledger with pre-registered acceptance gates ($\geq 90\%$ signal-capture rate, live expectancy inside the backtest band) run before any scaled deployment, converting a \$5-15K bet into a \$115 test plus a conditional bet. **Churn variance** (real edge, budget exhausted first): fully engineered above via budget and stagger. **Operational and terms-of-service failure:** reduced to near zero by an automated execution relay (idempotent webhook processing, an account-level guard tracking the trailing floor, scheduled end-of-day flattening, kill-switch endpoints) and by written per-firm confirmations of copy-trading and automation policy. A structural upgrade, a pre-registered anti-correlated second engine, sits in the pipeline with a fixed verdict date and is the only element permitted to alter the fleet's composition.

The forward pipeline formalizes what rejected-but-interesting looks like: five hypotheses (a macro-announcement drift with an 8x measured excess over base rate, a quarterly-expiration overnight effect its monthly control could not explain

away, a pre-holiday effect matching four decades of literature at $n=29$, a regime-conditional exit variant, and a conditioned carry tilt) each hold a dated future observation window, a pre-attached benchmark, and a prohibition on further historical testing. Promotion, if it ever occurs, will be earned exclusively on data that did not exist at filing time, the only evidence class immune to the multiple-testing arithmetic of Section 2.

Limitations register, in one place: IID sampling understates losing streaks; one firm's intraday trailing variant is modeled optimistically as EOD; era bias is inherited from a favorable 2023-2026 sample; the strategy's live capture of modeled fills is unproven until the forward ledger says otherwise; and every projection in Section 9 is conditional on Section 10's gates clearing in order. The program's final claim is deliberately modest: the numbers are only as good as the discipline that produced them, and the discipline is the part that transfers.

Appendix A. Toolchain

Fourteen purpose-built research tools (Python) were produced during the campaign: a dual-validated backtesting engine with session-timezone and exit-variant parameterization; a cross-asset lead-lag tester; a prior-day-conditioned session tester; a calendar/event-conditioned tester with computed holiday, expiration, and turn-of-month calendars; a flow-proxy builder; a three-hurdle correlation gate; three Monte Carlo simulators (evaluation, funded-policy grid, multi-firm fleet with per-firm fee and consistency profiles); a warehouse forensic auditor with deflated-Sharpe recomputation; and an automated execution relay (FastAPI, idempotent, guard-gated, kill-switchable) with bridge-service fanout for multi-account execution. All simulation figures in this document are reproducible from logged trade records via deterministic seeds.

Appendix B. Simulation parameter register

Simulator	Key mechanics	Paths	Principal outputs
Evaluation MC	IID daily bootstrap of joined book; ratcheting EOD floor locking at start; pass at target, absorb at floor; no time limit	10,000	pass 43.3%, median verdict 2d, E[cost] \$267
Funded policy grid	same barrier from funded start; phase-scaled sizing; sweep of (cushion, fraction) every 21 trading days; 12-month horizon	10,000/cell	interior optimum (\$6,000, 90%), \$7,638/acct-yr
Fleet simulator	one shared market day per step (correlation 1); per-firm fee, reset, activation, consistency and split profiles; staggered eval starts; hard budget cap on new spend; 36-month horizon	3,000/config	P(profit) 74/87/92.5% at \$5/10/15K; mean \$675K; median first payout month 2.9
Deflation auditor	schema-discovering warehouse forensics; DSR recompute with corrected observation counts and family denominators; sensitivity over N in {5,15,36}	-	legacy family closed at 0 survivors

All simulators share the same empirical daily P&L pool and deterministic seeds; cross-simulator agreement on the evaluation pass rate (within one point across two independent builds) is treated as the stack's own reconciliation test, mirroring Section 3.

Contact: Ali Asfar - [linkedin.com/in/aliasfar7](https://www.linkedin.com/in/aliasfar7) - github.com/aliasfar7

Postscript: what generalizes

Stripped of its redacted specifics, the program reduces to eight transferable rules, each purchased with a documented failure or a documented control in the sections above.

Rule	Purchased by
Write the kill criteria before the run; the reframe you invent after seeing a failure is the overfit, not the insight	Secs. 2.1, 4.1: every post-hoc rescue path closed
Count your trials at family level and let the bar rise with the count	Sec. 2.2: PF 3.0 cell routed to paper at $N > 60$
Run the control before believing the condition	Sec. 4.2: base rates deflated every glittering cell
A result that only wins the holdout was selected by the holdout	Sec. 5-6: exit variant with holdout-only edge sent to forward log
Regime beta wears strategy costumes; test the lean window	Sec. 4.1: seven cells executed by one rule
Skipping a trade is the only exit that pays no friction	Sec. 6: the substitution theorem
Under binding barriers, optimize extraction, not survival; under correlation one, accounts are leverage, not diversification	Secs. 8-9: interior optimum; N-squared variance
The one assumption simulation cannot test is the one the live ledger exists for	Sec. 10: sequencing as the cheapest probability upgrade

The strategy documented here may or may not survive its forward ledger; the framework does not depend on the answer. A research process is working exactly when it is allowed to say no to its own best-looking result, and this document is the record of it doing so 65 times in one week.