MAD-Portfolio: Portfolio Construction from the Moving-Average-Distance Signal

Lawson Arrington[1]

July 18, 2026

Abstract—Three companion studies developed and validated a trading signal built on Moving Average Distance—the percentage displacement of price from its 20-day average, standardized by trailing volatility—evaluating it one security at a time on unconstrained capital. This paper addresses the portfolio problem that design leaves open. Applied to 728 point-in-time S&P 500 constituents over 2016–2025, the gated entry generates a median of twelve simultaneous candidacies per signal day and exceeds the capacity of a concentrated book on 95% of days, so realized performance is determined largely by selection among simultaneous candidacies. Selection rules are evaluated in three stages: a cross-sectional ranking study over 42,340 resolved signal trades; nested scoring models fit strictly walk-forward; and dollar simulations under realistic execution, judged against matched random-selection envelopes and against capitalization-weighted (SPY) and equal-weight (RSP) index controls, with January 2022–December 2025 reported separately as the primary out-of-sample evidence. Trailing returns exhibit no predictive content for signal outcomes at any horizon examined; a displacement-based filter previously used in live trading selects no better than random draws; and fitted models attain the highest trade-level information coefficients but trail the unfitted volume rank in dollar simulation at every book size. The most robust selector is daily trading volume (monthly-mean t of 4.2 and 5.5 on its two statistics): conditional on a signal firing, heavily traded names earn materially more per trade over comparable holding times. The resulting estimation-free rule—eligibility restricted to the 100 highest-volume names, candidates accepted in volume order, ten equal slots—compounds $100,000 to $552,888 over 2016–2025 (18.7% annualized, Sharpe 0.90) and is the only simulated configuration whose confirmation window exceeds its full window: 19.7% annualized, Sharpe 0.97, and a −19.1% maximum drawdown, versus 9.5% annualized for SPY and 4.2% for RSP on the same dividend-free basis. All results are walk-forward, and all headline numbers reproduce from fixed seeds.

[1] *Disclaimer: MAD-Portfolio is an independent, self-directed research project developed exclusively on my personal time. It is not affiliated with, sponsored by, or conducted on behalf of the Department of the Army, the Department of War, or any U.S. Government entity.*

I. Introduction

Three companion studies developed and evaluated a trading signal built on Moving Average Distance (MAD), the percentage displacement of price from its 20-day average standardized by its own trailing volatility. Position within the displacement band forecasts the direction of the next regime transition under strictly walk-forward estimation [1]; a first-order decomposition of displacement changes into price-driven and moving-average-driven components converts that forecast into an entry rule that survives paired significance testing across the index cross-section [2]; and the joint displacement–velocity dynamics are stationary and shared across the index [3]. All three evaluations test each security independently, with unconstrained capital and every signal taken—a design that isolates the signal's properties and leaves the portfolio problem unaddressed.

The gap is not small. Applied to 728 point-in-time S&P 500 constituents over 2016–2025, the gated entry generates a median of twelve simultaneous candidacies per signal day, and on 95% of days

with any candidate the number of firings exceeds the free capacity of a concentrated book. The resulting allocation problem differs from classical portfolio selection [4] in two respects: the opportunity set is a stream of short-lived candidacies rather than a static universe, and the decision is which candidacies to accept rather than how to weight continuing holdings. Selection, not weighting, is the central margin.

This paper evaluates selection rules in three stages, with the statistical search confined to the earliest stage. First, a daily cross-sectional ranking study measures, over 42,340 resolved signal trades, whether candidate outcomes are predicted by trading volume, by each security's own history of converting signals into profit, or by trailing returns, each at multiple horizons. Second, nested linear scoring models are fit strictly walk-forward to test whether combining predictors improves on the best single predictor. Third, the surviving rules are evaluated in dollar simulations of a realistic account against three controls: matched random-selection envelopes, the capitalization-weighted index (SPY), and the equal-weight index (RSP). Every simulated configuration reports the January 2022–December 2025 sub-period separately; because the simulated menu was informed by the full sample, this sub-period is the primary out-of-sample evidence.

The results are asymmetric. Trailing-return rankings exhibit no predictive content for signal outcomes at any horizon between 5 and 80 days. A displacement-based filter previously applied in live trading selects no better than random draws from the same candidate pool. Fitted models achieve the highest trade-level information coefficients but trail the unfitted volume rank in dollar simulation at every book size. The most robust single selector is daily trading volume (monthly-mean t of 4.2 and 5.5 on its two statistics): conditional on a signal firing, candidates on heavily traded names earn materially more over comparable holding times, a conditional counterpart to the unconditional volume and liquidity effects [5], [6]. The final configuration—eligibility restricted to the 100 highest-volume names, candidates accepted in volume order, ten equal slots—compounds $100,000 to $552,888 over 2016–2025 (18.7% annualized) and is the only simulated configuration that performs better in the confirmation window than in the full window: 19.7% annualized against 9.5% for SPY on the same dividend-free basis, with a shallower maximum drawdown. The ordering matches prior evidence. Optimized portfolio rules routinely fail to outperform naive alternatives out of sample [7].

The paper proceeds as follows. Section II reviews related work. Section III restates the signal and the account rules inherited from the companions. Section IV describes the data and the walk-forward protocol. Section V reports the ranking study, Section VI the scoring models, and Section VII the simulations. Section VIII discusses limitations, and Section IX concludes.

II. Background and Related Work

A. Portfolio construction from signals

The classical treatment allocates weights over a fixed menu of assets given estimated moments [4]. A large subsequent literature documents how badly estimation error degrades that program in practice: no optimized rule reliably outperforms equal weighting out of sample [7]. The present

setting sharpens the issue. When the opportunity set is a daily stream of expiring candidacies, weights collapse to a sequence of accept-or-decline decisions under a slot constraint, and the quantity being estimated is not a covariance matrix but a ranking. The out-of-sample failure of fitted rules documented below is therefore a selection-problem analogue of a familiar allocation-problem result.

B. Volume and liquidity in the cross-section

Trading activity predicts returns unconditionally: individual stocks experiencing abnormally high volume earn excess returns over subsequent weeks [5], while in levels, illiquid stocks carry a return premium as compensation for trading costs [6], and turnover relates negatively to average returns [8]. The finding reported here is distinct from, and not in tension with, that literature: it is conditional on a mean-reversion signal having fired. Among simultaneous candidates, the heavily traded names—which the unconditional literature associates with lower required returns—are precisely the ones whose gated re-crossings pay best. Liquidity here functions as a quality filter on the signal rather than as a priced characteristic; relatedly, momentum profits are known to survive trading costs only in implementations tilted toward liquid names [9]. Restricting eligibility by volume also addresses implementability directly, since the resulting book holds names whose capacity dwarfs the simulated account.

C. Model complexity and evaluation discipline

With enough rules examined, some succeed by chance [10], and backtest overfitting inflates expected out-of-sample performance [11]. The design here responds in three ways. The predictor families and horizons were fixed before estimation; the simulations run a short pre-registered menu rather than a search, and are judged against random-selection envelopes constructed under identical mechanics, so that every reported percentile is a comparison against luck rather than against zero; and the 2022–2026 sub-period is reported for every configuration, with the paper's conclusions resting on it. The companion series' practice of reporting failed hypotheses alongside successful ones [1], [2], [3] is retained.

III. The Signal and the Account Rules

Everything this paper selects among is inherited, unchanged, from the companions [1], [2],

A. Construction

Following [1], let P denote the closing price. The trend reference is the w-day simple moving average, with w = 20 throughout:

$$\mathrm{SMA}_t = \frac{1}{w}\sum_{i=0}^{w-1} P_{t-i} \tag{1}$$

Moving Average Distance is the signed percentage displacement of price from that reference,

$$\mathrm{MAD}_t = 100 \cdot \frac{P_t - \mathrm{SMA}_t}{P_t} \tag{2}$$

positive when price trades above its average and negative when below. Because a fixed displacement is not equally meaningful across volatility regimes, MAD is standardized by its own trailing variability, the sample standard deviation over a window of W = 255 trading days,

$$\sigma_t = \text{std}(\text{MAD}_{t-W+1},\ \ldots,\ \text{MAD}_t) \tag{3}$$

and the standardized displacement

$$z_t = \frac{\text{MAD}_t}{\sigma_t} \tag{4}$$

expresses displacement in units of its trailing standard deviation, comparable across time and across assets. Every quantity in (1)–(4) is strictly trailing, so a value computed at one day's close is executable at the next day's open with no revision and no look-ahead. The constructions carry a warm-up of roughly one trading year, and every test in this paper begins only after it.

B. Regime classification

The standardized series is discretized into six ordered regimes at thresholds 0, ±1, and ±2, with no neutral state, so that every observation carries both a direction and a magnitude [1]:

Table I. Regime definitions.

Regime	Condition	Interpretation
+3	$z > 2$	extended above trend
+2	$1 < z \le 2$	above trend
+1	$0 < z \le 1$	just above trend (P > SMA)
−1	$-1 < z \le 0$	just below trend (P ≤ SMA)
−2	$-2 < z \le -1$	below trend
−3	$z \le -2$	deeply oversold

The two transitions this paper trades are defined on this state sequence: the entry references the crossing from −1 to +1, and the exit references the rollback from +2 to +1.

C. The velocity gate

A regime transition is, mechanically, a change in MAD, and that change is not a single thing. A first-order expansion over k = 5 trading days separates it into a price term and a moving-average term [2]:

$$\Delta\text{MAD}_t \approx \left(\frac{100 \cdot \text{SMA}_t}{{P_t}^2}\right) \cdot \Delta P_t \ - \ \left(\frac{100}{P_t}\right) \cdot \Delta\text{SMA}_t \tag{5}$$

where $\Delta X_t = X_t - X_{t-k}$. The first term is positive when price rises—displacement improving because the security is genuinely recovering. The second is positive when the moving average falls, typically because old highs are leaving the w-day window, so displacement can improve while price is flat or still declining. The signed share of the change attributable to price is

$$s_t = \frac{c_t^P}{|c_t^P| + |c_t^{MA}|} \tag{6}$$

where the numerator is the price term of (5) and the denominator sums the magnitudes of both terms. The share lies in [−1, +1]: values near +1 mark a price-dominated move, values near −1 a

moving-average artifact. The companion paper established that s functions as a qualifier on regime transitions, and that entries gated on it survive paired significance testing across the cross-section [2].

D. The entry, the exit, and candidacy

A security becomes a candidate when, at a close, its regime crosses from −1 to +1 and the concurrent share satisfies $s \geq 0.50$: the displacement is improving because price is moving, not because the average is catching up. Candidacy lasts exactly one day—the position is opened at the next session's open or not at all—and a security that did not trade the immediately preceding session is not a candidate, so a signal cannot survive a halt. The exit is the companion's only exit: when a held name's regime rolls back from +2 to +1 at a close, the position is sold at the next open. There is no stop-loss, no profit target, and no time limit. Three forced closures complete the rule set: a name removed from the index is sold at that day's open; a data splice closes the position at the prior close; a security whose series ends is closed at its final bar.

E. The account

Simulated capital starts at $100,000 in an account of N equal slots. A new position is sized at one-Nth of total account equity marked at its entry open, subject to available cash—the account never borrows—and then rides untouched until its exit: no rebalancing between positions, no averaging, no partial exits. All fills are market-on-open; every fill pays 5 basis points of trade value; shares are fractional; idle cash earns nothing. The book is long-only and unlevered. Prices are split-adjusted but dividend-free, a limitation of the archive; the index controls in Section VII are measured on the identical basis, so comparisons are internally consistent while all absolute levels understate total returns.

IV. Data and the Walk-Forward Protocol

A. Universe and prices

The universe is every security that was an S&P 500 constituent at any point between January 2016 and December 2025, as recorded in a point-in-time membership history [12]. Membership is reconstructed day by day from index snapshots: 728 distinct tickers touch the index during the window—between 501 and 507 on any given day, roughly 525 in any calendar year, with 19 to 29 additions and removals annually—and the set includes names subsequently acquired, delisted, or removed, so the tests are free of survivorship bias. Every layer of the study is gated on that day-by-day reconstruction: a security generates candidacies only on days it is a constituent, each evening's ranking universe is that evening's constituents, bench eligibility in Section VII requires current membership, and a held position is force-sold at the open of its first non-member day. A 2016 ranking therefore never sees a 2021 addition, and nothing entering the index mid-window carries history into rankings from before its add date: Tesla's first ranking day is its official addition date, December 21, 2020, on which it immediately ranks second by volume—the clearest possible evidence that high-volume non-members cannot contaminate the volume rankings. Spot checks of the reconstruction against officially announced effective dates match exactly for most additions and removals;

occasional boundaries are dated within one snapshot gap of the official date (Uber enters roughly two months early), an imprecision of the source history disclosed rather than repaired. Daily bars are drawn from Polygon.io, split-adjusted but without dividend reinvestment, and pinned to a fixed vintage (July 16, 2026) so that every number in this paper reproduces exactly. A security is tradeable only while a member; membership entry and exit dates are enforced daily. Archive artifacts receive explicit treatment: a bar whose one-day price ratio exceeds 5× in either direction, or that follows a gap of more than thirty calendar days between consecutive bars, marks a symbol-reuse splice or an era break, not a price path; the return chain is broken at such bars, positions held into one are closed at the last real close, capital idled by a vanished series is released after ten sessions, and every occurrence is logged. Those found include the 2018 Priceline–Booking rename and a 2020 spin-off adjustment; left untreated, the largest would have contributed a spurious 96,700% single-day return to an equal-weight benchmark.

B. The signal ledger

Applying the gate of Section III to the universe produces 42,340 candidacies over 2,514 trading days—the signal ledger on which the ranking study and the scoring models operate. Each candidacy is resolved independently through the full trade mechanics—next-open entry, rollback exit, forced closures—to an outcome: trade return, days held, and return per day of committed capital, the exposure-adjusted quantity this series treats as primary [2]. An outcome enters any subsequent computation only from the day its exit is knowable: a rollback exit on its execution day, and a delisting only after a recognition lag, since a final bar is not known to be final in real time. The ledger's composition testifies that the rule, not the archive, does the closing: 41,233 of the 42,340 candidacies (97.4%) resolve at the rollback exit, with 961 closed by index removal, 137 by delisting, and 9 at archive breaks. The median candidacy holds 24 trading days, wins 71.8% of the time, and returns +2.5% at the median. Because a security can fire again while an earlier candidacy is still open, overlapping candidacies on one name share an exit—82% of resolved rows share theirs with at least one other—and enter the ledger as correlated observations, the clustering the conventions of the next subsection exist to respect. Candidacies still open at the vintage date never enter estimation.

C. The walk-forward protocol and statistical conventions

Everything estimated in this paper follows one protocol. The daily ranking table of Section V is recomputed at each close from strictly trailing data—a stock's rank on a given evening uses only bars and resolved trades available at that close and is therefore actionable at the next open. The scoring models of Section VI refit on the first trading day of each month, training only on trades resolved before that day, and score the coming month's candidacies out of sample; models publish no score until 300 resolved trades exist, features with missing values impute to training-pool means and standardize per refit, and coverage after the burn-in is 98% of candidacies. Cross-sectional predictive claims are assessed uniformly: a statistic is computed within each day's cohort of simultaneous candidates, cohort statistics are averaged within calendar months, and t-statistics are computed on the monthly means—a convention that respects the strong cross-sectional and temporal clustering of signal outcomes, which pooled counts would overstate. Dollar simulations in

Section VII estimate nothing: their rules consume only the published ranks and scores, and their significance is assessed against random-selection envelopes run under identical mechanics with fixed seeds. The separation of roles across the three stages is deliberate: hypotheses may be found only in the ranking study, the models test combination, and the simulations confirm—a rule altered after seeing simulation output would forfeit the confirmation claim.

V. The Ranking Study

A. The instrument

The study's first instrument is a daily ranking table: every constituent, every trading day, ranked independently (rank 1 best, ties sharing the better rank) within that day's membership on twelve measures in three families. The VOLUME family is a single column, raw share volume that day. The SIGNAL-TO-PROFIT family ranks each security's mean trade return over its own past k completed signal trades, $k \in \{3, 5, 10, 15, 20\}$; a security with fewer than k resolved trades is unranked, and an outcome enters its history only from the day its exit is knowable. The RETURNS family ranks trailing close-to-close returns over $k \in \{5, 10, 15, 20, 40, 80\}$ of the security's own bars. Coverage differs by construction: volume and return ranks exist essentially always, while signal-to-profit coverage falls from 92% of candidacies at k = 3 to 67% at k = 20 as the history requirement deepens.

B. Evaluation frame

Each of the 42,340 candidacies is joined to the ranking table at its signal close—the table a trader would have held the evening before entry—and its outcome is the realized return per day of committed capital. Two statistics are computed per family within each day's cohort of five or more simultaneous candidates: a Spearman rank correlation between rank and outcome, signed so that positive values mean better-ranked candidates earned more; and a top-3 edge, the mean outcome of the three best-ranked candidates minus the cohort mean—the statistic closest to what a capacity-constrained book actually does. Both are averaged within calendar months, and t-statistics are computed on the monthly means.

C. Results

Table II reports the twelve families. The pattern is stark.

Table II. Rank families versus signal outcomes: within-cohort Spearman correlation and top-3 edge (bps per day held), t-statistics on monthly means, 2016–2025.

Family	Coverage %	IC	t (IC)	Top-3 edge	t (edge)
Volume	100.0	+0.034	4.15	+2.63	5.51
Signal-to-profit, k=3	92.5	+0.010	1.25	+1.37	2.74
Signal-to-profit, k=5	89.0	+0.007	0.72	+1.64	3.64
Signal-to-profit, k=10	81.5	+0.008	0.91	+1.30	2.54
Signal-to-profit, k=15	74.3	+0.005	0.25	+1.05	2.02
Signal-to-profit, k=20	67.3	+0.008	1.05	+1.37	2.73
Returns, k=5	100.0	+0.000	0.10	+0.82	0.78
Returns, k=10	100.0	+0.001	0.26	+0.35	0.82
Returns, k=15	100.0	−0.001	−0.43	+0.83	1.69
Returns, k=20	100.0	−0.000	−0.15	+0.26	0.79
Returns, k=40	100.0	+0.008	0.16	+0.83	1.44
Returns, k=80	99.9	−0.011	−1.34	+0.46	0.60

Volume dominates. Its information coefficient (t = 4.15) and top-3 edge (+2.63 basis points per day, t = 5.51) are the strongest of any single measure examined in this research program; the effect is monotone across the cross-section—candidates in the top quartile of same-day volume rank average 19.2 basis points per day held against 15.7 for the bottom quartile—and robust to excluding 2020–2021, where the edge is +2.5 (t = 5.0). It is a return effect, not a turnover effect: top- and bottom-quartile holding times are indistinguishable (35.3 versus 36.5 trading days on average), while mean trade returns differ by nearly a factor of two (+3.0% versus +1.7%). The decomposition is sharper still: win rates are indistinguishable across quartiles (71.8% against 72.2%), and the entire premium sits in the size of the winners—the average winning trade on a top-quartile name earns +7.6% against +5.8%—a magnitude effect, not a frequency effect. It is also persistent: the top-3 edge is positive in all ten calendar years, ranging from +1.5 to +3.7 basis points per day, and in 68% of the 120 months. Within the hundred most-traded names themselves, however, the residual volume ordering is directional but unproven (+0.8, t = 0.7): the effect lives at the boundary between liquid and illiquid, not among the most liquid—a shape Section VII's decomposition recovers independently. Signal-to-profit history is real but modest: every window's top-3 edge is positive, the strongest at k = 5 (t = 3.6), but combining it with volume adds nothing—and not because the two overlap: names with strong signal histories sit at the median of the volume spectrum, and the weaker ranking simply dilutes the stronger when averaged into it. Trailing returns carry nothing at any horizon: the largest magnitude among six t-statistics is 1.78, and signs alternate. Momentum, at least as expressed through a stock's own recent performance, does not identify which mean-reversion signals pay.

Two intuition-based selectors complete the negative results. Ranking by the displacement magnitude at the signal—the preference for shallow entries that the author applied in live trading—produces a top-3 edge of −0.6 basis points per day (t = −1.2): indistinguishable from, and pointed against, chance. Ranking by the velocity share s itself, the quantity that gates entry, produces −0.4 (t = −0.8): the gate's magnitude carries no ranking information beyond passing the gate, a finding the scoring models of Section VI confirm independently.

VI. The Scoring Models

A. Design

The ranking study evaluates one measure at a time. The natural next question is whether a fitted combination improves on the best single measure, and the test is designed to answer it without contaminating the answer: three nested linear models are fit side by side under one protocol. Model R uses only the ranking table, with each rank converted to a daily percentile: volume and the five signal-to-profit windows. Model S uses only the signal's own state at the signal close: the share s, MAD, z, five-day relative volume, and the 20-day trailing return. Model RS uses both sets. Each is a ridge regression on the outcome—return per day held, winsorized at ±250 basis points—refit on the first trading day of each month using only trades resolved before that day, with features standardized and missing values imputed to training-pool means per refit. No model publishes a score until 300

resolved trades exist; coverage thereafter is 98% of candidacies. Every score is therefore out of sample with respect to the trade it scores.

B. Trade-level results

Table III evaluates the three models in the frame of Section V.

Table III. Nested scoring models versus signal outcomes; the volume percentile alone is shown for reference. Same frame and conventions as Table II.

Model	IC	t (IC)	Top-3 edge	t (edge)
R (ranks only)	+0.020	2.30	+1.31	2.45
S (signal state)	+0.052	4.23	+3.44	3.81
RS (both)	+0.050	4.34	+3.38	3.69
volume percentile, unfitted	+0.034	4.15	+2.63	5.51

Two results deserve emphasis. First, fitting can subtract. Model R, handed the ranking table's information, delivers a top-3 edge of +1.31—half of what its own strongest input earns unfitted in the identical frame (+2.63, t = 5.51). The final refit shows why: the volume percentile receives a correctly signed weight (−1.42; lower percentile is better), but the five signal-to-profit percentiles—overlapping windows of the same underlying history, individually weak and mutually correlated—receive alternating signs (−1.00, +0.38, +0.76, +0.17, −0.57), the signature of a regression chasing in-pool noise across collinear inputs. The estimation itself, not the information, is the liability [7], [11]. Second, combining does not add: model RS is statistically indistinguishable from model S, so the ranking table contributes nothing beyond what the signal state already carries at the trade level.

C. What the fitted model learns

Model S's final refit is interpretable and confirms Section V from an independent direction. Its weight concentrates in an opposed pair: raw displacement MAD at +7.27 and standardized displacement z at −7.48, near-equal and opposite. Jointly they select candidates whose displacement is large in raw percentage terms but small relative to their own volatility—volatile names early in their standardized move. The share s, which gates every candidacy, receives −0.38: within the gated pool, the gate's own magnitude adds nothing, matching the ranking result. Relative volume and trailing return receive weights indistinguishable from zero. Model S is the statistical winner of this section; whether its advantage survives realistic implementation—costs, slots, cash, and one realized path rather than an average over cohorts—is the question Section VII answers, in the negative.

VII. The Simulations

A. The menu and its controls

The simulations run a fixed menu: four pickers—the volume rank used directly (RANKVOL) and the three fitted scores of Section VI—crossed with book sizes $N \in \{5, 10, 20\}$, plus, at N = 10, a bench variant of each picker in which eligibility is restricted to the 100 highest-volume names as of the prior month's last ranking table. Every run uses the account of Section III.E. Three controls discipline the results. First, random-selection envelopes: 200 books per book size (and 200 more under the bench restriction) that accept candidates at random under identical mechanics, seeds fixed; each menu run is placed within its matching envelope as an exact Monte Carlo percentile, so every comparison

is against luck operating under the same rules. Second and third, buy-and-hold positions in SPY and in RSP, the equal-weight S&P 500 fund—the fairer of the two for an equal-slot book—entered at the window's first open and measured on the same dividend-free basis. The 2022–2025 sub-period is reported for every run and carries the evidentiary weight.

B. Results

Table IV reports every run.

Table IV. All simulated configurations and controls: $100,000 initial capital, 5 bps per side, January 2016–December 2025, with the January 2022–December 2025 sub-period reported separately. Envelope percentile is the exact Monte Carlo placement within the matching 200-seed random-selection envelope.

Run	Final $	CAGR	Sharpe	MaxDD	2022+ CAGR	2022+ Sharpe	2022+ MaxDD	Env. pctile
SPY	339,957	13.1%	0.77	−34.1%	9.5%	0.59	−25.4%	—
RSP	253,931	9.8%	0.60	−39.6%	4.2%	0.33	−22.5%	—
RANKVOL, N=5	651,013	20.7%	0.95	−29.8%	17.8%	0.81	−24.5%	100.0
RANKVOL, N=10	426,302	15.6%	0.81	−31.8%	9.0%	0.52	−21.3%	99.0
RANKVOL, N=20	385,276	14.5%	0.80	−34.0%	7.6%	0.50	−19.6%	99.5
Score S, N=5	383,310	14.4%	0.66	−36.2%	12.5%	0.62	−33.3%	93.0
Score S, N=10	280,111	10.9%	0.57	−35.9%	5.2%	0.35	−26.6%	72.0
Score S, N=20	234,371	8.9%	0.51	−37.7%	5.3%	0.36	−19.8%	40.5
Score R, N=5	486,059	17.2%	0.83	−39.9%	7.6%	0.45	−22.7%	98.0
Score R, N=10	318,217	12.3%	0.69	−36.0%	6.4%	0.43	−18.4%	88.0
Score R, N=20	239,510	9.2%	0.56	−37.8%	1.5%	0.17	−21.4%	45.5
Score RS, N=5	176,209	5.8%	0.35	−58.3%	3.2%	0.25	−36.3%	23.5
Score RS, N=10	185,874	6.4%	0.39	−54.8%	2.3%	0.21	−28.7%	15.0
Score RS, N=20	213,931	7.9%	0.47	−43.2%	3.8%	0.29	−21.1%	21.5
RANKVOL, bench, N=10	552,888	18.7%	0.90	−38.7%	19.7%	0.97	−19.1%	88.5
Score S, bench, N=10	405,671	15.1%	0.76	−32.7%	11.9%	0.62	−19.9%	39.5
Score R, bench, N=10	360,648	13.7%	0.72	−38.4%	10.7%	0.60	−20.0%	17.0
Score RS, bench, N=10	346,002	13.2%	0.66	−46.1%	12.2%	0.62	−20.3%	14.0

Three regularities organize the table. First, the volume rank dominates its envelope wherever selection bites: RANKVOL places no lower than the 99th percentile of 200 random books at every book size, and at N = 5 it beats all 200 (p = 0.01, the envelope's floor); no fitted score matches it. Second, the trade-level ordering of Section VI inverts in dollars: the fitted models trail the volume rank at the same book size in every window, and their confirmation-window results are uneven—from 1.5% to 12.5% annualized across the nine score books, against 7.6–19.7% for the volume rules. A rule that must be re-estimated is exposed to every way the future differs from the pool it was fit on; a rank is not. Third, the bench raises every picker's confirmation-window return, with confirmation-window drawdowns uniformly shallower than the unbenched counterparts'.

The strongest configuration combines both volume effects: RANKVOL on the bench at N = 10 compounds $100,000 to $552,888 (18.7% annualized, Sharpe 0.90) and is the only configuration in the menu whose confirmation window is stronger than its full window: 19.7% annualized, Sharpe 0.97, maximum drawdown −19.1%, against 9.5%, 0.59, and −25.4% for SPY and 4.2%, 0.33, and −22.5% for RSP. One honesty note: its full-window drawdown (−38.7%) is deeper than the index's, a cost of concentration paid in the 2020 drawdown; its confirmation-window drawdown is materially shallower. Its edge decomposes cleanly. Eligibility alone is the larger part: the bench-constrained

random envelope's median final equity is $435,913 against $242,100 unrestricted—restricting the pool to liquid names raises the median random book by four-fifths, and the winner beats 100% of the unrestricted envelope. Ranking within the bench contributes the remainder—19.7% against the benched-random median of 10.5% in the confirmation window—but places at the 88.5th percentile of its own envelope (p = 0.24): directional, not individually significant. The honest decomposition is that eligibility is the statistically overwhelming component and the within-bench ranking a consistent but unproven increment.

C. The confirmation window, year by year

Table V reports calendar-year returns for the winning configuration against both controls.

Table V. Calendar-year returns, dividend-free basis, 2016–2025.

Year	RANKVOL bench N=10	SPY	RSP
2016	+22.8%	+11.4%	+14.8%
2017	+16.3%	+19.4%	+16.6%
2018	+3.2%	−6.3%	−9.5%
2019	+36.1%	+28.8%	+26.6%
2020	+8.2%	+16.2%	+10.2%
2021	+24.4%	+27.0%	+27.6%
2022	−3.4%	−19.5%	−13.2%
2023	+38.1%	+24.3%	+11.7%
2024	+14.3%	+23.3%	+11.0%
2025	+34.3%	+16.4%	+9.3%

The pattern is asymmetric participation. In both down years of the window the strategy is the best performer by a wide margin: +3.2% in 2018 against −6.3% for SPY and −9.5% for RSP, and −3.4% in 2022 against −19.5% and −13.2%. The mechanism is structural rather than forecast-based: the exit rule liquidates positions as displacements roll over, so sustained declines find the book increasingly in cash, and the bench confines what it does hold to liquid mega-capitalization names. The 2022 gap is where the confirmation window's shallower drawdown and higher Sharpe originate. The cost is participation lag in melt-up years led by names the rule exits early—2020, 2021, and 2024 are the strategy's weakest relative years—and the net of the two asymmetries over the decade is the 18.7% against 13.1% recorded in Table IV.

VIII. Discussion and Limitations

A. What multiplicity remains

The statistical search was confined to the ranking study, but the simulation menu it informed still contains sixteen configurations, and the envelope placements in Table IV are not corrected for that multiplicity: a Bonferroni adjustment across the menu would leave no single placement significant at conventional levels, and the 200-seed envelopes cannot resolve p below 0.01 in any case. The paper's claim therefore does not rest on any one percentile. It rests on consistency across frames that multiplicity does not link: volume predicts outcomes at the trade level (t = 4.15 and 5.51 on two statistics), monotonically across quartiles, excluding the pandemic years, at every book size, in the full window and in the confirmation window—and the configuration that combines its two expressions is the only configuration in the menu that strengthens out of sample at all. A false

discovery can survive one of those tests by luck; surviving all of them requires the luck to be coordinated.

B. Dividends, costs, and taxes

The dividend convention favors the strategy, slightly. All series are measured without dividends. The controls are always invested and forgo roughly the index yield—about 1.5% annually—while the strategy, approximately 97% deployed on average, forgoes almost as much. The comparison in Table IV is therefore biased in the strategy's favor by the exposure difference times the yield, on the order of a tenth of a percentage point annually: real, disclosed, and far smaller than the reported gaps. Taxes are outside scope entirely; at roughly seventy round trips per year the strategy realizes short-term gains, and taxable deployment would meaningfully narrow its after-tax advantage over an unrealized buy-and-hold position.

C. One decade, one index, one mechanism

The membership archive begins in 2016, so every result lives in a single decade containing two drawdowns—2018 and 2022—both of which the strategy weathered well, and no episode resembling 2008's liquidity collapse, in which the liquid-bench premise itself might fail. The out-of-sample superiority concentrates in 2022, 2023, and 2025; the mechanism is structural—rollback exits cash the book as displacements deteriorate, and the bench confines holdings to names that can be exited at will—but a structural mechanism demonstrated in one bear market is a hypothesis about the next one, not a guarantee. Relatedly, the paper documents the conditional volume effect without explaining it. The direction opposes the unconditional liquidity premium [6], which suggests the effect is a property of the signal's interaction with liquid names—faster information incorporation producing cleaner re-crossings is one candidate mechanism—but no test here distinguishes candidates, and an undocumented mechanism can decay without warning.

D. The scope of the model result

The fitted models that failed out of sample were ridge regressions over modest feature sets, refit monthly—the complexity class a careful practitioner reaches for first, not the frontier. The result licenses a comparative claim, not a universal one: within this signal stream, at this sample size, estimation added fragility rather than edge, and the unfitted rank captured most of what the best model knew. Whether richer function classes, longer histories, or regime-conditioned refits reverse that ordering is untested. Finally, execution is idealized at the margin that matters least for liquid names—market-on-open fills at 5 basis points per side—and at the $100,000 scale simulated, capacity is not a binding concern; neither claim extends to institutional size without a market-impact model this paper does not contain.

IX. Conclusion

The companion studies built a signal; this paper built the portfolio the signal implies, and found that nearly everything of consequence happens in the step between them. The gated entry fires far more often than a concentrated book can act, so realized performance is set by the answer to one question—which candidacies deserve capital—and the answer that survives every frame this paper

can construct is the least sophisticated one available. Trailing returns predict nothing. A filter the author traded by conviction selects no better than chance. Fitted models win the statistical contest and lose the economic one. Daily trading volume—one column, no estimation—predicts trade outcomes with the strongest and most stable statistics in this research program, and expressing it twice, as an eligibility bench and as the acceptance order, produces the one configuration that beats both index controls while strengthening out of sample: 19.7% annualized against 9.5% for SPY and 4.2% for RSP over January 2022 through December 2025, at a shallower maximum drawdown, on the identical dividend-free basis.

Across four studies the series has now traced one object from indicator to account: position within the displacement band forecasts the next transition [1]; the composition of the displacement change identifies which transitions are genuine [2]; the joint dynamics those two coordinates define are stationary and shared across the index [3]; and the resulting signal converts into a portfolio through liquidity, not optimization. A regularity repeats along that path. At every stage the durable finding was the simple, interpretable one—a monotone forecast, a two-term decomposition, a shared flow field, a volume rank—while each attempt at additional sophistication either restated the simple result or failed to survive out of sample. The final trading rule fits in two sentences: when the gated signal fires, trade only names in the current top hundred by volume, accepting the most-traded first into ten equal slots; sell each position at the first rollback. Nothing in it requires estimation, and all of it can be checked by eye against a public table.

Three directions follow. The conditional volume effect deserves a mechanism—distinguishing information-incorporation accounts from microstructure ones would turn a documented regularity into an understood one. The evaluation deserves harder eras: membership histories reaching into 2008 would test the liquid-bench premise in the one environment this decade never supplied. And the discipline deserves richer challengers: whether any function class can beat the rank under the same walk-forward, confirmation-window standard is an open and well-posed question. Until one does, the operating summary of this paper is its title's answer: the signal says when; volume says which.

Appendix A. Reproducibility

The full pipeline is released as four Python scripts, together with every result file cited in this paper. Price data is not redistributed; script 01 rebuilds the archive from polygon.io under a user-supplied key, constructing the universe from the membership history [12] and computing the indicator stack, signal flags, and day-by-day membership of Section III—the same columns the published runs consumed from the companion pipeline (MAD-Velocity-Signal). The data vintage is pinned (July 16, 2026) and the index-control series are cached alongside the results, so every number in the paper, including each envelope percentile and Monte Carlo p-value, reproduces exactly: all randomness derives from named seed sequences, and reruns are byte-identical on a given machine. Each engine was additionally verified against hand-constructed fixtures with known answers—trade mechanics, walk-forward knowability, membership gating, era-break handling, and dollar accounting checked to the cent—before touching market data; the fixture suites ship in the repository's tests/ directory.

Table A1. Analysis pipeline (Python; pandas/numpy; polygon.io for prices under a user key).

Script	Output	Role
01_data.py	data/raw/, data/<TICKER>.csv	Universe (point-in-time membership [12], rename and predecessor stitching) + 20-year daily bars + the backtest-ready columns
02_ranking.py	results/02_ranks.csv	The daily ranking table: every constituent, every day, twelve ranks in three families (Table II)
03_score.py	results/03_scores.csv, 03_verdict.csv	Signal ledger (every candidacy resolved through the trade mechanics) + three nested walk-forward scoring models (Table III)
04_sim.py	results/04_sim_*.csv	The simulation menu, random-selection envelopes, and SPY/RSP controls under realistic account mechanics (Tables IV–V)

References

[1] L. Arrington, "Position Within Trend: The MAD-Markov Model for Calibrated Regime-Transition Forecasting," independent research, July 2026.

[2] L. Arrington, "MAD-Velocity: Constructing and Evaluating a Moving-Average-Distance Trading Signal," independent research, July 2026.

[3] L. Arrington, "MAD-Manifold: Mapping and Testing the Phase-Space Flow of Moving-Average Distance," independent research, July 2026.

[4] H. Markowitz, "Portfolio Selection," Journal of Finance, vol. 7, no. 1, pp. 77–91, 1952.

[5] S. Gervais, R. Kaniel, and D. H. Mingelgrin, "The High-Volume Return Premium," Journal of Finance, vol. 56, no. 3, pp. 877–919, 2001.

[6] Y. Amihud, "Illiquidity and Stock Returns: Cross-Section and Time-Series Effects," Journal of Financial Markets, vol. 5, no. 1, pp. 31–56, 2002.

[7] V. DeMiguel, L. Garlappi, and R. Uppal, "Optimal Versus Naive Diversification: How Inefficient Is the 1/N Portfolio Strategy?," Review of Financial Studies, vol. 22, no. 5, pp. 1915–1953, 2009.

[8] V. T. Datar, N. Y. Naik, and R. Radcliffe, "Liquidity and Stock Returns: An Alternative Test," Journal of Financial Markets, vol. 1, no. 2, pp. 203–219, 1998.

[9] R. A. Korajczyk and R. Sadka, "Are Momentum Profits Robust to Trading Costs?," Journal of Finance, vol. 59, no. 3, pp. 1039–1082, 2004.

[10] R. Sullivan, A. Timmermann, and H. White, "Data-Snooping, Technical Trading Rule Performance, and the Bootstrap," Journal of Finance, vol. 54, no. 5, pp. 1647–1691, 1999.

[11] D. H. Bailey, J. M. Borwein, M. López de Prado, and Q. J. Zhu, "Pseudo-Mathematics and Financial Charlatanism: The Effects of Backtest Overfitting on Out-of-Sample Performance," Notices of the American Mathematical Society, vol. 61, no. 5, pp. 458–471, 2014.

[12] fja05680, "S&P 500 Historical Components & Changes," GitHub repository, github.com/fja05680/sp500, accessed July 2026.