

Stocks as Time-Varying Investment States: Cross-Ticker Similarity, Dynamic Allocation, and the Limits of Predictive Transfer

Shenggang Li

Independent Researcher

datalev@gmail.com

Working Paper – September 2026

Abstract

A stock is a persistent legal entity but not a persistent investment opportunity. This paper represents equities as time-varying stock-time states and asks whether similarity across companies supports predictive transfer and adaptive allocation. The analysis uses 4,597 stock-time observations for 30 U.S.-listed equities, 16 price- and market-state variables, and a 63-trading-day forward log-return target under walk-forward label maturity. In the raw historical pool, the nearest cross-ticker state is closer than the nearest same-ticker state in 93.3% of 2,562 comparable cases. That result is strongly affected by candidate availability: after matching the number of cross-ticker candidates to the available same-ticker history, the corresponding probability is 40.4% (95% moving-block interval 38.7–42.0%). Same-ticker observations also appear 1.60 times as often among the 30 nearest pooled states as their share of the eligible history would predict. Thus a large cross-sectional pool supplies many close analogues, while ticker identity retains measurable local structure. Predictive transfer is weaker. At $k = 30$, same-ticker k NN has MAE 14.14%, versus 14.60% for cross-ticker-only and 14.64% for pooled k NN; the paired same-minus-cross 95% interval is $[-1.03, 0.09]$ percentage points. On the common 2,549-row sample, a pooled historical-mean forecast has lower MAE (13.71%) than any of the three k NN rules, and moving-block intervals for monthly cross-sectional rank correlations all include zero. A separate Bayesian-Thompson incremental-allocation demonstration produces a 39.58% time-weighted CAGR versus 39.63% for equal-weight dollar-cost averaging, with a modestly smaller maximum drawdown (-36.61% versus -39.38%). Because the 30-stock universe is ex-post selected and not survivorship-free, these portfolio returns are descriptive rather than evidence of alpha. The evidence supports a narrower methodological conclusion: state representation, local similarity, and return predictability are distinct objects. A more defensible next model is hierarchical partial pooling that learns how much information to share across state similarity and ticker identity.

Keywords: dynamic equity selection; investment states; contextual bandits; Thompson Sampling; cross-ticker transfer; portfolio allocation; walk-forward backtesting; nonstationarity.

JEL codes: G11, G17, C44, C53.

1 Introduction

A ticker identifies a legal and accounting entity, but it does not freeze the investment opportunity represented by that entity. Growth, valuation, risk, competitive position, market regime, and the investor's existing portfolio all change through time. For a decision problem, it is useful to distinguish corporate identity from the state in which the company is encountered.

The motivating representation is

$$\text{company identity} \neq \text{investment state.} \quad (1)$$

For example,

$$\text{NVDA}_{2023} \neq \text{NVDA}_{2026} \quad (2)$$

means only that the two observations may describe materially different investment opportunities even though the underlying company is the same. We represent a stock-time observation by $(i, t, x_{i,t})$, where i is ticker identity, t is decision time, and $x_{i,t}$ is the observed state vector.

This representation is consistent with empirical asset-pricing models in which characteristics and exposures vary over time and with sequential-decision models in which actions are conditioned on states rather than permanent labels (Kelly et al., 2019; Gu et al., 2020; Sutton and Barto, 2018). The representation by itself makes no forecasting claim. The empirical question is whether observations that are close in the chosen state space transfer useful information across companies.

The allocation implementation that motivated the study omits ticker identity from its pooled predictive features. It combines an empirical-Bayesian linear model with pooled nearest-neighbor information, uses a posterior draw for Thompson-style ranking, and then applies portfolio-aware penalties before allocating incremental capital. That design embodies a strong pooling idea: a similar historical state may be more useful than the same company’s distant history. The paper treats that idea as a hypothesis to be tested rather than a premise.

Three questions organize the analysis.

1. **State geometry and candidate availability:** When the full historical pool is searched, how often is the nearest analogue drawn from another ticker, and how much of that result remains after equalizing the number of same- and cross-ticker candidates?
2. **Predictive transfer:** Do pooled or cross-ticker-only neighbors predict the next 63-trading-day return more accurately than same-ticker history, and do any of these local-state rules improve on simple historical-mean forecasts?
3. **Decision use:** When pooled state learning is embedded in an incremental allocation rule, does it materially change where new capital goes, and how does the resulting path compare with simple baselines inside the same eligible universe?

The geometry has two distinct features. In the raw historical pool, the nearest cross-ticker state is closer than the nearest same-ticker state in 93.3% of the 2,562 cases for which both types of history are available. Cross-ticker observations also make up 94.46% of the 30 nearest pooled states. Those raw proportions reflect the much larger number of other-company observations in the search pool. Once the cross-ticker candidate count is matched to the available same-ticker history, the probability that a cross-ticker state supplies the closer minimum falls to 40.4%. Consistently, same-ticker observations account for 5.54% of the 30 nearest pooled states even though their average share of eligible history is only 3.46%. The data therefore show both abundant cross-sectional analogues and persistent ticker-specific clustering.

Prediction is less favorable to local state matching. Same-ticker k NN has lower full-sample MAE than cross-ticker-only or pooled k NN for $k \in \{5, 12, 30\}$, but the paired moving-block intervals include zero. The ordering changes across subperiods, and mean monthly cross-sectional Spearman correlations are near zero or negative. A further benchmark is more revealing: on the common sample, a pooled historical-mean forecast has MAE 13.71%, below the 14.14% same-ticker k NN result and below the pooled and cross-ticker k NN results. Local proximity in this price-based state space is therefore not sufficient for useful return prediction.

The portfolio experiment reaches a similarly restrained result. The Thompson and deterministic dynamic rules produce time-varying allocation paths and modestly lower drawdowns than equal-weight dollar-cost averaging, but the long-run CAGRs are essentially the same. Because the 30-stock universe was selected for demonstration and contains many firms that became strong historical performers, comparisons with SPY or QQQ are descriptive only and cannot establish market-beating

alpha.

The contribution is a representation framework and an empirical stress test of its strongest pooling implication. The results separate three concepts that are easy to conflate: the availability of cross-ticker analogues, local clustering in state space, and predictive sufficiency. They point toward partial pooling rather than either extreme of ignoring ticker identity or fitting thirty unrelated models.

The remainder of the paper proceeds as follows. Section 2 relates the study to empirical asset pricing, Bayesian sequential allocation, reinforcement-learning applications, and backtesting. Section 3 formalizes the state and allocation framework. Section 4 describes the data and validation protocol. Section 5 reports the empirical results, including candidate-count matching and simple forecast benchmarks. Section 6 develops the interpretation. Section 7 discusses limitations and extensions. Section 8 concludes. Technical details and full uncertainty tables appear in the appendices.

2 Related Literature

2.1 Time-varying stock characteristics and return prediction

Expected returns and risk exposures need not be stable attributes of a ticker. Kelly et al. (2019) use characteristics to instrument time-varying factor loadings, while Gu et al. (2020) study nonlinear machine-learning methods for the cross section of U.S. equity returns. This paper asks a narrower question. Rather than estimating a broad factor model or constructing a comprehensive point-in-time characteristics database, it tests whether a simple price- and market-state representation supports information transfer across stock-time observations.

The distinction is consequential because the present feature vector omits variables that would ordinarily matter for an economic description of investment state, including point-in-time valuation, revenue growth, profitability, balance-sheet quality, and industry fundamentals. The empirical exercise should be read as a first layer of the state-representation problem, not as a complete model of the investment opportunity set.

2.2 Contextual bandits, Thompson Sampling, and nonstationarity

A contextual bandit chooses an action from observed context and learns from realized rewards. Thompson Sampling implements Bayesian randomized decision-making by drawing a plausible model from the posterior and acting on that draw (Agrawal and Goyal, 2013). The allocation experiment uses this mechanism for posterior-randomized ranking of current stock states.

The analogy to a classical partial-feedback bandit is not exact. Ex-post market returns are observable for listed stocks even when a particular stock was not purchased, so the purpose of the posterior draw here is not to create information by hiding unchosen rewards. It is a disciplined way to propagate parameter uncertainty into a sequential ranking rule. For that reason, the paper uses the terms *Thompson-style ranking* or *Bayesian randomized allocation* rather than treating the backtest as a canonical bandit-learning experiment.

Nonstationarity remains central. Relationships between state variables and returns can drift with interest rates, technology cycles, competition, investor expectations, and market structure. Nonstationary bandit theory formalizes the cost of changing reward processes and the tension between reusing history and emphasizing recent information (Besbes et al., 2014). The empirical design uses rolling historical training sets and reports subperiod results instead of presuming that one transfer rule is stable through time.

2.3 Sequential portfolio allocation and reinforcement-learning connections

When today’s allocation changes tomorrow’s holdings, concentration, cash, and risk capacity, the decision problem becomes sequential and is naturally related to reinforcement learning (Sutton and Barto, 2018). Financial applications include deep reinforcement learning for trading and portfolio choice (Zhang et al., 2020; Han, 2023; Ayala et al., 2025); Borkar and Jadhav (2024) survey the broader stock-trading literature. The implementation here is deliberately simpler. It allocates incremental new capital, recomputes portfolio penalties after each micro-purchase, and permits cash to remain unallocated. It is a sequential allocation demonstration rather than a solved Markov decision process or a deep-RL trading system.

The equal-weight baseline is especially relevant. DeMiguel et al. (2009) show that estimation error can erase the theoretical advantages of more sophisticated allocation rules relative to naive $1/N$ diversification. In the present setting, complexity is not evidence of success; the dynamic rule must be judged against a simple strategy using the same eligible stock universe and cash-flow schedule.

2.4 Backtesting, evidence, and policy evaluation

Financial machine learning is vulnerable to small samples, repeated specification search, look-ahead bias, and ex-post universe construction. Arnott et al. (2019) emphasize disciplined backtesting protocols in this setting. The design here enforces target maturity and time-ordered training, while retaining a major limitation: the demonstration universe is not a point-in-time, survivorship-free reconstruction.

Formal offline policy evaluation is a separate layer. Companion work examines when historical partial-feedback evidence can resolve allocation-policy choices and when apparently better value estimates remain too uncertain for deployment (Li, 2026a,b). The present paper does not use importance weighting or doubly robust OPE. Its scope is earlier in the pipeline: define the state representation, test cross-ticker transfer, and audit a walk-forward allocation mechanism.

3 Time-Varying Investment-State Framework

3.1 From ticker identity to stock-time states

Let $i \in \{1, \dots, N\}$ index stocks and let t denote a decision time. The observable state is

$$x_{i,t} \in \mathbb{R}^d, \quad (3)$$

and we index each observation as

$$z_{i,t} = (i, t, x_{i,t}). \quad (4)$$

Ticker identity remains part of the observation, but the pooled predictive model intentionally does not use i as a feature. Here $d = 16$ and the state contains multi-horizon returns, volatility, drawdown, moving-average gaps, relative performance versus SPY, and common market-regime variables.

Let $d(x, x')$ denote Euclidean distance after standardizing each feature with means and standard deviations estimated from the contemporaneously eligible training set. The representation allows

$$d(x_{i,t}, x_{j,s}) < d(x_{i,t}, x_{i,u}) \quad (5)$$

for $j \neq i$: a historical observation from another company can be closer in the chosen feature space than an observation from the same company. Because the cross-ticker pool is much larger, however, the raw nearest-neighbor comparison mixes state geometry with candidate availability. Section 4 therefore evaluates both the raw search pool and a candidate-count-matched diagnostic.

For prediction, define three inverse-distance estimators for a current state $x_{i,t}$:

$$\hat{m}_k^{\text{pool}}(x_{i,t}) : \text{the } k \text{ nearest eligible historical states from all tickers,} \quad (6)$$

$$\hat{m}_k^{\text{same}}(x_{i,t}) : \text{the same calculation restricted to ticker } i, \quad (7)$$

$$\hat{m}_k^{\text{cross}}(x_{i,t}) : \text{the same calculation restricted to tickers } j \neq i. \quad (8)$$

If $\mathcal{N}_k(x)$ is the selected neighbor set and $w_n \propto 1/\max(d_n, 10^{-6})$, then

$$\hat{m}_k(x) = \sum_{n \in \mathcal{N}_k(x)} w_n y_n, \quad \sum_{n \in \mathcal{N}_k(x)} w_n = 1. \quad (9)$$

The empirical issue is not whether cross-ticker information must dominate. It is whether the observed state vector contains enough transferable information to justify pooling. Persistent company-specific information omitted from $x_{i,t}$ can make aggressive pooling harmful even when other companies provide many close analogues.

3.2 Future-return target

The target horizon is 63 trading days. For adjusted price $P_{i,t}$,

$$y_{i,t} = \text{clip} \left(\log \frac{P_{i,t+63}}{P_{i,t}}, -1, 1 \right). \quad (10)$$

Clipping limits extreme labels in the linear model. A label observed at signal date t is not eligible for training until its entire 63-day horizon has elapsed. This maturity rule prevents a rolling model from using future-return information that overlaps the decision date.

3.3 Empirical-Bayesian pooled opportunity model

Let X denote standardized training features with an intercept and y the vector of mature forward returns. The pooled linear model is

$$y = X\beta + \varepsilon, \quad \varepsilon \sim \mathcal{N}(0, \sigma^2 I), \quad (11)$$

with Gaussian prior

$$\beta \sim \mathcal{N}(0, \tau^2 I). \quad (12)$$

The implementation estimates σ^2 from a lightly regularized preliminary fit and forms

$$\Sigma_{\beta|D} = \left(\tau^{-2} I + \sigma^{-2} X^\top X \right)^{-1}, \quad (13)$$

$$\mu_{\beta|D} = \Sigma_{\beta|D} \sigma^{-2} X^\top y. \quad (14)$$

The posterior mean prediction for a current state is $\hat{q}_{i,t}^B$ and its model-based epistemic standard deviation is $u_{i,t}^B$.

The allocation implementation blends the Bayesian prediction with pooled k NN:

$$\hat{q}_{i,t} = (1 - \omega) \hat{q}_{i,t}^B + \omega \hat{m}_{30}^{\text{pool}}(x_{i,t}), \quad \omega = 0.25. \quad (15)$$

For Thompson-style ranking, one common posterior coefficient vector is drawn at each decision date,

$$\tilde{\beta}_t \sim \mathcal{N}(\mu_{\beta|D_t}, \Sigma_{\beta|D_t}), \quad (16)$$

and the sampled Bayesian component replaces its posterior mean. A fixed master seed of 42 is used in the reported backtest, with a deterministic date-specific seed derived from it.

The code also forms a composite uncertainty proxy by linearly blending the Bayesian epistemic standard deviation and the k NN standard-error estimate. This quantity enters the allocation penalty below. It should be interpreted as a model-based uncertainty score, not as a calibrated posterior standard deviation for the blended forecast.

3.4 From standalone opportunity to a next-dollar allocation score

A high standalone return estimate need not identify the best next purchase. Let $w_{i,t}$ be the current portfolio weight, $s_{i,t}$ the current weight of the stock's sector, $\rho_{i,t}$ the stock's recent correlation with the held portfolio, $v_{i,t}$ its recent annualized volatility, $D_{i,t}$ its absolute drawdown, and $u_{i,t}$ the composite uncertainty score. The implementation uses the heuristic next-dollar score

$$A_{i,t} = q_{i,t} - \lambda_v v_{i,t} - \lambda_D D_{i,t} - \lambda_w w_{i,t} - \lambda_s s_{i,t} \quad (17)$$

$$- \lambda_\rho \max(\rho_{i,t}, 0) - \lambda_u u_{i,t}. \quad (18)$$

The frozen coefficients are

$$(\lambda_v, \lambda_D, \lambda_w, \lambda_s, \lambda_\rho, \lambda_u) = (0.04, 0.05, 0.18, 0.04, 0.008, 0.10). \quad (19)$$

These coefficients define an allocation score rather than the derivative of a fully specified expected-utility function. Cash has score zero. A candidate must also respect a 12% single-stock cap and a 35% sector cap.

Monthly deployable cash is divided into 20 micro-purchases. After each purchase, weights and portfolio penalties are recomputed before the next unit of capital is allocated. Each step approximately chooses

$$\arg \max_{i \in \mathcal{U}_t \cup \{\text{cash}\}} A_{i,t}. \quad (20)$$

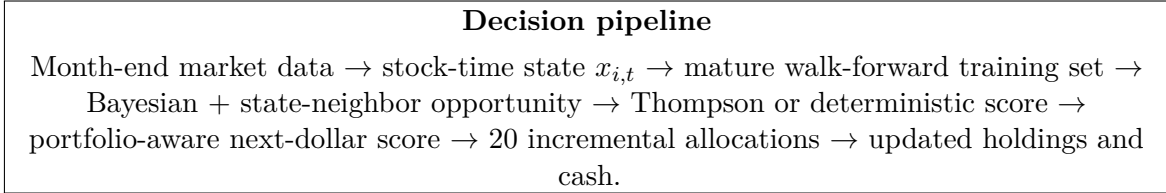


Figure 1: Research pipeline. Portfolio state affects the destination of the next unit of capital even when the standalone opportunity estimate is unchanged.

4 Data and Empirical Protocol

4.1 Research universe and state panel

The analysis uses adjusted-close data for a demonstration universe of 30 U.S.-listed stocks together with SPY and QQQ benchmarks. The stock universe is

NVDA, MSFT, GOOG, AMZN, META, AAPL, AVGO, AMD, TSM, ASML, ANET, MRVL, MU, LRCX, KLAC, AMAT, CRWD, PANW, ORCL, VRT, ETN, GEV, APH, LLY, ISRG, COST, WMT, MA, JPM, CAT.

The stored state panel contains 4,597 stock-time rows from January 31, 2013 through September 2, 2026. Regular state dates are month-end through August 31, 2026; the September 2 snapshot is retained for the most recent allocation demonstration. Late-listed companies enter only after enough price history exists to construct the state variables. The target is the 63-trading-day forward log return. Table 1 summarizes the 16 features.

Table 1: Price- and market-state variables

Feature group	Variables
Own-stock return state	21-, 63-, 126-, and 252-day returns
Own-stock risk state	21- and 63-day annualized volatility; 252-day drawdown
Trend state	Price gap to 50-day and 200-day moving averages
Relative state	63- and 126-day return relative to SPY
Common market state	SPY 63- and 252-day return; SPY 63-day volatility; SPY 252-day drawdown; QQQ minus SPY 63-day return

Ticker identity is not included in the pooled predictive feature vector. Point-in-time accounting and valuation variables are also absent; Section 7 treats that omission as an important limitation rather than an innocuous simplification.

4.2 Walk-forward timing

At signal date t , the model uses up to eight years of historical observations satisfying

$$s < t, \quad \text{target_available_date}_{i,s} \leq t, \quad y_{i,s} \text{ observed.} \quad (21)$$

At least 400 pooled training rows are required. Portfolio signals use the previous trading day’s close and execution occurs on the first trading day of the next month.

The realized state-transfer validation starts in 2019 and ends on May 29, 2026, because later 63-day targets had not matured in the frozen panel. It contains 2,572 stock-month observations across 89 signal dates. A nearest same-ticker versus cross-ticker comparison is possible for 2,562 observations; 10 early-listing cases have no eligible own-ticker history. Same-ticker k NN forecasts require at least five own-history observations and are available for 2,549 rows.

4.3 State geometry and candidate-count matching

Distances are computed after standardizing features with the contemporaneously eligible training sample. Four diagnostics are used:

1. whether the nearest cross-ticker state is closer than the nearest same-ticker state in the full eligible history;
2. the fraction of the pooled 30 nearest neighbors that share the current ticker;
3. the same-ticker fraction expected from the ticker’s share of the eligible historical pool;
4. an equal-candidate-count comparison that removes the mechanical minimum-distance advantage of the much larger cross-ticker pool.

The fourth diagnostic is important because a minimum drawn from thousands of cross-ticker candidates will tend to be smaller than a minimum drawn from a much shorter own-ticker history even when the underlying distance distributions are similar. For a current observation, let n_s be the

number of eligible same-ticker candidates, let N_c be the number of cross-ticker candidates, and let R count cross-ticker states closer than the nearest same-ticker state. If $n = \min(n_s, N_c)$ cross-ticker candidates are sampled without replacement, the exact probability that the matched cross-ticker sample contains at least one closer state is

$$p_{\text{matched}} = 1 - \frac{\binom{N_c - R}{n}}{\binom{N_c}{n}}. \quad (22)$$

Averaging this probability over comparable stock-months gives the candidate-count-matched statistic. This calculation is exact conditional on each historical pool and introduces no Monte Carlo sampling noise.

For the raw and matched geometry statistics, uncertainty intervals are obtained by applying the same three-month moving-block bootstrap over the 89 date-level averages, with 4,000 resamples and seed 42.

4.4 Predictive-transfer tests and simple benchmarks

For each realized current target, pooled, cross-ticker-only, and same-ticker inverse-distance forecasts are computed for $k \in \{5, 12, 30\}$. The primary error metric is

$$\text{MAE} = \frac{1}{M} \sum_{m=1}^M |\hat{y}_m - y_m|, \quad (23)$$

and RMSE and monthly cross-sectional Spearman rank correlation are also reported.

Three deliberately simple forecasts provide a scale for the k NN results: zero return, the mean mature return in the current pooled training window, and the mean mature return for the current ticker. Comparisons involving same-ticker history use the common 2,549-row sample. The zero-return and pooled-mean forecasts are constant within a decision date, so a cross-sectional rank correlation is not defined for them.

Adjacent monthly 63-day targets overlap, so row-level independence is not assumed. Paired MAE contrasts are first averaged within signal date and then evaluated with a three-month moving-block bootstrap over dates using 4,000 resamples and seed 42. The same date-block procedure is used to form uncertainty intervals for mean monthly Spearman IC. These intervals are dependence-aware empirical diagnostics rather than formal finite-sample guarantees for the full data-adaptive process.

4.5 Incremental portfolio experiment

The allocation backtest runs from January 2, 2019 through September 2, 2026. All strategies receive the same external cash flows: initial capital of \$20,000 and monthly contributions of \$2,000. The two dynamic strategies may also reconsider 25% of previously idle cash. Each deployable budget is split into 20 micro-purchases, and transaction cost is fixed at 5 basis points per purchase.

The principal internal benchmark is equal-weight DCA over the same currently eligible demonstration universe. SPY and QQQ DCA are included only for context. Because the 30-stock list was chosen for demonstration rather than reconstructed ex ante from a point-in-time universe, the stock-selection history is exposed to survivorship and ex-post selection bias. Index-relative performance is therefore not interpretable as alpha.

Table 2 records the frozen settings used in the reported implementation.

Table 2: Frozen settings for the allocation demonstration

Item	Setting
Stock universe	30 demonstration stocks; late listings enter after sufficient history
State frequency	Month-end, plus final September 2, 2026 snapshot for the latest demonstration
Forward target	63-trading-day clipped log return; clipping bound $[-1, 1]$
Training lookback	Up to 8 years; minimum 400 pooled rows; mature targets only
k NN blend	30 pooled neighbors; inverse-distance weighting; blend weight $\omega = 0.25$
Bayesian model	Gaussian empirical-Bayesian linear model; prior variance 1.0
Randomized ranking	One common posterior coefficient draw per signal date; master seed 42
Capital process	\$20,000 initial capital; \$2,000 monthly contribution; 20 micro-purchases
Constraints	12% single-stock cap; 35% sector cap; cash score 0
Trading friction	5 bps illustrative cost per purchase
Dynamic cash	25% of previously idle cash may be reconsidered each month
Execution	Signal on prior trading day; execution on next month’s first trading day

5 Empirical Results

5.1 Raw and candidate-count-matched state geometry

Searching the full eligible history produces a striking raw nearest-neighbor result, but candidate availability changes its interpretation. Table 3 reports both views. Among 2,562 stock-months with at least one same-ticker and one cross-ticker historical candidate, the nearest cross-ticker state is closer in 93.33% of cases. Median standardized distance is 1.80 for the nearest cross-ticker observation and 2.37 for the nearest same-ticker observation.

The cross-ticker pool is much larger, however. After matching the number of cross-ticker candidates to the available same-ticker history, the probability that the matched cross-ticker sample contains a state closer than the same-ticker minimum falls to 40.43%. The three-month moving-block 95% interval is 38.72–41.99%, entirely below 50%. The raw 93.33% result is therefore best read as a statement about the availability of analogues in a large cross-sectional library, not as evidence that cross-ticker states are intrinsically more tightly clustered around the current observation.

Table 3: Raw and candidate-count-matched geometry of historical investment states

Diagnostic	Value
Evaluation stock-months	2,572
Evaluation dates	89
Comparable nearest-same/cross cases	2,562
Raw $P(\text{nearest cross} < \text{nearest same})$	93.33%
Raw moving-block 95% interval	[92.30%, 94.27%]
Candidate-count-matched $P(\text{cross closer})$	40.43%
Matched moving-block 95% interval	[38.72%, 41.99%]
Median nearest cross-ticker standardized distance	1.80
Median nearest same-ticker standardized distance	2.37
Observed same-ticker share among 30 nearest pooled states	5.54%
Expected same-ticker share from eligible history	3.46%
Observed/expected same-ticker neighbor ratio	1.60
Cross-ticker share among 30 nearest pooled states	94.46%

The neighbor-composition diagnostic points in the same direction as the matched comparison.

Same-ticker states occur 1.60 times as often as random pool composition would imply. Most pooled neighbors are still other companies because other-company observations dominate the candidate pool, but ticker identity remains visible in local state geometry.

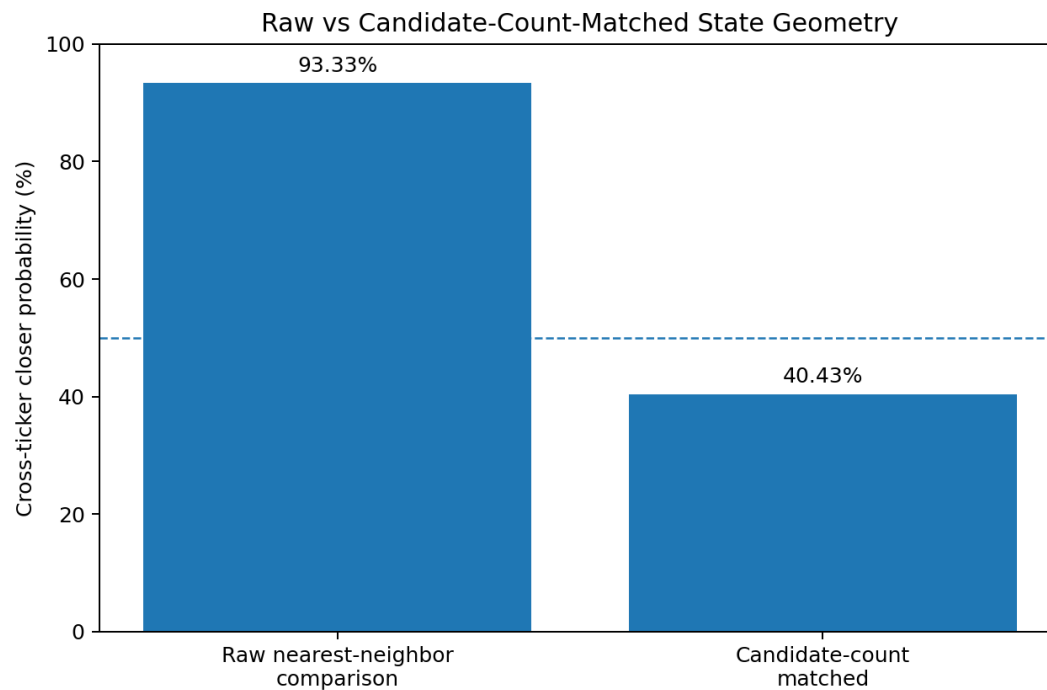


Figure 2: Raw versus candidate-count-matched nearest-state geometry. The full cross-ticker pool almost always contains a closer observation, but that advantage reverses after equalizing candidate counts. The dashed line marks 50%.

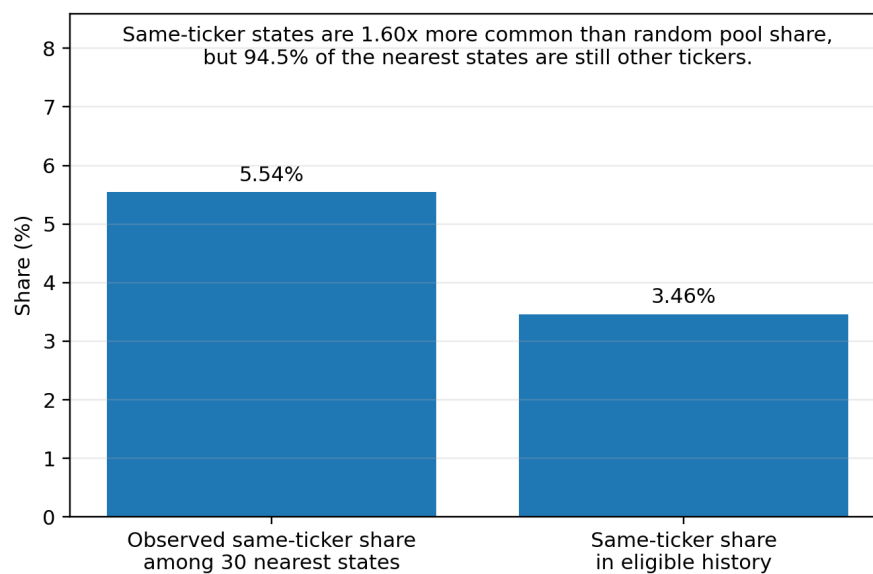


Figure 3: Ticker identity among the 30 nearest pooled states. Same-ticker observations account for 5.54% of neighbors versus 3.46% expected from their share of eligible history.

5.2 Local state matching does not beat simple forecast benchmarks

Table 4 compares out-of-sample 63-day return forecasts. Same-ticker k NN has the lowest MAE and RMSE for each tested neighborhood size. At $k = 30$, same-ticker MAE is 14.14%, compared with 14.60% for cross-ticker-only k NN and 14.64% for pooled k NN.

Table 4: Walk-forward nearest-neighbor return forecasting

k	History restriction	n	MAE	RMSE	Mean monthly Spearman IC
5	Pooled all tickers	2,572	15.77%	20.88%	-0.015
5	Cross-ticker only	2,572	15.71%	20.78%	-0.005
5	Same ticker only	2,549	15.39%	20.56%	-0.054
12	Pooled all tickers	2,572	15.16%	20.09%	-0.021
12	Cross-ticker only	2,572	15.15%	20.06%	-0.018
12	Same ticker only	2,549	14.66%	19.55%	-0.044
30	Pooled all tickers	2,572	14.64%	19.45%	-0.042
30	Cross-ticker only	2,572	14.60%	19.43%	-0.021
30	Same ticker only	2,549	14.14%	18.89%	-0.044

On the 2,549 paired rows with sufficient same-ticker history, the date-level same-minus-cross MAE difference at $k = 30$ is -0.41 percentage points, with a three-month moving-block 95% interval of $[-1.03, 0.09]$ percentage points. The same-minus-pooled difference is -0.45 percentage points with interval $[-1.08, 0.07]$. The point estimates favor same-ticker history, but neither contrast is statistically decisive.

A more demanding benchmark compares the local-state rules with forecasts that use no local distance information. Table 5 uses the same 2,549-row sample for all entries. The pooled historical mean has MAE 13.71%, and the same-ticker historical mean has MAE 13.82%; both are below the 14.14% same-ticker k NN result. The zero-return forecast is worse at 14.89%, but still close enough to emphasize that the predictive content of the tested price-state neighborhood is weak.

Table 5: Simple forecast benchmarks on the common same-history-eligible sample

Forecast	n	MAE	RMSE
Zero return	2,549	14.89%	19.60%
Pooled historical mean	2,549	13.71%	18.29%
Same-ticker historical mean	2,549	13.82%	18.45%
Same-ticker k NN, $k = 30$	2,549	14.14%	18.89%
Cross-ticker k NN, $k = 30$	2,549	14.56%	19.38%
Pooled k NN, $k = 30$	2,549	14.60%	19.40%

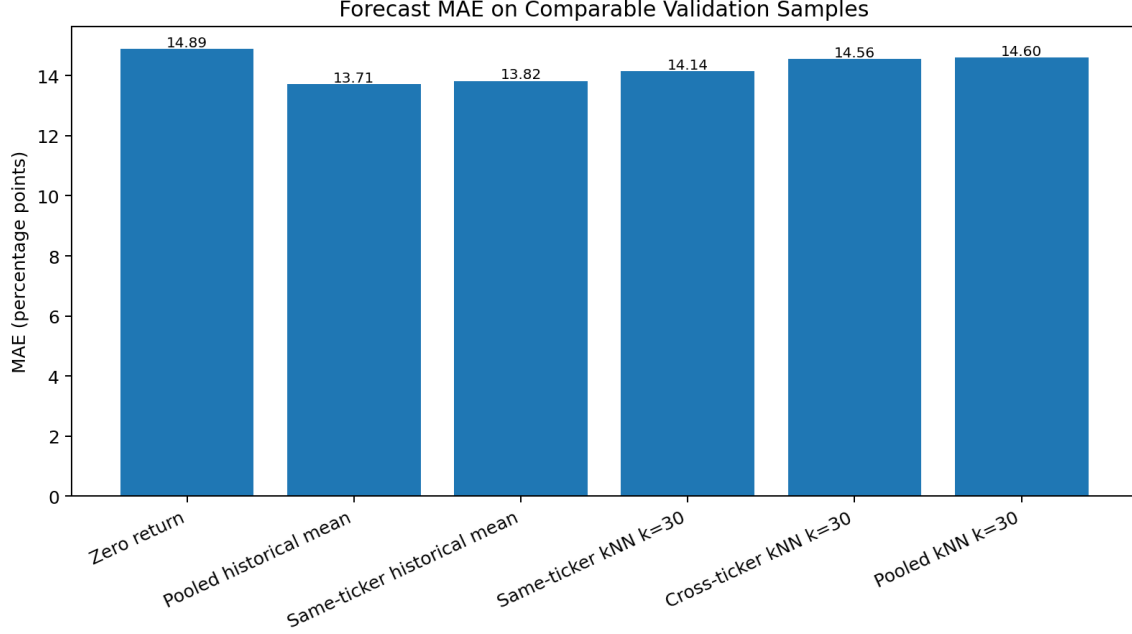


Figure 4: MAE for simple and local-state forecasts on comparable validation samples. Historical-mean forecasts outperform the tested k NN rules in absolute error.

Cross-sectional ranking is also weak. For $k = 30$, mean monthly Spearman IC is -0.042 for pooled, -0.021 for cross-ticker-only, and -0.044 for same-ticker k NN. Their respective three-month moving-block 95% intervals are $[-0.101, 0.009]$, $[-0.080, 0.031]$, and $[-0.104, 0.021]$. All tested IC intervals for $k \in \{5, 12, 30\}$ include zero; Appendix E reports the full set. The price-only state vector and Euclidean distance therefore do not establish a stable cross-sectional return-ranking signal.

5.3 The transfer ordering changes over time

Subperiod results reinforce the nonstationarity of the transfer relationship. Table 6 reports $k = 30$ MAE by period. Cross-ticker-only and pooled forecasts are slightly better in 2019–2021, while same-ticker history is better in 2022–2024 and 2025–2026.

Table 6: Subperiod robustness of 30-neighbor transfer forecasts

Period	Pooled all tickers	Cross-ticker only	Same ticker only
2019–2021	11.89%	11.86%	12.07%
2022–2024	15.74%	15.66%	14.69%
2025–2026	17.90%	17.92%	17.15%

A fixed decision to pool or not to pool is therefore difficult to justify from these data. The degree of borrowing across companies is more naturally treated as a quantity to estimate, potentially as a function of market regime and company identity.

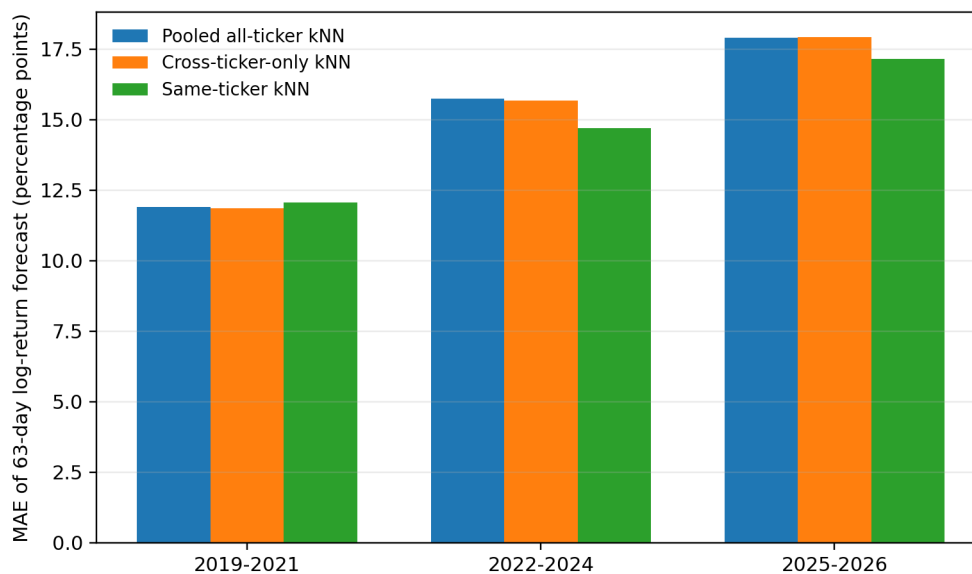


Figure 5: Subperiod MAE for 30-neighbor forecasts. The relative performance of cross-ticker and same-ticker histories changes over time.

5.4 The Bayesian-kNN blend remains a weak ranking signal

The allocation model does not use k NN alone. It blends pooled nearest-neighbor information with an empirical-Bayesian linear model. Across 2,599 realized decision-score rows, the blended posterior-mean score has MAE 14.32%, compared with 14.43% for the Bayesian component and 14.61% for the pooled k NN component. The small reduction in absolute error is not accompanied by a positive ranking signal: mean monthly Spearman IC is -0.034 for the blend, -0.028 for the Bayesian component, and -0.035 for pooled k NN.

This result limits the interpretation of the allocation backtest. Uncertainty draws, concentration limits, cash, and portfolio penalties can produce different allocation paths even when the stock-level return-ranking signal is weak. The blended score should not be described as a validated cross-sectional alpha model.

5.5 Portfolio context changes the next-dollar decision

The allocation rule distinguishes an asset-level forecast from a portfolio-level purchase decision. Existing position weight, sector exposure, recent correlation, volatility, drawdown, and model-based uncertainty can lower the score of a stock that has a high standalone return estimate. The resulting quantity is a heuristic allocation score, not a formally derived expected-utility gradient. This distinction matters because the portfolio experiment is intended to audit the mechanics of sequential allocation rather than establish an optimal-control theorem.

5.6 New capital moves across stocks rather than following a permanent favorite

Figure 6 plots monthly new-capital allocation for the Thompson strategy. The destination of contributions changes across periods and capital is sometimes left in cash. Across the full run, 29 of the 30 stocks receive some allocation; the median month directs new capital to one stock, reflecting the deliberately concentrated next-dollar rule rather than broad rebalancing.

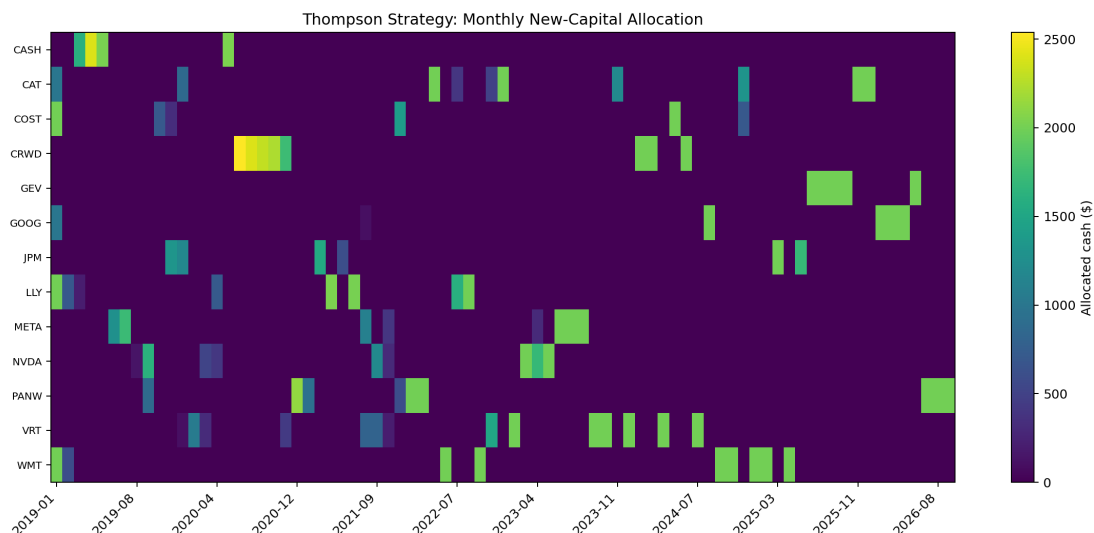


Figure 6: Monthly new-capital allocation heatmap for the Thompson strategy. Preferred destinations vary through time, and cash remains available as an action.

5.7 Dynamic allocation does not beat equal weighting on long-run return

Table 7 reports time-weighted performance within the fixed demonstration universe. The Thompson strategy has a 39.58% CAGR, essentially identical to the 39.63% equal-weight DCA baseline. The deterministic dynamic strategy is slightly lower at 39.00%. Relative to equal weighting, the Thompson path has modestly lower annualized volatility, a higher descriptive Sharpe ratio (computed with a zero risk-free rate), and a smaller maximum drawdown.

Table 7: Historical performance within the fixed demonstration universe, January 2019–September 2026

Strategy	TWR CAGR	Ann. vol.	Sharpe ($r_f = 0$)	Max drawdown	Calmar
Proposed Thompson	39.58%	29.33%	1.285	-36.61%	1.081
Deterministic dynamic	39.00%	29.20%	1.275	-36.80%	1.060
Equal-weight DCA	39.63%	29.90%	1.267	-39.38%	1.006
SPY DCA [†]	17.41%	19.30%	0.929	-33.72%	0.516
QQQ DCA [†]	22.71%	23.96%	0.974	-35.12%	0.646

[†]Descriptive only. The ex-post 30-stock universe is not a valid alpha comparison against SPY or QQQ. Sharpe ratios use a zero risk-free rate and are reported only for descriptive within-sample comparison.

The long-run return result is a useful constraint on the interpretation of the framework. Dynamic allocation changes the path and slightly improves drawdown and risk-adjusted metrics in this sample, but it does not improve CAGR relative to the simple internal baseline. That outcome is consistent with the broader evidence that estimation error can offset the gains from more elaborate portfolio rules (DeMiguel et al., 2009).

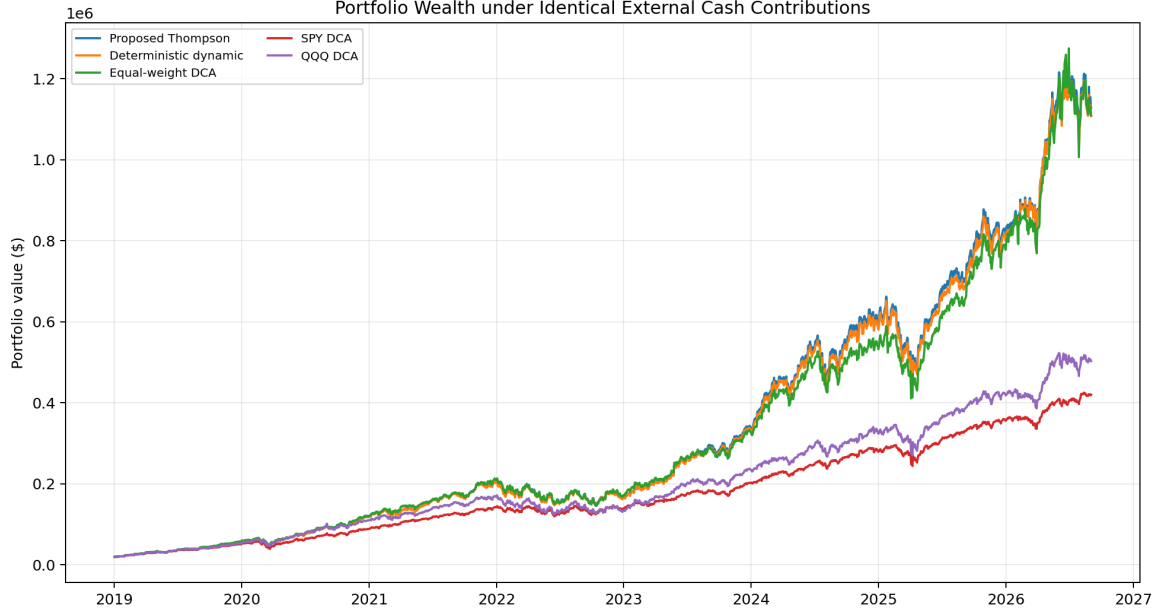


Figure 7: Portfolio wealth under identical external cash contributions. Legend labels match the strategy names used in Table 7. SPY and QQQ are shown only for scale; the ex-post stock universe prevents an alpha interpretation of the index-relative gap.

6 Robustness and Interpretation

6.1 Neighborhood size does not rescue predictive transfer

The predictive ordering is not specific to $k = 30$. At $k = 5$, same-ticker MAE is 15.39% versus 15.71% for cross-ticker history; at $k = 12$, the corresponding values are 14.66% and 15.15%. In every case the moving-block interval for the paired same-minus-cross difference includes zero. More smoothing reduces absolute error across all three local-state rules without creating decisive evidence for cross-ticker predictive superiority.

6.2 Candidate availability and clustering are different geometric questions

The raw and matched geometry statistics answer different questions. The 93.33% raw probability asks what an investor is likely to find when searching the entire historical library. With many more observations from other companies, a cross-ticker analogue is usually the closest available point. The 40.43% matched probability instead asks whether a cross-ticker pool of equal size would still be more likely to contain the closer observation. It would not. The 1.60 observed-to-expected same-ticker neighbor ratio independently points to ticker-specific clustering.

Neither result makes ticker identity a permanent investment state. The evidence says something more limited: the observed price-state vector does not absorb all persistent company-level structure, while a large cross-sectional history still expands the set of available analogues.

6.3 Geometry and prediction answer different questions

A close observation need not have a transferable conditional return distribution. In particular,

$$P(y \mid x, i) = P(y \mid x) \quad (24)$$

need not hold. Same-ticker history can carry persistent omitted information about business quality, industry structure, valuation, accounting characteristics, management, or idiosyncratic exposure that is absent from the 16 price- and market-state variables.

The simple-mean benchmarks sharpen this point. The pooled historical mean has lower MAE than any tested k NN rule on the paired sample. Local geometry therefore fails a stronger test than merely losing to same-ticker neighbors: in this specification, conditioning on Euclidean proximity does not improve absolute forecast accuracy over a simple rolling location estimate. The near-zero IC results further show that lower absolute error should not be confused with useful cross-sectional ranking.

6.4 A partial-pooling interpretation

The evidence motivates a hierarchical model rather than a binary choice between one global model and thirty unrelated ticker models. A generic form is

$$y_{i,t} = f(x_{i,t}; \theta_0) + g_i(x_{i,t}; \theta_i) + \varepsilon_{i,t}, \quad (25)$$

with, for example,

$$\theta_i \sim \mathcal{N}(\theta_0, \Sigma_\theta). \quad (26)$$

The common component f captures transferable state structure; g_i absorbs ticker-specific deviations. The data then determine how strongly each company should borrow information from the cross-sectional pool.

This formulation preserves the central time-varying-state idea while accommodating both matched-sample clustering and the modest full-sample advantage of same-ticker prediction. It also makes the strength of pooling an estimand rather than a fixed modeling choice.

6.5 The backtest is a decision audit, not an alpha test

All three stock-universe strategies use the same external cash-flow schedule and eligible universe, so the allocation experiment is informative about how posterior randomization and portfolio context redirect new capital. It is not a tradable alpha test. The universe was selected ex post, the return-ranking signal is weak, and the strategy primarily allocates new capital instead of executing a full sell/rebalance policy with taxes, market impact, and institutional constraints.

The appropriate reading is correspondingly narrow: the state-based next-dollar mechanism is computationally coherent and produces materially different allocations, but the present evidence does not show that those allocations raise expected return.

7 Limitations and Research Implications

7.1 Ex-post and survivorship-biased universe

The largest empirical limitation is the 30-stock demonstration universe. Many constituents became strong historical performers, and the list was not generated by a point-in-time selection rule that could have been executed in 2012 or 2019. The resulting opportunity set is unusually favorable, which makes the large absolute CAGR unsurprising. A research-grade alpha study would reconstruct the candidate universe at each historical date, include delisted and failed firms, and apply ex-ante screening rules using only information available at that time.

7.2 Price-only states omit economically important variables

The conceptual state includes growth, profitability, valuation, competitive position, and common economic conditions. The empirical state uses only prices and market variables because current fundamentals cannot safely be backfilled into historical dates. This choice protects timing integrity but omits the firm-level variables most likely to make cross-company similarity economically meaningful.

A stronger test requires point-in-time fundamentals that preserve filing dates and restatements. Candidate additions include valuation ratios, revenue and earnings growth, gross and operating margins, free-cash-flow measures, leverage, analyst-expectation revisions, and industry-level capital-expenditure or competition indicators.

7.3 Distance metric and representation learning

Euclidean distance after z-score standardization gives every feature the same local geometry. It is transparent, but there is no reason to expect it to be optimal for returns. Mahalanobis distance, supervised metric learning, learned embeddings, or regime-conditional distances could produce a different transfer map. Any such metric must be estimated inside the walk-forward training sample; learning a representation on the full history would reintroduce look-ahead bias.

The candidate-count-matched diagnostic also shows why geometry should be audited before a distance model is interpreted economically. A large comparison pool can produce visually compelling nearest-neighbor results through order statistics alone.

7.4 Overlapping targets and finite historical evidence

Monthly 63-day targets overlap. Three-month moving-block resampling accounts for an important source of temporal dependence, but it is not a general finite-sample theorem for the full model-selection process. Longer blocks, non-overlapping quarterly evaluation dates, and repeated universe-level resampling would provide stronger inference.

The historical-mean benchmark is another useful guardrail. It shows that the present local-state rules do not improve absolute error over a simple rolling mean. Future representation changes should be required to beat such low-complexity forecasts before they are carried into a portfolio policy.

7.5 The incremental rule is not full reinforcement learning

The strategy primarily allocates incoming capital and reconsidered idle cash. It does not solve a full Markov decision process with systematic sells, tax lots, transaction-cost state, and future capital needs. A fuller portfolio state could be written

$$s_t = (X_t, w_t, C_t, M_t), \quad (27)$$

with action $a_t = \Delta w_t$ and objective

$$\max_{\pi} \mathbb{E}_{\pi} \left[\sum_{t=0}^T \gamma^t r_t \right]. \quad (28)$$

That extension is conceptually natural but would create additional estimation and evaluation risk. The present use of Thompson Sampling should likewise be understood as posterior-randomized allocation under observable market outcomes, not as a classical partial-feedback exploration problem.

7.6 Offline evaluation should follow policy construction

A reconstructed historical backtest answers what would have happened on one path under specified rules. It does not by itself establish whether the logged evidence is sufficient to support deployment of a new policy. Formal OPE, uncertainty analysis, and policy-resolution diagnostics form a separate research stage (Li, 2026a,b). A useful progression is

$$\text{state representation} \rightarrow \text{predictive transfer} \rightarrow \text{portfolio policy} \rightarrow \text{offline policy evaluation.} \quad (29)$$

7.7 Priority extension: hierarchical partial pooling in a point-in-time universe

The most direct next experiment is a point-in-time hierarchical state model that estimates the degree of pooling across ticker identity. A practical comparison is

$$\text{Model A: pooled state-only model,} \quad (30)$$

$$\text{Model B: ticker-specific model,} \quad (31)$$

$$\text{Model C: hierarchical partial-pooling state model,} \quad (32)$$

under identical walk-forward splits and a survivorship-controlled dynamic universe. The primary endpoints should be cross-sectional ranking and calibration of forward-return distributions, not terminal portfolio wealth alone. A second-stage allocation experiment can then ask whether any predictive improvement survives portfolio constraints and transaction costs.

The new validation results make this extension more, not less, relevant. Raw search pools contain abundant cross-ticker analogues, matched samples retain ticker-specific clustering, and local k NN forecasting does not beat simple historical means. The natural next step is to estimate how much information should be shared rather than imposing either complete pooling or complete separation.

8 Conclusion

A ticker is stable; the investment opportunity it represents is not. Treating equities as stock-time states is a useful way to express changing return, risk, market, and portfolio conditions. The empirical results also show why that representation should not be confused with a forecasting result.

In the full historical library, the nearest cross-ticker state is closer than the nearest same-ticker state in 93.3% of comparable cases. Once candidate counts are equalized, the cross-ticker closer probability falls to 40.4%, and same-ticker observations appear 1.60 times as often among pooled nearest neighbors as their historical availability would predict. The data therefore contain two effects at once: a broad cross-sectional pool supplies many available analogues, while ticker identity retains persistent local structure.

Neither effect is enough to produce a strong return signal. Same-ticker k NN has modestly lower full-sample MAE than cross-ticker or pooled k NN, but the paired uncertainty intervals include zero and the ordering changes across subperiods. More importantly, a simple pooled historical-mean forecast has lower MAE than all three k NN rules on the common sample. Moving-block intervals for all nine tested k NN rank-correlation specifications include zero. Local similarity in the current price-state representation is therefore not predictive sufficiency.

The allocation demonstration reaches a compatible conclusion. Bayesian randomized ranking and portfolio-aware penalties create adaptive next-dollar paths and a somewhat smaller historical drawdown, but long-run CAGR is essentially the same as equal-weight DCA inside the same strong demonstration universe. Because that universe is ex-post selected and not survivorship-free, the backtest supports the mechanics of the decision architecture rather than an alpha claim.

The practical implication is not to abandon state representations or to treat ticker identity as permanent investment quality. It is to model both. Hierarchical partial pooling offers a direct next step: share information when state evidence is transferable, retain company-specific structure when the data require it, and evaluate the resulting policy under point-in-time data and explicit uncertainty controls.

Code and Data Availability

The empirical analysis was conducted from a frozen research package containing the Python allocation implementation, cached adjusted-price data, the 4,597-row state panel, walk-forward decision scores, monthly allocations, portfolio-weight histories, performance summaries, and figures. A separate validation script reconstructs the state-geometry and predictive-transfer results from the frozen panel and produces the candidate-count-matched geometry, simple forecast benchmarks, moving-block MAE contrasts, and rank-correlation intervals reported in this version. Replication code, validation scripts, configuration files, derived results, and figures are available in the public GitHub repository associated with this paper: <https://github.com/SequentialDecisionResearch/adaptive-capital-allocation-rl>. The repository contains the frozen allocation implementation and the validation routines used to reproduce the state-geometry and predictive-transfer results reported here. Redistribution of raw vendor market data remains subject to the data provider’s terms.

AI-Assisted Research and Editing Disclosure

Generative-AI tools were used for language editing, programming assistance, and implementation support for validation analyses and document preparation. The author reviewed and approved the empirical procedures, executed the reported analyses, verified the numerical outputs against the underlying research data, reviewed the references and interpretations, and is responsible for the final manuscript.

A Detailed Feature Definitions

For stock i at state date t , the analysis panel contains

```
ret_21, ret_63, ret_126, ret_252,
vol_21, vol_63, dd_252,
ma50_gap, ma200_gap,
rel_spy_63, rel_spy_126,
market_spy_63, market_spy_252,
market_vol_63, market_dd_252, qqq_minus_spy_63.
```

Volatility is annualized from daily returns. Drawdown is measured relative to the rolling 252-day maximum. Relative-return variables subtract the corresponding SPY return. The QQQ-minus-SPY variable is a simple proxy for the relative growth/technology market regime.

B Predictive-Transfer Bootstrap

For each k , define the date-level paired error difference

$$\delta_t^{\text{same-cross}} = \text{MAE}_t(\text{same}) - \text{MAE}_t(\text{cross}). \quad (33)$$

The reported point estimate is $\bar{\delta} = T^{-1} \sum_t \delta_t$. To retain approximate dependence induced by overlapping 63-day targets, contiguous three-month blocks of $\{\delta_t\}$ are sampled until T dates are reconstructed. The procedure uses 4,000 resamples and seed 42. Table 8 reports the validation-run intervals used in the revised manuscript.

Table 8: Paired moving-block bootstrap for MAE contrasts

k	Contrast	Mean difference	95% bootstrap interval
5	Same minus cross	-0.27 pp	[-0.96, 0.26] pp
5	Same minus pooled	-0.33 pp	[-1.02, 0.21] pp
12	Same minus cross	-0.44 pp	[-1.17, 0.12] pp
12	Same minus pooled	-0.44 pp	[-1.17, 0.11] pp
30	Same minus cross	-0.41 pp	[-1.03, 0.09] pp
30	Same minus pooled	-0.45 pp	[-1.08, 0.07] pp

Negative values favor same-ticker history. Every interval includes zero. The evidence does not establish cross-ticker predictive superiority, but it also does not prove that pooling is uniformly inferior.

C Nearest-State Example

A nearest-state report for the most recent MU signal provides an intuitive illustration of the representation. The closest historical state is an earlier MU observation, while the next closest states include VRT and other tickers. The example is descriptive only. One neighbor list cannot establish transferability, which is why the main analysis evaluates all eligible stock-months and separately controls for candidate-pool size.

D Final Protocol Checklist

Table 9: Final protocol checklist

Item	Research choice
Price source	Cached adjusted-close series used by the research implementation; no current-fundamental backfill
State dates	Month-end through August 31, 2026; final September 2 snapshot retained for the latest allocation demonstration
Stock-time panel	4,597 rows; 30 demonstration stocks; January 2013–September 2, 2026
Forward target	63-trading-day log return clipped to $[-1, 1]$
Target maturity	Historical label enters training only after its forward horizon has elapsed
Training window	Maximum 8 years; minimum 400 pooled rows
State standardization	Training-sample mean and standard deviation at each decision date
Distance	Euclidean in standardized 16-dimensional state space
Raw geometry evaluation	2,562 comparable nearest-same/cross cases across 89 dates
Matched geometry	Cross-ticker candidate count matched to available same-ticker history; exact without-replacement probability

Item	Research choice
Transfer tests	Pooled, cross-ticker-only, and same-ticker inverse-distance k NN; $k \in \{5, 12, 30\}$
Forecast evaluation	2,572 realized stock-months; 2,549 rows with sufficient same-ticker history
Simple forecast benchmarks	Zero return; pooled historical mean; same-ticker historical mean
Bootstrap	4,000 moving-block resamples; block length 3 monthly decision dates; seed 42
Bayesian model	Gaussian empirical-Bayesian linear model; prior variance 1.0
Thompson score	One common posterior coefficient draw per signal date; master seed 42
k NN blend	25% pooled 30-neighbor estimate, 75% Bayesian estimate
Uncertainty term	Linear composite of Bayesian epistemic SD and k NN SE; used as a heuristic allocation penalty
Portfolio penalties	Volatility, drawdown, current stock weight, sector weight, positive correlation, uncertainty
Capital constraints	12% single-stock cap; 35% sector cap; cash action
Execution	Signal on prior trading day; trade on first trading day of next month
External cash flows	\$20,000 initial capital; \$2,000 monthly contribution
Micro-allocation	20 pieces per deployable monthly budget
Transaction cost	5 bps illustrative purchase friction
Principal internal baseline	Equal-weight DCA over the same eligible demonstration universe
Index benchmarks	SPY and QQQ DCA, descriptive only
Primary validity caveat	Demonstration universe is not point-in-time or survivorship-free

E Rank-Correlation Uncertainty Intervals

Table 10 reports moving-block intervals for the mean monthly cross-sectional Spearman information coefficient. The resampling unit is decision date, with three-month blocks, 4,000 resamples, and seed 42. All nine intervals include zero.

Table 10: Moving-block intervals for mean monthly Spearman IC

k	History restriction	Mean IC	95% interval
5	Pooled all tickers	-0.015	[-0.069, 0.025]
5	Cross-ticker only	-0.005	[-0.056, 0.035]
5	Same ticker only	-0.054	[-0.099, 0.000]
12	Pooled all tickers	-0.021	[-0.080, 0.027]
12	Cross-ticker only	-0.018	[-0.075, 0.030]
12	Same ticker only	-0.044	[-0.096, 0.013]
30	Pooled all tickers	-0.042	[-0.101, 0.009]
30	Cross-ticker only	-0.021	[-0.080, 0.031]
30	Same ticker only	-0.044	[-0.104, 0.021]

References

- Agrawal, S., and N. Goyal. 2013. “Thompson Sampling for Contextual Bandits with Linear Payoffs.” *Proceedings of the 30th International Conference on Machine Learning*, PMLR 28(3):127–135.
- Arnott, R., C. R. Harvey, and H. Markowitz. 2019. “A Backtesting Protocol in the Era of Machine Learning.” *The Journal of Financial Data Science* 1(1):64–74. doi:10.3905/jfds.2019.1.064.
- Ayala, J., J. Axt, C. Sepulveda, E. Canessa, and J. Atkinson. 2025. “A Stock Trading Strategy Based on Deep Reinforcement Learning from Multiple Price Charts.” *The Journal of Financial Data Science* 7(2):62–83. doi:10.3905/jfds.2025.1.182.
- Besbes, O., Y. Gur, and A. Zeevi. 2014. “Stochastic Multi-Armed-Bandit Problem with Non-stationary Rewards.” *Advances in Neural Information Processing Systems* 27:199–207.
- Borkar, S., and A. Jadhav. 2024. “Reinforcement Learning Techniques for Stock Trading: A Survey of Current Research.” *The Journal of Financial Data Science* 6(3):38–56. doi:10.3905/jfds.2024.1.161.
- DeMiguel, V., L. Garlappi, and R. Uppal. 2009. “Optimal Versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy?” *The Review of Financial Studies* 22(5):1915–1953. doi:10.1093/rfs/hhm075.
- Gu, S., B. Kelly, and D. Xiu. 2020. “Empirical Asset Pricing via Machine Learning.” *The Review of Financial Studies* 33(5):2223–2273. doi:10.1093/rfs/hhaa009.
- Han, S. 2023. “Risk Budgeting Portfolio Optimization with Deep Reinforcement Learning.” *The Journal of Financial Data Science* 5(4):86–99. doi:10.3905/jfds.2023.1.137.
- Kelly, B. T., S. Pruitt, and Y. Su. 2019. “Characteristics are Covariances: A Unified Model of Risk and Return.” *Journal of Financial Economics* 134(3):501–524. doi:10.1016/j.jfineco.2019.05.001.
- Li, S. 2026a. “Can Old Market Data Reliably Evaluate a New Policy? A QQQ Contextual-Bandit Study of Historical Memory, Off-Policy Evaluation, and Evidence under Nonstationarity.” SSRN preprint. doi:10.2139/ssrn.7322039.
- Li, S. 2026b. “Historical Evidence for Off-Policy Evaluation of ETF Allocation Policies: When Better Value Estimates Do Not Resolve Policy Choice.” SSRN working paper, September 2026.
- Sutton, R. S., and A. G. Barto. 2018. *Reinforcement Learning: An Introduction*. 2nd ed. Cambridge, MA: MIT Press.
- Zhang, Z., S. Zohren, and S. Roberts. 2020. “Deep Reinforcement Learning for Trading.” *The Journal of Financial Data Science* 2(2):25–40. doi:10.3905/jfds.2020.1.030.

Research and educational paper. Historical results are not evidence of future investment performance and are not investment advice.