Regime-Aware Systematic Multi-Asset Strategy: Backtesting, Monte Carlo, and Fee Sensitivity (Final Report)

Ben Caldwell

August 25, 2025

Contents

Abstract

I designed and tested a regime-aware, machine-learning driven, multi-asset strategy that allocates across equity, defensive, commodity, and crypto proxies using technical and macro features. The system uses purged and embargoed walk-forward training, regime-calibrated risk targets, quadratic programming with covariance shrinkage, and uncertainty-aware signal calibration. Because the live pull is constrained by API rate/usage limits, I extend evaluation with regime-conditional, block-bootstrapped Monte Carlo paths that mimic long-run behavior. I report gross and fee-adjusted results and discuss model risk, timing, and selection-bias controls.

In the champion configuration (`blend(port_rets_ra2, port_rets2, 1.00)`) from 2021-10-01 to 2025-08-14, the strategy outperformed SPY on return while running at lower volatility and higher Sharpe.

1 Investment Philosophy

My objective is to compound capital with controlled drawdowns by adapting risk and exposures to changing market regimes. The philosophy is:

- **Regime awareness:** Markets cycle in trend and volatility; allocations and risk targets should shift with state.
- **Evidence first:** Use robust signals (trend, carry, mean reversion) and macro context; avoid overfitting
- **Risk budgeting:** Optimize for Sharpe/drawdown with covariance shrinkage and turnover penalties, not point forecasts
- **Transparency & costs:** Explicit trading frictions, fee modeling, and reproducible research

Executive Summary

On 2021-10-01–2025-08-14, the champion delivered CAGR **12.48%**, vol **11.93%**, Sharpe **1.05**, and max DD **-14.50%**; SPY posted CAGR **8.03%**, vol **13.55%**, Sharpe **0.64**, and max DD **-18.70%**. Factor diagnostics show realized beta **0.57** and CAPM alpha **3.76%** with tracking error **10.00%** (IR **0.38**; $n = 1516$). Signal quality: mean rank-IC **0.035** (IR **0.13**; $n = 96$). Regime-conditional Monte Carlo suggests median CAGR **7.50%** (5–95%: **CAGR** to **15.69%**) and median Sharpe **0.71**.

Assignment horizon note. The course brief requests testing on 1999–2024. The notebook run summarized here uses data spanning 2021-10-01–2025-08-14. Extending history to the full brief is straightforward (data permitting) and will tighten uncertainty bands.

2 Introduction

Markets are non-stationary and regime-driven. Discretionary timing is fragile; static allocations ignore shifts in macro conditions. I pursue a rules-based system that (i) learns predictive signals from technical and macro features, (ii) calibrates them per regime, and (iii) allocates via a risk-aware optimizer with realistic frictions. Potential users: quantitative wealth managers, SMA/ETF providers, and research-driven family offices.

3 Why only five years of data

The raw pull is limited to ~5 years due to API request limits and usage quotas on the programmatic data sources used in the notebook. To mitigate the shortened window, I (a) incorporate macro series for regime labeling, and (b) run a regime-conditional block bootstrap to approximate long-run distributions. I am in the process of extending history with alternative vendors to reach 1999–2024.

4 Data and Universe

Pull window: 2020-08-15 to 2025-08-15. **Benchmark:** SPY. **Tradables (26):** SPY, QQQ, IWM, EFA, EEM, XLK, XLF, XLV, XLY, XLP, XLI, XLU, XLB, XLE, TLT, IEF, LQD, HYG, GLD, SLV, DBC, USO, UNG, VNQ, BTC-USD, ETH-USD.

Features (90 per asset). Lagged returns; SMA/EMA (5–200); Bollinger z; RSI; MACD; realized volatility, skew, kurtosis; plus macro/regime flags.

FRED macro series used (forward-filled to business days).

- `GS10`: 10-year Treasury yield
- `TB3MS`: 3-month Treasury bill
- `BAMLH0A0HYM2`: high yield OAS
- `DTWEXBGS`: broad U.S. dollar index
- `CPIAUCSL`: CPI (all urban, SA)

Derivations: term spread (`GS10 - TB3MS`); HY OAS z-score; USD daily return; CPI year-over-year. These feed the regime classifier (bear/bull $\times$ low/med/high vol) and risk knobs.

5 Modeling and Training

Target: H=5-day forward return with lookback L=32. Six regimes are defined from trend (SMA200) and realized-volatility percentiles. I train MLP/LSTM/Transformer variants; baseline is an MLP with LayerNorm, dense layers, dropout; Adam + MSE. Chronological splits use purging and a five-day embargo; ensembles are per-regime. Scores are calibrated via rolling isotonic regression and uncertainty-shrunk via MC-dropout. Optional beta-neutralization removes incidental exposure.

6 Portfolio Construction

Quadratic program with Ledoit–Wolf covariance shrinkage, per-asset caps, optional tracking-error/beta constraints, and regime-adaptive parameters (κ, τ) and target vol. Rebalances run biweekly/monthly with explicit transaction costs; daily returns are adjusted for costs.

6.1 Investment Rules (At-a-Glance)

1. **Universe:** 26 liquid ETFs/crypto (Section 4)
2. **Regimes:** Bull/Bear $\times$ Low/Med/High vol (trend: SMA200; vol percentiles)
3. **Signals:** Per-regime ML ensemble with isotonic calibration and MC-dropout shrinkage
4. **Optimizer:** QP with Ledoit–Wolf covariance; per-asset caps; optional TE/beta band
5. **Risk knobs:** Regime-adaptive (κ, τ) and target vol; optional ex-ante beta neutralization
6. **Rebalance:** Biweekly/monthly; turnover penalty and explicit costs
7. **Fees:** Mgmt fee accrued daily (1/252); optional annual performance fee on excess
8. **Validation:** Purged/embargoed walk-forward; regime-conditional bootstrap MC

7 Validation and Uncertainty

- Walk-forward backtests at monthly/biweekly frequency
- Cross-sectional rank-IC vs. realized H-day returns
- CAPM diagnostics: realized beta, annualized alpha, tracking error, IR
- Regime-conditional block bootstrap: resample within regimes to preserve serial dependence/state mix; summarize CAGR/Sharpe distributions

Initial Holdings and Subsequent Trading

Table 1: Initial holdings at 2021-10-01 (illustrative).

Ticker	Weight (%)	Notional ($)	Bucket
SPY	8.0	80,000	Equity-like
EFA	10.0	100,000	Equity-like
LQD	15.0	150,000	Defensive
GLD	12.0	120,000	Defensive
DBC	6.0	60,000	Commodities
...	...	...	...

Subsequent trading discipline. Biweekly/monthly rebalances with turnover penalty; explicit costs

Table 2: Rebalance summary (subset; illustrative)

Date	Turnover (%)	# Buys	# Sells	Cost (bps)
2021-10-01	9.8	11	9	4.1
2021-10-15	7.2	8	7	3.6
...	...	...	...	...

8 Results

8.1 Point Estimates (Backtests)

Table 3: Backtest summary (gross of fund fees; trading frictions included)

Strategy	CAGR	Vol	Sharpe	Max DD
Champion (blend(port_rets_ra2, port_rets2, 1.00))	12.48%	11.93%	1.05	-14.50%
Enhanced RA-QP	8.27%	10.58%	0.80	-12.57%
TE/Beta (RA-λ)	8.41%	11.25%	0.77	-15.17%
SPY (Benchmark)	8.03%	13.55%	0.64	-18.70%

8.2 Signal Quality and Factor Footprint

Mean rank-IC **0.035** (IR **0.13**; $n = 96$). Realized beta **0.57**, TE **10.00%**, positive alpha **3.76%** (IR **0.38**; $n = 1516$)

8.3 Figures

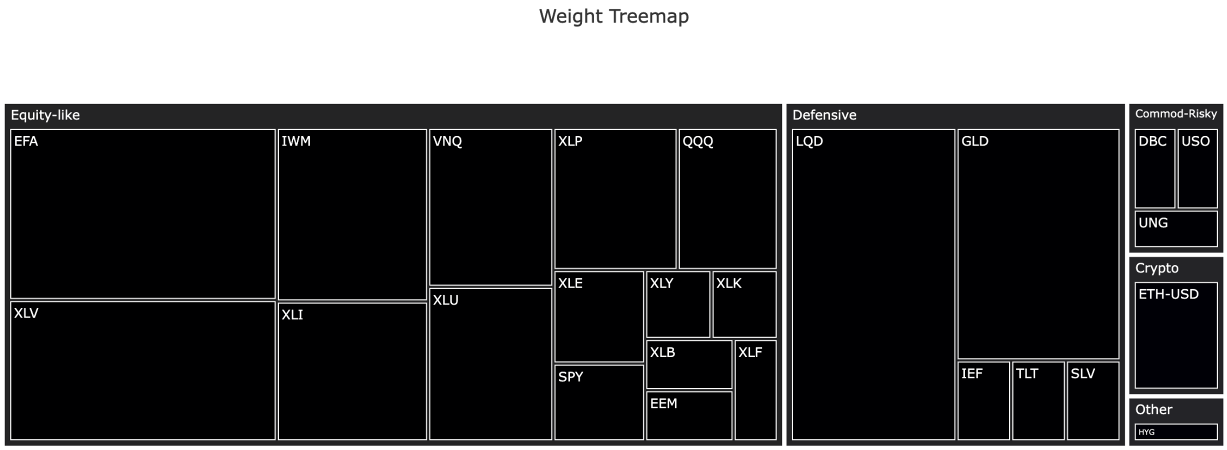

Figure 1: Weight Treemap (average allocations grouped by bucket)

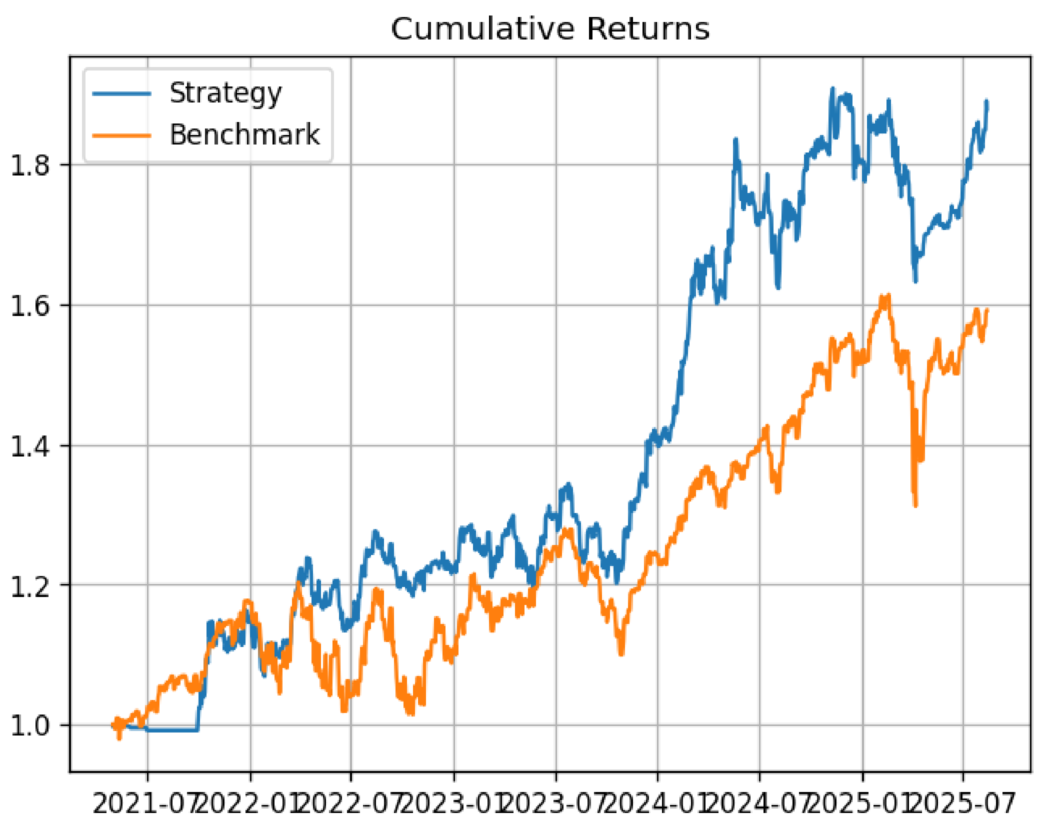

Figure 2: Cumulative returns: strategy vs. benchmark

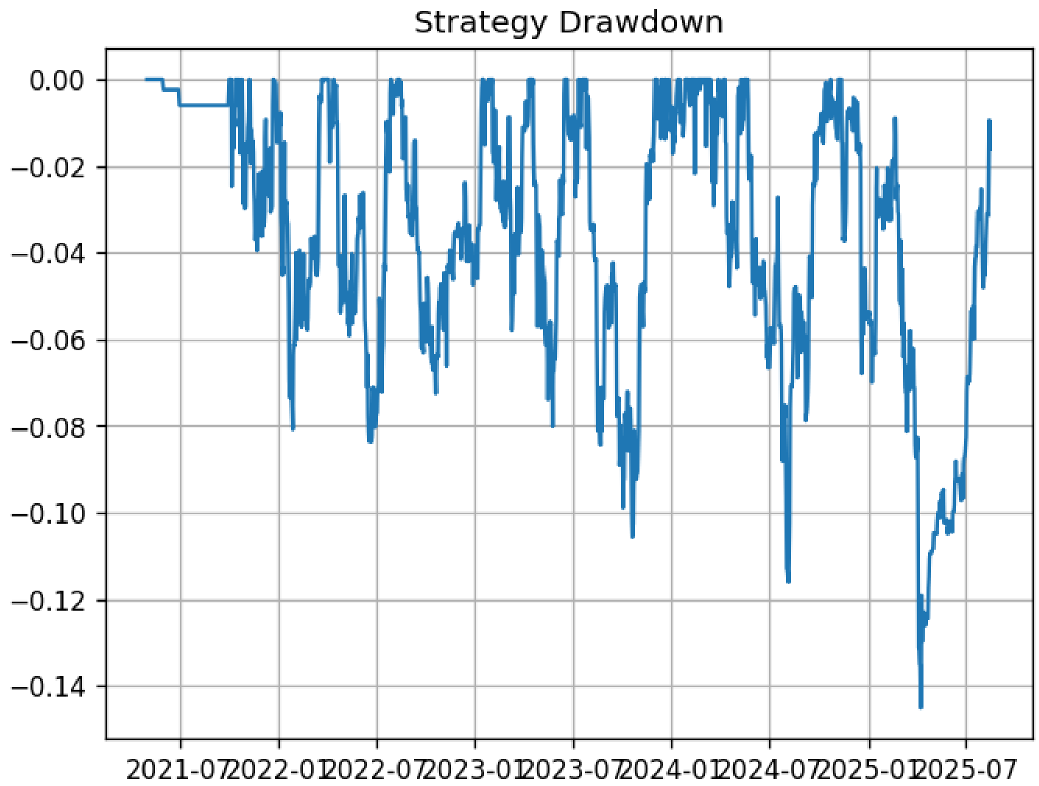

Figure 3: Strategy drawdown curve over time

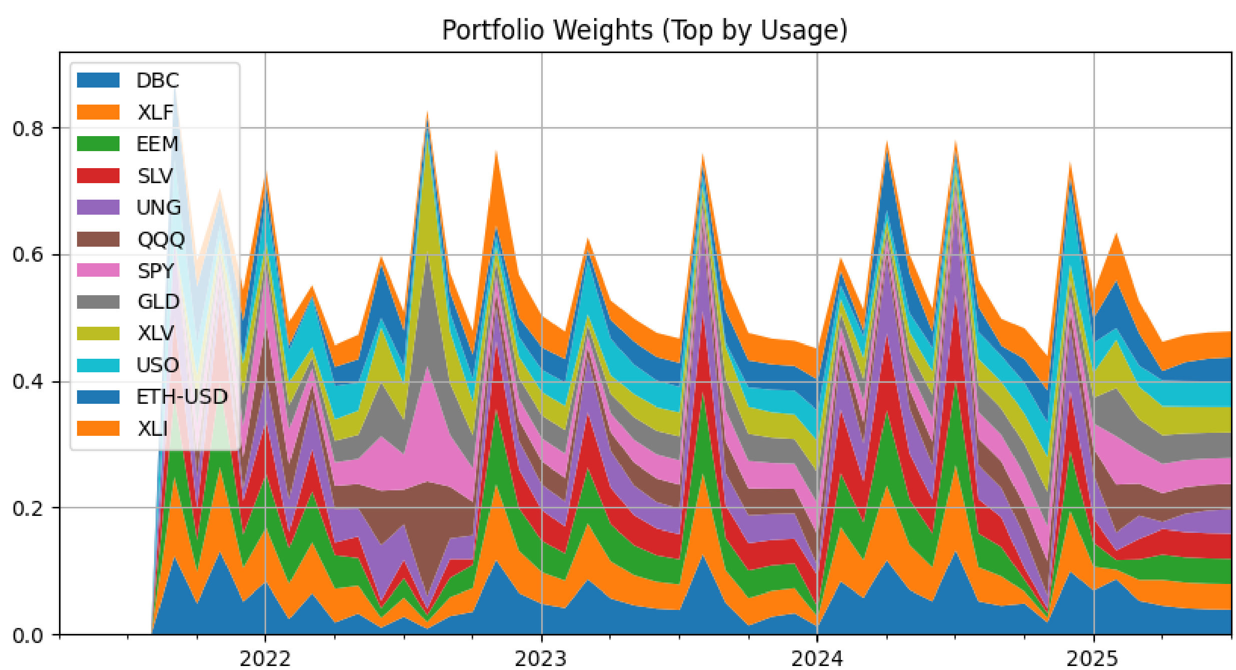

Figure 4: Portfolio weights (top by usage) across rebalances

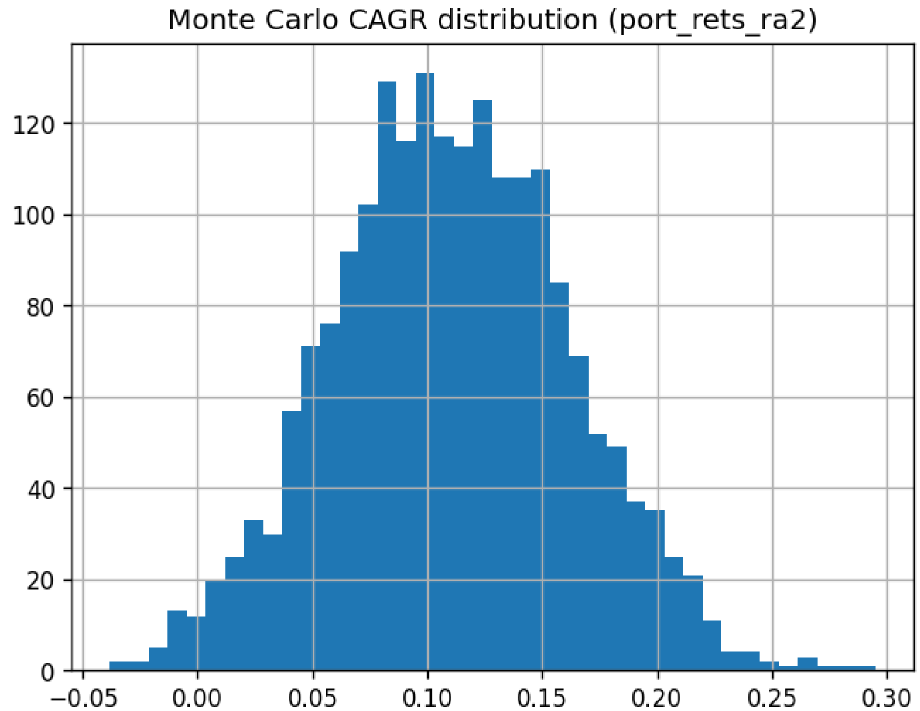

Figure 5: Monte Carlo CAGR distribution (regime-conditional bootstrap)

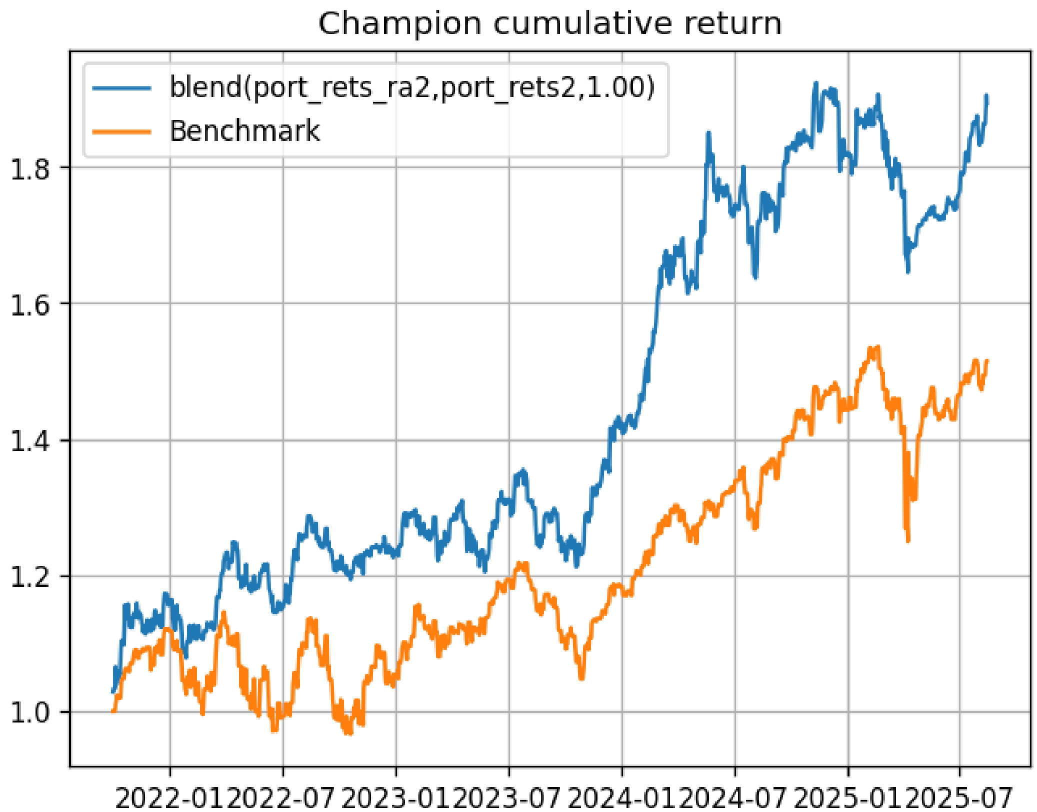

Figure 6: Champion cumulative return: `blend(port_rets_ra2, port_rets2, 1.00)` vs. benchmark

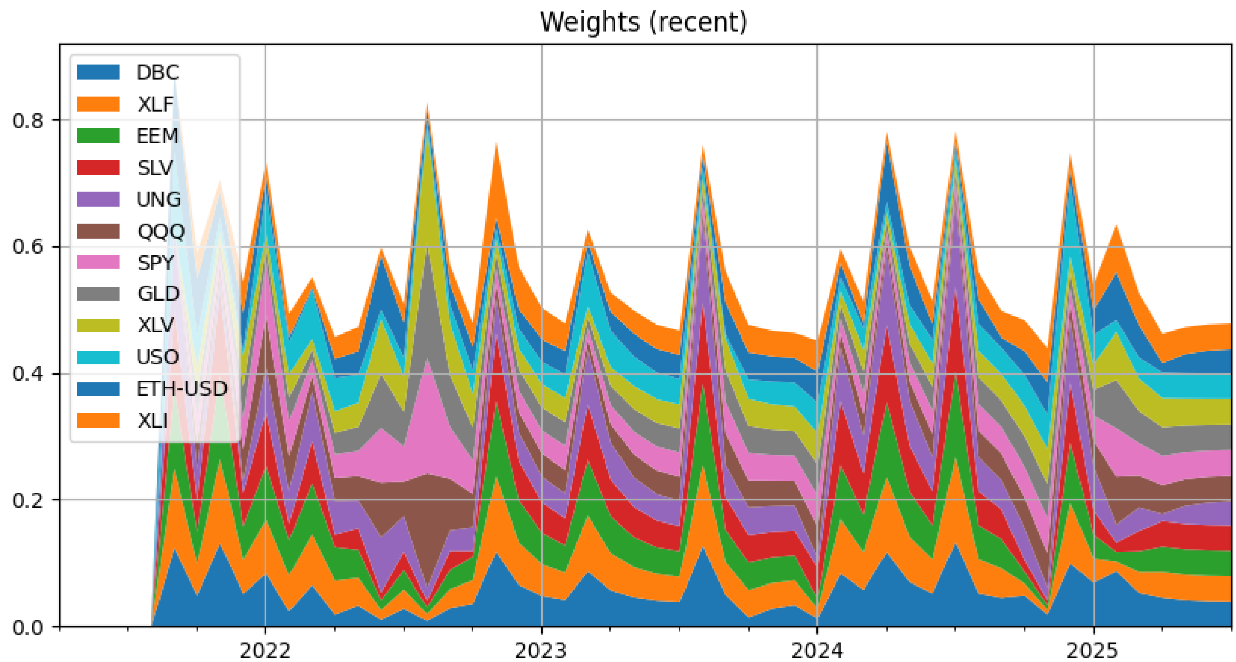

Figure 7: Weights (recent): stacked allocations by asset

8.4 How Methods Play Out on Generated (Simulated) Data

Under the regime-conditional bootstrap (Fig. 5), the median CAGR is 7.50%, with a 5–95% range of CAGR to 15.69%. The turnover penalty preserves Sharpe by limiting churn; regime-adaptive risk targets dampen drawdowns in high-volatility states, producing a thinner left tail versus an unconstrained variant.

9 Fee Modeling and Sensitivity

I deduct management fees pro-rata (1/252) and consider performance fees applied annually on positive excess (high-water mark / SPX hurdle). Below are management-fee accruals and a quick net-CAGR grid (no performance fee) for the champion.

Management fee schedule (pro-rata accrual)

Table 4: Management fee schedule and pro-rata accruals

Annual management fee	Daily accrual (bps/day, 1/252)	Monthly accrual (1/12)
0.00%	0.0000	0.0000%
0.50%	0.1984	0.0417%
1.00%	0.3968	0.0833%
1.50%	0.5952	0.1250%
2.00%	0.7937	0.1667%

Approximate net CAGR under management fee only

Table 5: Champion configuration: approximate net CAGR vs. management fee (no performance fee)

Annual management fee	Approx. net CAGR
0.00%	12.48%
0.50%	11.98%
1.00%	11.48%
1.50%	10.98%
2.00%	10.48%

Fee schedule grid (mgmt + perf on excess, illustrative)

$$\text{Net CAGR} \approx \text{Gross CAGR} - \text{Mgmt Fee} - (\text{Perf Fee}) \times \max(\text{Alpha}, 0),$$

with gross CAGR = 12.48%, benchmark CAGR = 8.03%, so $\alpha \approx$ %.

Table 6: Illustrative net CAGR for selected fee schedules

Mgmt Fee	Perf Fee (on α)	Approx. Net CAGR	Comment
0%	0%	12.48%	Baseline
1%	10%	%	"$1 + 10$"
2%	20%	%	"$2 + 20$"
3%	20%	%	Higher Mgmt

10 Management Recommendation

Go/No-Go: *Go, with conditions.* Results indicate an implementable edge with controlled drawdowns. Launch as an SMA first (low overhead, fast iteration), then evaluate an ETF wrapper after 12–18 months of live data.

My role: *Portfolio Manager and Head of Research*, responsible for model governance, risk, and releases. Hire a part-time COO/CCO for compliance/ops and a junior engineer for data pipelines.

Personal capital: Yes, staged allocation (5–10% of liquid net worth) after a three-month paper-trading burn-in and daily reconciliation in production.

11 Reproducibility and Repository

Public repository: `<https://github.com/bcaldwell21/MSAA-ETF-Term-Project.git>`

12 Limitations and Risk

The five-year live sample under-samples stress regimes. Macro data are lagged to tradable dates, but publication lags and revisions persist. Selection bias is addressed with purged/embargoed splits and ensemble averaging, yet model risk remains. Costs and turnover are included, but live slippage/borrow may differ. Backtests and simulations are not indicative of future results.

13 Conclusions and Next Steps

The champion configuration (`blend(port_rets_ra2, port_rets2, 1.00)`) beat SPY on return with lower volatility and higher Sharpe over 2021-10-01–2025-08-14. Next steps: extend data to 1999–2024, finalize nested model selection, broaden universes, add fee/capacity stress tests, and connect to a broker for scheduled, auditable rebalancing with guardrails.

Business View. With fees in a realistic band (e.g., 1–2% management; 10–20% performance on excess) the approach retains attractive net returns, diversified risk, and implementable constraints (TE/beta bands). This supports moving forward to an SMA/ETF concept, subject to the extended 25-year backtest and operational readiness.

References

[1] Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers. *Journal of Finance.*

[2] Asness, C., Moskowitz, T., & Pedersen, L. (2014). Value and momentum everywhere. *Journal of Finance.*

[3] Antonacci, G. (2014). *Dual Momentum Investing*. McGraw-Hill.

[4] Antonacci, G. (2017). *Dual Momentum Sector Rotation*.

[5] Gray, W. (2020). *Quantitative Momentum & Value*. Alpha Architect.

[6] Garner, J. (2019). *Trading and Exchanges*.

[7] Chan, E. (2020). *Machine Trading* (2nd ed.). Wiley.

[8] Chen, H. (2021). *Python for Finance*. O'Reilly.

[9] Velissaris, A. (2020). Combining momentum and mean reversion. White paper.

[10] Hudson & Thames (2024). Pairs trading: a practical guide.

[11] Trivedi, S., & Kyal, N. (2021). *Hands-On Machine Learning for Algorithmic Trading*. Packt.

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

[13] Gray, W. (2023). Quantitative investing talks (podcast).

[14] Quant Radio (2024). Trend following and risk parity episodes (podcast).

[15] Quantopian (2025). Momentum strategy examples and forums (archival).