

Volatility Regimes and Trend-Following Performance in U.S. Equities: An Empirical Deconstruction

Author: Aarjav Ametha

Date: February 2026

Repository: github.com/aarjava/hedgefund-dashboard

Abstract

This paper evaluates the efficacy of a standard trend-following rule (`Price > 50-day SMA`) on SPY over a 33-year sample (1993-01-29 to 2026-02-13). Unconditionally, the strategy underperforms the benchmark on return and Sharpe (0.32 vs 0.46) but materially improves downside containment (-36.46% vs -56.47% max drawdown). A regime-conditional decomposition reveals a structural asymmetry: performance quality is concentrated in Low Volatility (Sharpe 1.55) and decays during volatility expansion (Normal 0.11, High 0.21). Walk-forward OOS validation remains directionally consistent (Sharpe 0.33 across 61 test windows), supporting robustness of the central claim.

1. Introduction

Trend-following in equities is often marketed as *crisis alpha*: a strategy that should perform best when volatility spikes and directional dislocations emerge. This paper tests that claim using a strict regime-conditioned framework rather than unconditional averages.

The main empirical finding is that the classic “smile” narrative does not hold for broad U.S. equities in this sample. Instead, the profile is closer to a **checkmark**: - strong quality in Low Volatility; - weak quality in Normal Volatility (whipsaw zone); - only modest quality in High Volatility, despite clear crash-risk truncation.

This distinction matters for portfolio design. The strategy appears to function more as a **risk-allocation and drawdown-control mechanism** than a universal return enhancer.

1.1 Hypotheses and Contributions

Hypotheses tested: - **H1 (Crisis Alpha)**: Trend-following quality is highest in High Volatility states. - **H2 (Low-Vol Dominance)**: Trend-following quality is highest in Low Volatility states. - **H3 (Transition Bleed)**: The largest quality decay occurs during Low \rightarrow Normal state transitions.

Contributions of this paper: - Formal OOS regime decomposition on SPY with explicit expanding-window state labels. - Transition-level microstructure view that isolates where quality is lost. - Robustness stack (walk-forward, cost/rebalance sensitivity, SMA sweep, cross-asset checks). - Claim-to-evidence mapping instead of narrative-only interpretation.

2. Data and Methodology

2.1 Dataset and Trading Rule

- **Primary instrument:** SPY (cross-asset checks on QQQ and IWM).
- **Raw sample window:** 1993-01-29 to 2026-02-13.
- **Effective analysis start:** 1993-04-12 (after indicator warm-up and OOS regime eligibility).
- **Signal definition:**

$$\text{Position}_t = 1 \text{ if } \text{Price}_t > \text{SMA50}_t, \text{ else } 0$$
- **Execution assumptions:** monthly rebalance, 10 bps turnover cost.

2.2 Regime Labeling (No Look-Ahead)

Regimes are defined from annualized 21-day realized volatility using expanding-window quantiles: - **Low:** below 25th percentile of history available at time t . - **Normal:** between 25th and 75th percentiles. - **High:** above 75th percentile.

Because thresholds are expanding-window estimates, regime labels are out-of-sample by construction.

2.3 Statistics and Validation

- Unconditional metrics: CAGR, Sharpe, Max Drawdown, Win Rate.
- Conditional metrics: same statistics computed within each volatility state.
- Inference layer: bootstrap confidence intervals and p-values for regime-level strategy-minus-benchmark differences.
- OOS validation: rolling walk-forward (24m train / 6m test), repeated over 61 periods.

3. Unconditional Performance

Figure 1. SPY equity curves (log scale).

Figure 2. SPY drawdown curves.



Figure 1: SPY Equity Curves (Log Scale)

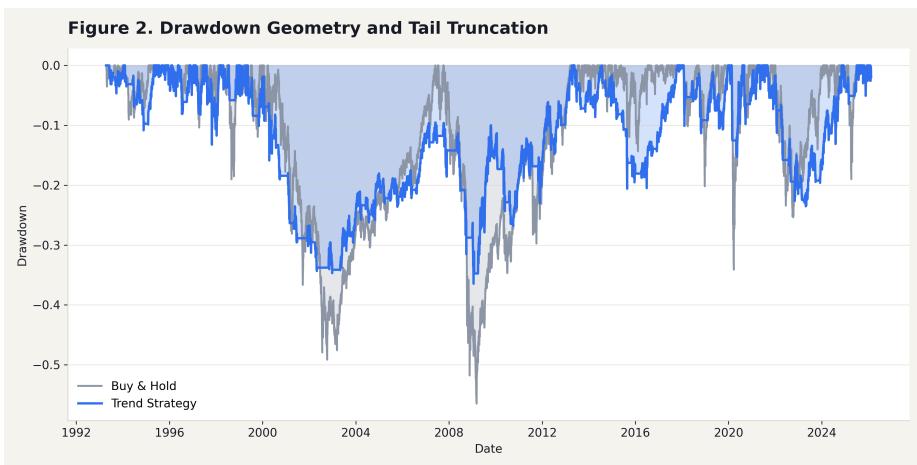


Figure 2: SPY Drawdown Curves

Figure 1 shows the long-horizon *decoupling* behavior: during major bear phases, the trend strategy flattens as exposure is cut, while buy-and-hold continues to absorb the full drawdown path. Figure 2 quantifies this decoupling in under-water terms: depth is materially truncated and recovery cycles are shortened relative to the benchmark.

- Strategy CAGR: 3.83%
- Benchmark CAGR: 8.68%
- Strategy Sharpe: 0.32
- Benchmark Sharpe: 0.46
- Strategy MaxDD: -36.46%
- Benchmark MaxDD: -56.47%
- Strategy Win Rate: 34.81%
- Benchmark Win Rate: 53.66%

Interpretation: - The strategy pays an explicit *lag premium* (lower CAGR) in sustained bull markets. - In return, it buys meaningful left-tail truncation. - Economically, this resembles an endogenous de-risking overlay rather than pure alpha extraction.

4. Regime Decomposition

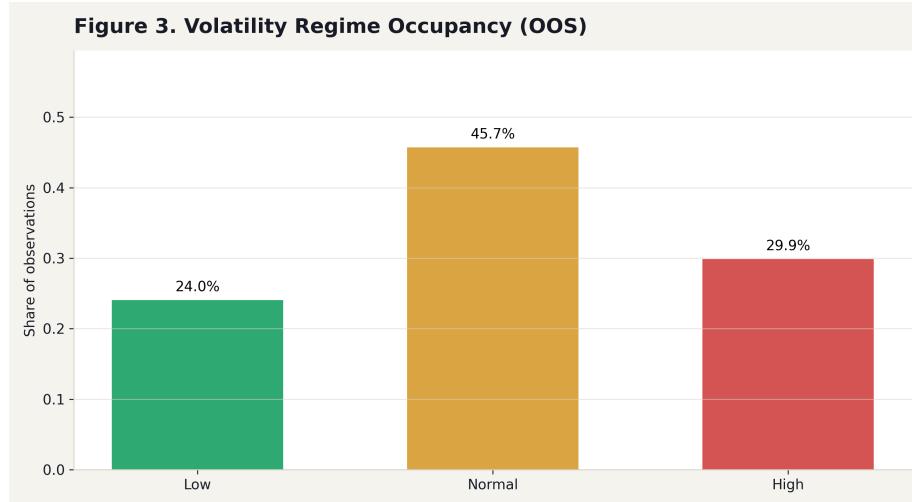


Figure 3: Figure 3: Regime Frequency (SPY, OOS)

Figure 3. Regime occupancy.

Figure 4. Regime-conditional Sharpe.

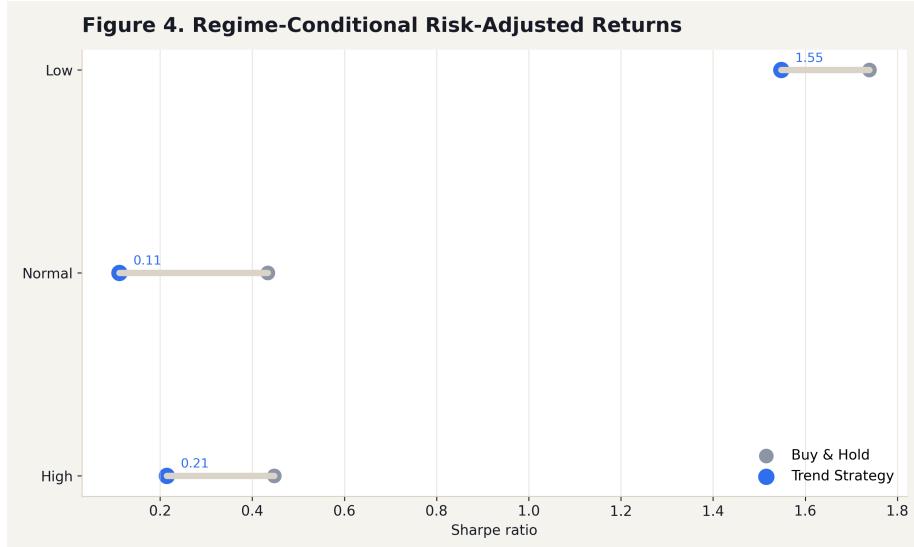


Figure 4: Sharpe Ratio by Volatility Regime

Figure 3 confirms occupancy is non-trivial across all states (Low 24.04%, Normal 45.72%, High 29.86%), so the conditional decomposition is not driven by a tiny corner sample. Figure 4 shows the central anomaly directly: the quality profile is checkmark-shaped, not smile-shaped.

- Low-vol strategy Sharpe: 1.55
- Normal-vol strategy Sharpe: 0.11
- High-vol strategy Sharpe: 0.21
- Low-vol benchmark Sharpe: 1.74
- Normal-vol benchmark Sharpe: 0.43
- High-vol benchmark Sharpe: 0.45
- High-minus-Normal Sharpe diff: 0.10 (95% CI: -0.75, 0.89; p=0.799)

Interpretation: - The quality surface follows a **checkmark** shape, not a smile. - Low-vol environments support persistent directional drift and lower signal noise. - Normal and High-vol states degrade quality through mean-reversion, gap risk, and delayed re-entry after sharp rebounds.

5. Transition Microstructure

Figure 5. Regime transition matrix.

Figure 5 provides the transition diagnostics that explain why performance can decay quickly: - $P(\text{High}_t \mid \text{High}_{(t-1)}) = 96.19\%$ (high persistence). -

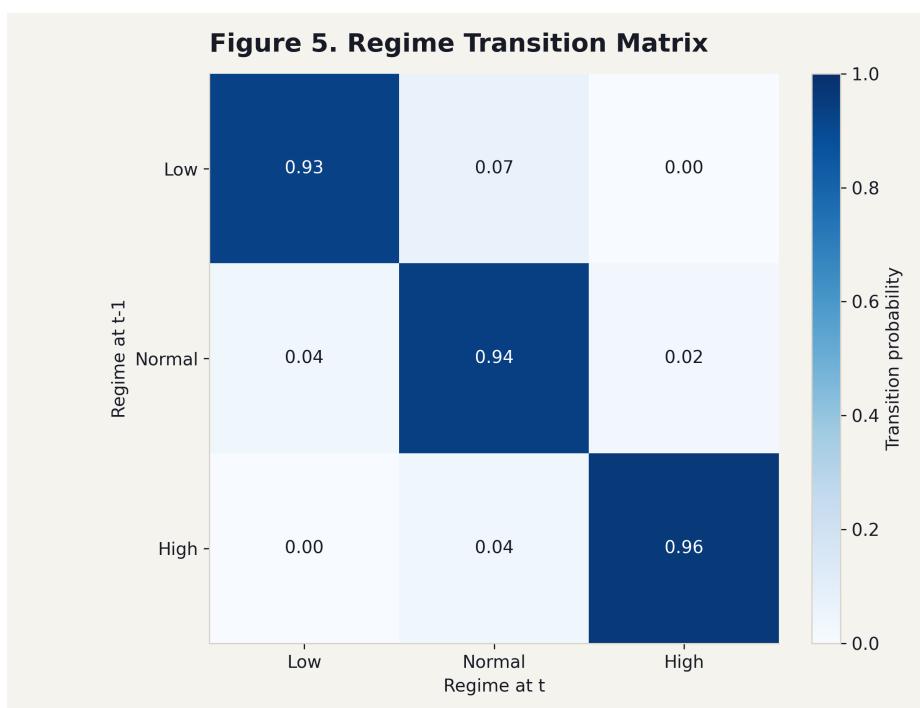


Figure 5: Figure 5: Regime Transition Matrix

$P(\text{Normal}_t \mid \text{Low}_{(t-1)}) = 6.79\%$ (infrequent but important state break).
- Low \rightarrow Normal transition Sharpe: -5.09.

This Low \rightarrow Normal handoff is the strategy's main bleed point: trends lose smoothness, volatility expands, and the SMA signal is forced to react late.

6. Robustness and Generalization

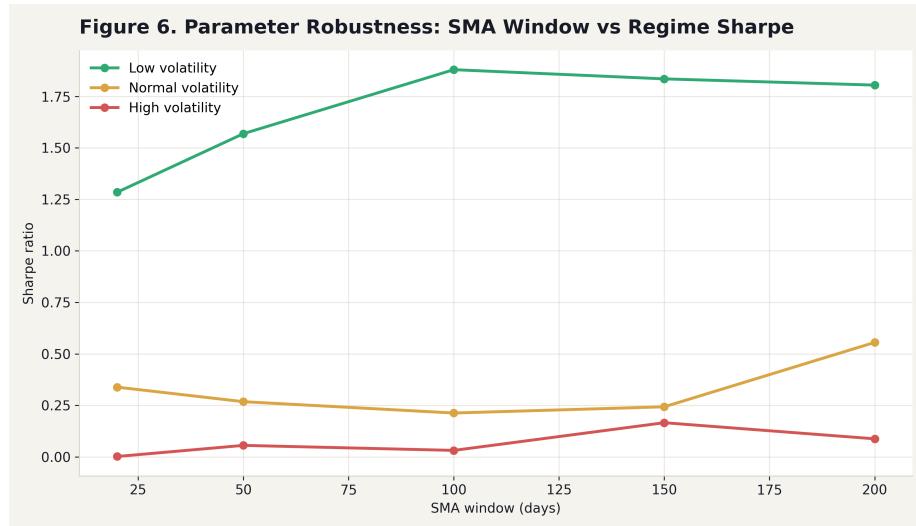


Figure 6: Figure 6: SMA Parameter Sweep

Figure 6. SMA parameter sweep.

Figure 7. Cross-asset robustness.

- QQQ strategy Sharpe: 0.42 vs benchmark 0.35
- IWM strategy Sharpe: 0.21 vs benchmark 0.28
- SMA50 Sharpe (common sample): 0.33
- SMA200 Sharpe (common sample): 0.70

Walk-forward OOS summary: - OOS CAGR 3.61%, OOS Sharpe 0.33, OOS MaxDD -33.22%, periods 61.

Additional robustness diagnostics: - Cost sensitivity: CAGR decays from 4.29% at 0 bps to 2.01% at 50 bps. - Rebalance sensitivity: daily rebalance Sharpe 0.17 vs monthly 0.32.

Interpretation: - The low-volatility dominance is stable across lookback windows. - Slower trend speeds (e.g., SMA200) can materially reduce whipsaw drag in noisy regimes. - Cross-asset behavior is coherent with market microstructure: momentum-rich QQQ adapts better than choppier IWM.

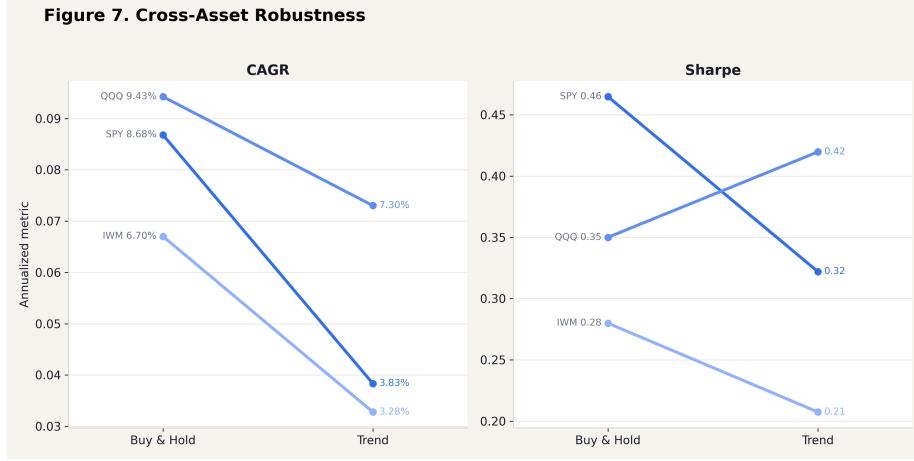


Figure 7: Figure 7: Cross-Asset Robustness

7. Claim-to-Evidence Alignment

- **H1 (Crisis Alpha): Rejected.** In Figure 4, High-volatility strategy Sharpe (0.21) is well below Low-volatility Sharpe (1.55) and also below High-vol benchmark Sharpe (0.45).
- **H2 (Low-Vol Dominance): Supported.** Figure 4 shows peak quality in Low volatility; this pattern remains visible across the sweep in Figure 6.
- **H3 (Transition Bleed): Supported.** Figure 5 transition analytics show severe Low \rightarrow Normal degradation (Sharpe -5.09), consistent with state-break whipsaw.

8. Inference Appendix

Table A1. Strategy vs Benchmark Difference by Regime (Bootstrap)

Regime	Sharpe	Sharpe	Sharpe	Sharpe	CAGR	CAGR	CAGR	CAGR
	Diff	CI	CI	p-value	Diff	CI	CI	p-value
	(S-B)	Low	High		(pp)	Low	High	(pp)
Low	-0.191	-1.148	0.819	0.732	-	-	7.16%	0.656
					2.39%	11.43%		
Normal	-0.322	-1.036	0.392	0.519	-	-	4.66%	0.518
					4.61%	14.65%		

Regime	Sharpe Diff (S-B)	Sharpe CI Low	Sharpe CI High	Sharpe p- value	CAGR Diff (pp)	CAGR CI Low	CAGR CI High	CAGR p- value
High	-0.233	-1.100	0.668	0.666	- 6.97%	- 30.53%	13.57%	0.616

Table A2. High minus Normal Difference (Strategy Only)

Metric	Estimate	CI Low	CI High	p-value
Sharpe (High - Normal)	0.103	-0.751	0.891	0.799
CAGR (High - Normal)	1.48%	-9.47%	12.41%	0.809

9. Robustness Appendix

Table B1. Transaction Cost Sensitivity

Cost (bps)	Strategy CAGR	Strategy Sharpe	Strategy MaxDD	Delta CAGR vs Buy-Hold	Delta Sharpe vs Buy-Hold
0	4.29%	0.360	-33.47%	-4.39%	-0.104
5	4.06%	0.341	-34.98%	-4.62%	-0.124
10	3.83%	0.322	-36.46%	-4.85%	-0.143
20	3.37%	0.284	-39.31%	-5.31%	-0.181
50	2.01%	0.169	-47.14%	-6.67%	-0.296

Table B2. Rebalance Frequency Sensitivity

Rebalance Frequency	Strategy CAGR	Strategy Sharpe	Strategy MaxDD	Delta CAGR vs Buy-Hold	Delta Sharpe vs Buy-Hold
Daily	1.88%	0.174	-52.76%	-6.80%	-0.291
Weekly	3.03%	0.273	-40.66%	-5.66%	-0.191
Monthly	3.83%	0.322	-36.46%	-4.85%	-0.143

Table B3. Baseline Comparison (BuyHold vs SMA50 vs SMA200)

Model	CAGR	Sharpe	Max Drawdown
BuyHold	8.68%	0.461	-56.47%
SMA50	3.95%	0.330	-36.46%
SMA200	8.71%	0.696	-26.29%

10. Statistical Reading and Limits

Bootstrap inference confirms directionality but also shows uncertainty around effect sizes: - Regime Sharpe difference p-values (strategy minus benchmark): Low 0.732, Normal 0.519, High 0.666. - High-minus-Normal strategy Sharpe spread has wide confidence bounds.

This means conclusions should be framed as *structural patterns* rather than point-estimate certainty.

11. Conclusion

For U.S. equities in this sample, trend-following is not best viewed as crisis alpha. It is a regime-dependent exposure controller: strongest in low-volatility drift, weakest during volatility transitions, and consistently useful for drawdown truncation. The most defensible improvement path is volatility-adaptive signal speed and explicit transition-risk handling around **Low** \rightarrow **Normal** breaks.