Momentum Returns Across Volatility and Dispersion Regimes: Evidence from the S&P 500

Ed Zhao

May 27, 2025

Contents

1 Introduction **2**

2 Data and Universe Construction **3**

3 Momentum Factor Definition **3**

4 Volatility and Dispersion Regime Classification **3**

5 Methodology **3**

5.1 Cross-Sectional Return Sorting 4

5.2 Walk-Forward Validation 4

5.3 Bootstrap Testing 4

6 Results **4**

6.1 Baseline Comparison 5

6.2 Sensitivity Analysis 6

7 Discussion and Limitations **6**

8 Conclusion **7**

Abstract

We examine whether market volatility and cross-sectional return dispersion influence the stability of momentum factor returns. Using daily S&P 500 data from 2010 to 2025, we construct regime labels based on volatility and dispersion, applying static, rolling, and expanding quantile definitions. Factor returns are computed via a decile-spread momentum strategy using 12-month returns that skip the most recent month, and evaluated using walk-forward validation with bootstrapped confidence intervals.

Results show that momentum returns are directionally stronger in low-volatility regimes under expanding quantile definitions, while dispersion-based conditioning yields less consistent patterns. No return differentials are statistically significant. While statistical significance is limited, the consistency of directional results under volatility regimes warrants further study.

1 Introduction

Momentum signals are widely used in equity factor models, but their performance is known to vary across market conditions [1, 2, 3]. Prior work has shown that momentum returns are particularly vulnerable to sharp reversals during volatile periods, motivating efforts to condition or dynamically adjust momentum exposure based on volatility signals.

One hypothesis is that these signals are more stable and effective in environments with low market volatility or low cross-sectional dispersion—conditions under which price trends may be less disrupted by macroeconomic shocks or investor disagreement.

This project investigates how equity momentum returns differ across regimes defined by both market volatility and cross-sectional return dispersion. Unlike prior studies that rely on a single regime construct or static thresholding, we evaluate three regime methodologies—static, rolling, and expanding quantiles—applied to both volatility and dispersion measures. This allows us to assess the robustness of conditional momentum factor behavior across both regime variables (volatility vs. dispersion) and quantile estimation methods (static, rolling, expanding). Static regimes use full-sample quantiles (non-causal); rolling regimes use fixed-length trailing windows; expanding regimes grow cumulatively and are causal.

We compute long-short decile spread returns using a momentum signal defined as 12-month return skipping the most recent month, applied to a universe of S&P 500 equities from 2010 to 2025. To ensure causality and reflect real-world latency constraints, regime labels assigned to returns on day t are based on volatility and dispersion measured at day $t - \text{lag}$. Returns are validated via walk-forward analysis using rolling test windows, with bootstrap inference applied to assess statistical significance and stability.

Our results show that momentum returns are generally stronger in both low-volatility and low-dispersion regimes, but that none of the return differentials are statistically significant. Some comparisons also yielded inverted results. The sensitivity of results to regime type and definition highlights the need for caution in drawing inference from conditional factor performance. This instability echoes challenges noted in prior studies, which have proposed dynamic volatility scaling [2] or crash-risk modeling [3] as more robust alternatives.

2 Data and Universe Construction

We use a universe of S&P 500 constituents with daily adjusted prices from EODHD covering 2010 to 2025. Stocks are required to have $\geq$85% non-missing price coverage. No minimum price filter was applied given the inherent liquidity of S&P 500 constituents. The coverage filter resulted in a final universe of 450 stocks.

3 Momentum Factor Definition

We define momentum as the relative price change between two lagged points in time, following the formula:

$$\text{Momentum}_t = \frac{P_{t-21}}{P_{t-252}} - 1 \tag{1}$$

where P_{t-21} is the price lagged by 21 trading days (approximately 1 month), and P_{t-252} is the price lagged by 252 trading days (approximately 1 year). This specification avoids using recent returns to mitigate short-term reversal effects.

Momentum values are computed cross-sectionally for all eligible stocks using adjusted close prices. The resulting factor is then used to sort stocks into deciles for performance evaluation under different market regimes.

4 Volatility and Dispersion Regime Classification

We construct two types of market regimes:

- **Volatility Regime:** Based on the rolling 20-day standard deviation of average index-level log returns. For each quantile estimation method—static, rolling, or expanding—volatility values are assigned high or low labels using quantile thresholds computed with the corresponding approach.

- **Dispersion Regime:** Based on the rolling 20-day average of cross-sectional standard deviation of log stock returns across the universe on each day. As with volatility, quantile thresholds for high and low dispersion states are defined using static, rolling, or expanding windows depending on the regime specification.

The static method computes quantile thresholds using the full sample period, which introduces lookahead bias by incorporating future information into regime classification. On the other hand, rolling quantile definitions produce unstable regime labels due to high sensitivity to short-term noise, particularly during volatile periods with erratic return dispersion. In contrast, the expanding method avoids lookahead bias and offers greater label stability by incorporating only past data in a progressively growing window.

5 Methodology

We evaluate the conditional performance of the momentum factor using walk-forward validation from January 2010 through May 2025. To preserve causality and reflect implementation delays, each day is classified using the regime label from 5 trading days earlier (i.e., the regime state at $t-5$ is used on day t).

Throughout, we focus on the expanding quantile method and define low and high regimes using the 30th and 70th percentiles, respectively. This choice reflects the superior temporal validity and stability of expanding quantile definitions: static regimes incorporate future data, and rolling regimes yield unstable labels.

5.1 Cross-Sectional Return Sorting

On each trading day:

- Stocks are ranked by momentum.
- A long-short return is computed: average return of top decile minus bottom decile.
- Only test days in the target regime are retained.

5.2 Walk-Forward Validation

- No training window is used (0-day), as there is no predictive model or parameter estimation.
- Each test segment spans 21 days.
- The process advances in 21-day fixed steps to reduce computational load without significantly compromising sampling frequency.
- Factor returns are evaluated only on test days where the delayed regime matches the target.

5.3 Bootstrap Testing

- Bootstrapping is applied to long-short return spreads to assess statistical significance. 95% confidence intervals and p-values are computed for return differences between regimes.
- Testing used 1000 bootstrap iterations. For each regime, samples were drawn with replacement, matching the length of the original series.

After applying coverage filters, constructing regimes, and applying the 5-day delay, the final number of testable trading days ranged from 918 to 1373 depending on the regime variable (volatility vs. dispersion) and quantile method (static, rolling, expanding).

This setup defines our baseline configuration. In subsequent analyses, we vary the regime lookback window, test window length, and lag (delay) days to evaluate sensitivity across parameter choices.

6 Results

Figure 1 shows that long-short momentum returns are directionally stronger in low-volatility regimes, with the bootstrapped 95% confidence interval fully above zero. In contrast, Figure 2 illustrates that dispersion-based regime conditioning yields less consistent results, with wide confidence intervals and negligible separation between high and low regimes.

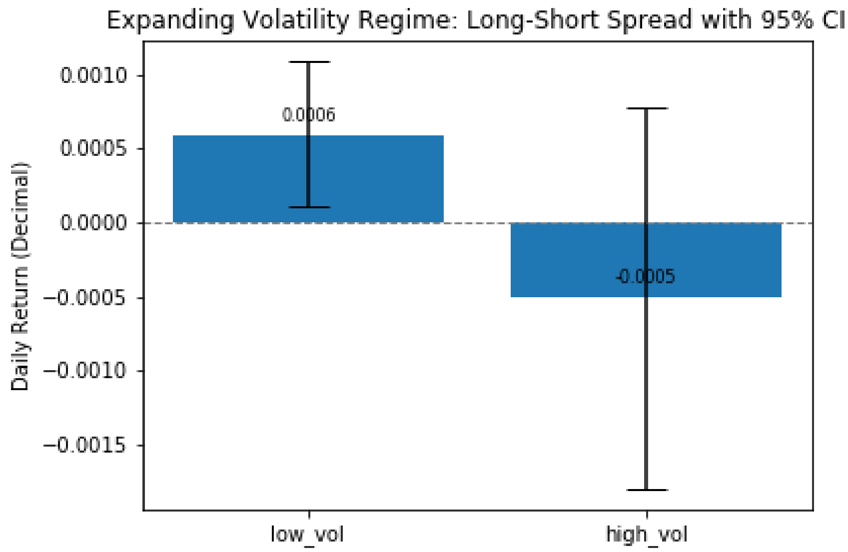


Figure 1: Expanding Volatility Regime: Long-short momentum returns with 95% CI.

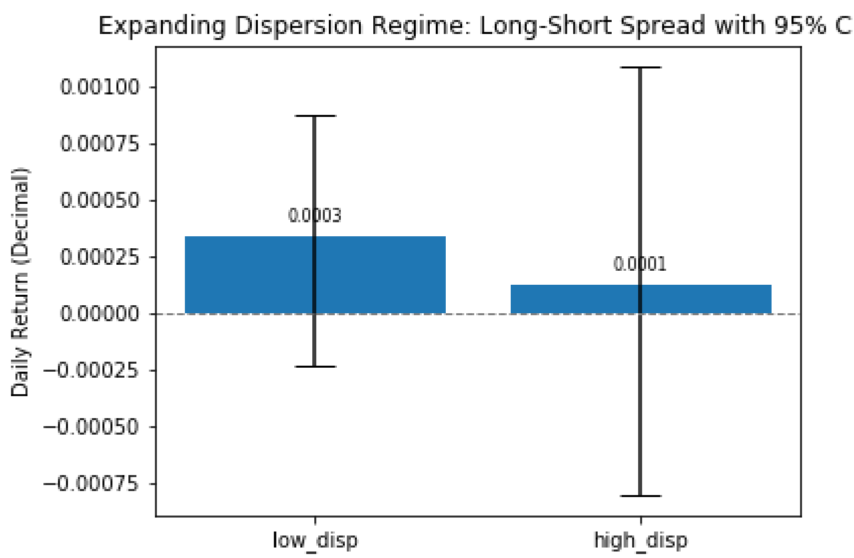


Figure 2: Expanding Dispersion Regime: Long-short momentum returns with 95% CI.

6.1 Baseline Comparison

We report bootstrapped long-short return statistics across expanding regime definitions for both volatility and dispersion. All results use walk-forward validation (21-day test) with a 5-day regime delay to ensure causality. See Table 1 for regime-specific mean returns, confidence intervals, and sample sizes.

For volatility regimes, the low-volatility group outperforms the high-volatility group with a mean difference of 0.001092. The bootstrapped confidence interval is wide (CI: [-0.001265, 0.001168]), and the p-value of 0.077 indicates marginal significance.

Table 1: Summary Statistics by Regime (Expanding Quantile Definition)

Regime Type	Regime	Mean Return	95% CI	N	Total Days	p-value
Volatility	Low	0.000585	[0.000083, 0.001119]	1269	3843	
Volatility	High	-0.000507	[-0.001737, 0.000686]	918	3843	
Volatility	Low - High	0.001092	[-0.001265, 0.001168]	1000	—	0.077
Dispersion	Low	0.000338	[-0.000172, 0.000876]	939	3843	
Dispersion	High	0.000128	[-0.000873, 0.001131]	1373	3843	
Dispersion	Low - High	0.000210	[-0.001186, 0.001166]	1000	—	0.743

For dispersion regimes, the mean difference is smaller (0.000210), with a p-value of 0.743, suggesting no meaningful effect.

All reported p-values are based on two-sided bootstrap tests. However, under a one-sided hypothesis aligned with the expectation that momentum returns are stronger in low-volatility regimes, the expanding volatility result crosses the conventional significance threshold of 5%. All other regime comparisons remain statistically insignificant.

6.2 Sensitivity Analysis

We tested regime delays of 0, 5, and 10 days to reflect different assumptions about recognition delay. These were crossed with volatility lookback windows of 10, 21, 63 days, and return test windows of 5, 10, and 63 days, to evaluate robustness across regime definitions and holding periods without excessive computational burden. These values were chosen to span common trading horizons: short-term (5–10 days), intermediate (21 days), and monthly-to-quarterly cycles (63 days $\approx$ 3 months).

Note: While infeasible, a delay of 0 was included for testing purposes. Also, some walk-forward configurations use a test window shorter than the step size, resulting in non-overlapping return windows and fewer total evaluations. This does not affect the validity of bootstrap estimates.

Figure 3 in Appendix A shows that conditioning momentum returns on volatility regimes generally produces higher return spreads across a wide range of parameter configurations. In contrast, Figure 4 illustrates that dispersion-based regimes yield less consistent effects, with several parameter combinations producing inverted or negligible differentials.

This reinforces the view that volatility is a more stable and informative conditioning variable than dispersion for momentum-based strategies.

7 Discussion and Limitations

Sensitivity analysis indicates that volatility-based regime conditioning tends to produce relatively robust performance differentials. In contrast, dispersion-based regimes exhibit greater variability, with more parameter settings producing inverted or negligible effects. In most comparisons, momentum returns were directionally higher in low-volatility and low-dispersion regimes, but differentials were not statistically significant under two-sided testing. One-sided tests yielded borderline significance only in the expanding volatility regime, suggesting regime states may influence factor performance, but not reliably enough to serve as standalone signals.

The weaker performance under dispersion-based conditioning likely stems from the higher noise inherent in cross-sectional return spread metrics, which is often microstructure-driven. Dispersion fluctuates with sector rotation and idiosyncratic shocks, lacking the systemic coherence of volatility, which aggregates index-level stress and reflects broad risk sentiment. Consequently, volatility offers more stable and interpretable conditioning across momentum factor strategies. This aligns with prior literature noting dispersion's limited predictive utility outside of severe dislocation periods.

The study is limited by its reliance on daily data, a non-predictive decile-based sorting model, and the restriction to the S&P 500 universe. In high-volatility regimes, reduced coverage and sample instability likely contributed to result noise. Using delayed regime labels helps preserve causality but further reduces usable sample size.

Future work will incorporate additional signals, such as reversal and volatility exposure. We also plan to test predictive models, including machine learning, to estimate regime-dependent factor payoffs. Finally, we aim to evaluate the framework on international markets and in broader regime-aware portfolio construction.

8 Conclusion

Momentum signals appear more stable and consistent in low-volatility regimes, despite a lack of statistical significance. In contrast, dispersion-based regimes yield weaker and less reliable differentiation. This highlights the potential value of volatility-aware conditioning in factor timing.

References

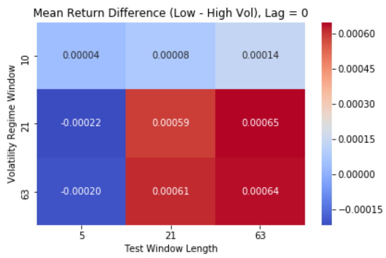
[1] Tobias J Moskowitz, Yao Hua Ooi, and Lasse Heje Pedersen. Time series momentum. *Journal of Financial Economics*, 104(2):228–250, 2012.

[2] Pedro Barroso and Pedro Santa-Clara. Momentum has its moments. *Journal of Financial Economics*, 116(1):111–120, 2015.

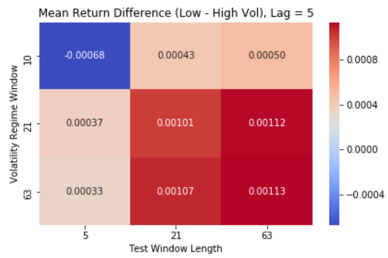
[3] Kent Daniel and Tobias J. Moskowitz. Momentum crashes. *Journal of Financial Economics*, 122(2):221–247, 2016.

Appendix

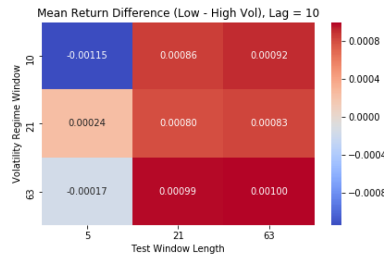
Regime Sensitivity Heatmaps (Grid Analysis)



(a) Vol. Lag 0

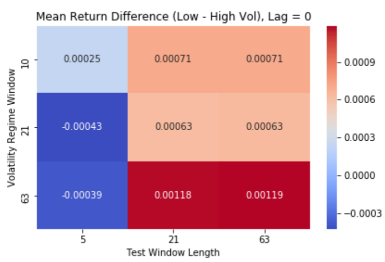


(b) Vol. Lag 5

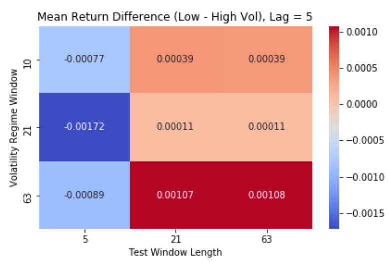


(c) Vol. Lag 10

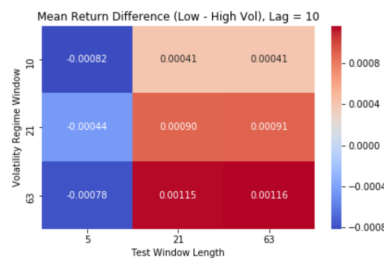
Figure 3: Momentum: Mean Return Difference by Volatility Regime



(a) Disp. Lag 0



(b) Disp. Lag 5



(c) Disp. Lag 10

Figure 4: Momentum: Mean Return Difference by Dispersion Regime