Do Momentum Portfolio Returns Exhibit Abnormal Performance After Controlling for Fama-French Five Factors?

A Regime-Dependent Analysis of Factor-Adjusted Returns

Syed Bashir Hydari   Pablo Rocha Gomez   Amaan Khan

2025-12-16

1) Introduction

Momentum—the tendency for assets with strong recent performance to continue outperforming—represents one of the most persistent yet puzzling phenomena in financial markets. Since Jegadeesh and Titman's (1993) seminal work documenting this anomaly, momentum strategies have generated consistent positive returns across markets and time periods, challenging the efficient market hypothesis and prompting extensive academic inquiry. Despite decades of research, a fundamental question remains unresolved: do momentum returns represent genuine abnormal performance, or are they simply compensation for exposure to established risk factors?

This study examines whether momentum portfolio returns exhibit statistically and economically significant alpha after controlling for the Fama-French Five-Factor model, the current consensus framework in empirical asset pricing. Using monthly U.S. equity data spanning January 2010 through December 2024, we test whether momentum's apparent profitability persists after accounting for market, size, value, profitability, and investment factors. Our analysis encompasses 180 monthly observations, capturing a period of substantial market volatility including the European debt crisis, multiple Federal Reserve policy shifts, the COVID-19 pandemic, and recent inflationary pressures.

Beyond testing for unconditional abnormal returns, we extend the analysis to investigate whether momentum's effectiveness varies with market conditions. Specifically, we examine whether momentum's alpha and factor exposures differ across volatility regimes—periods of high versus low market turbulence. This regime-dependent perspective recognizes that momentum strategies may

exhibit asymmetric behavior during calm versus turbulent markets, potentially reflecting behavioral biases that intensify during specific market states. If momentum's abnormal returns prove state-dependent, such findings would have important implications for dynamic portfolio allocation and risk management.

Our exploratory analysis reveals intriguing patterns. Momentum (UMD) earned an average monthly return of 0.25% with volatility of 3.57% over the sample period. Notably, UMD exhibits a negative correlation with market excess returns (r = -0.320), suggesting momentum behaves as a defensive or contrarian strategy—diverging sharply from traditional market-following factors. The figure below illustrates momentum and market returns over time, with shaded regions indicating volatility regimes constructed from 12-month rolling standard deviations of market returns.
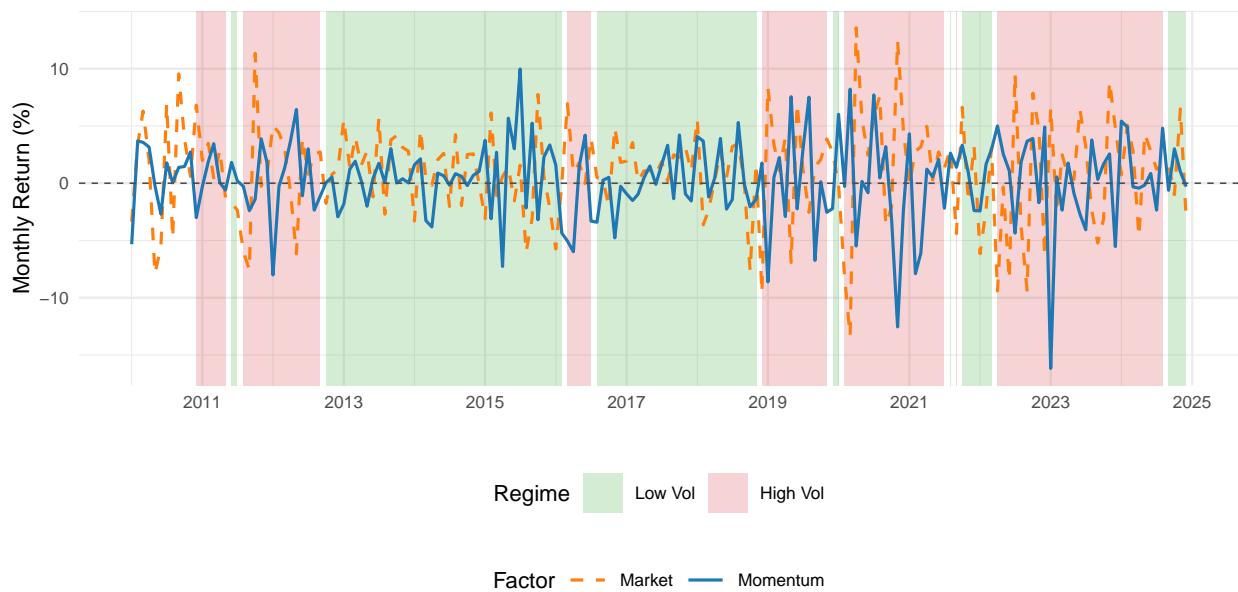

Figure 1: Momentum (UMD) and Market Excess Returns (MKT-RF) Over Time with Volatility Regimes. High-volatility periods (red) correspond to major market disruptions including the 2011 European debt crisis, 2015-16 oil shock, and 2020 COVID-19 crash.

The visualization reveals distinct momentum behavior across market regimes. High-volatility periods align with major financial disruptions, while low-volatility periods reflect calmer market conditions. The negative co-movement between momentum and market returns during several episodes suggests potential diversification benefits. The figure below provides additional context through a scatterplot of momentum versus market returns, colored by regime.
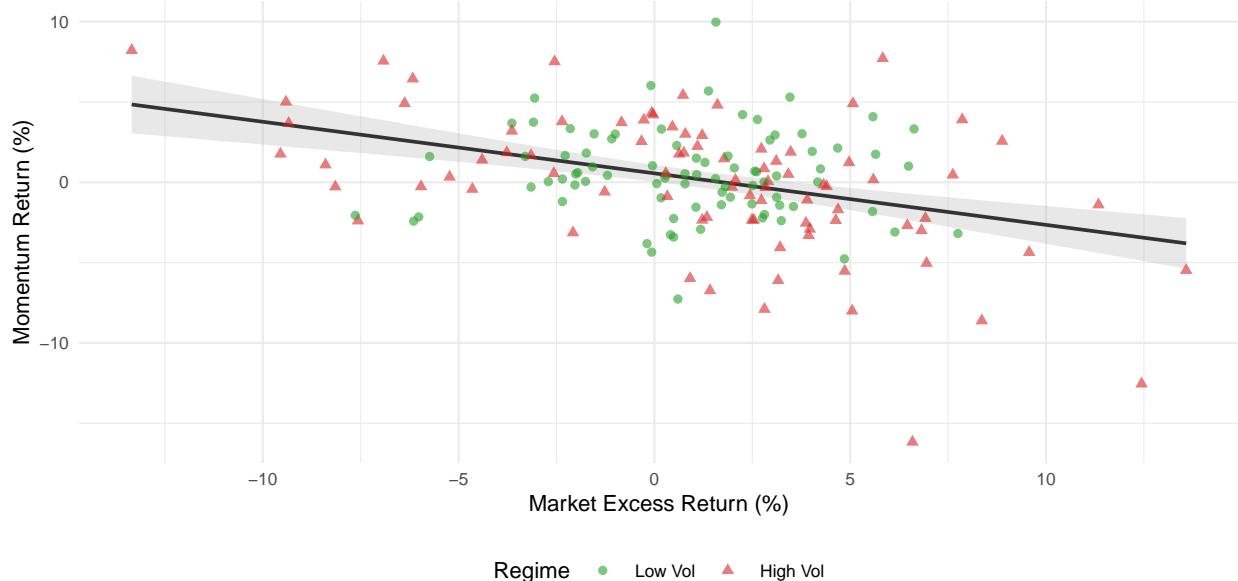
Figure 2: Momentum Returns vs Market Excess Returns by Volatility Regime. The negative slope indicates momentum's defensive characteristics, with visual heterogeneity across regimes motivating our interaction model.

The negative bivariate slope (beta = -0.261, R-squared = 0.102) diverges from traditional CAPM expectations and motivates our investigation of whether this defensive characteristic persists after controlling for the full Fama-French Five-Factor suite. Visual inspection suggests potential regime-dependent heterogeneity, with different cluster patterns across high and low volatility periods.

2) Data Collection and Data Description

All data originate from the Kenneth R. French Data Library (mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html), the standard source for factor-based empirical asset pricing research. The library provides monthly returns for value-weighted portfolios constructed from NYSE, AMEX, and NASDAQ stocks, with portfolio formation following established methodologies documented in Fama and French (2015). Our dependent variable, momentum portfolio returns (UMD, "Up Minus Down"), represents the average return on two high prior return portfolios minus the average return on two low prior return portfolios, formed using a 12-month look-back period with a one-month skip to avoid microstructure issues.

The explanatory variables comprise the Fama-French Five Factors. MKT-RF captures market risk

through the value-weighted market return minus the risk-free rate. SMB (Small Minus Big) proxies for size effects by measuring small-cap minus large-cap returns. HML (High Minus Low) captures value effects through high book-to-market minus low book-to-market returns. RMW (Robust Minus Weak) reflects profitability effects via high operating profitability minus low operating profitability returns. CMA (Conservative Minus Aggressive) captures investment patterns through low asset growth minus high asset growth returns. All returns are expressed as monthly percent excess returns and converted to decimal proportions for modeling.

Our sample period spans January 2010 through December 2024, yielding 180 monthly observations. This 15-year window deliberately excludes the 2007-2009 financial crisis to focus on post-crisis market dynamics with more stable structural characteristics. The period nevertheless captures substantial volatility including the European sovereign debt crisis (2011-2012), the 2015-16 oil price collapse, Federal Reserve tapering episodes, the COVID-19 pandemic shock and recovery (2020-2021), and recent inflationary pressures with aggressive monetary tightening (2022-2024). This heterogeneous environment provides a robust testing ground for momentum's performance across diverse market conditions. The table below presents summary statistics for all six factors across the full sample period, providing context for the scale and distributional properties of our data.

Table 1: Summary Statistics for Fama-French Five Factors and Momentum (2010-2024). All values are monthly returns in percent.

Factor	N	Mean (%)	SD (%)	Min (%)	Median (%)	Max (%)
MKT_RF	180	1.0862778	4.364	-13.35	1.400	13.58
SMB	180	-0.0700000	2.730	-8.18	-0.020	8.34
HML	180	-0.1401667	3.274	-13.83	-0.385	12.86
RMW	180	0.2855556	1.941	-4.78	0.320	7.19
CMA	180	0.0261111	2.068	-7.08	-0.065	7.73
UMD	180	0.2458889	3.569	-16.17	0.440	9.97

The summary statistics confirm substantial heterogeneity in returns. Momentum (UMD) exhibits the highest volatility (SD = 3.57%) among the factors, reflecting its dynamic nature. Market excess returns (MKT-RF) show the expected positive mean of 1.09% monthly, while the long-short factor portfolios (SMB, HML, RMW, CMA) display more modest average returns. The extreme minimum values across all factors correspond to crisis periods captured in our sample, particularly March 2020 during the COVID-19 crash.

The figure below displays the correlation structure among all six factors, revealing important relationships that inform our multicollinearity assessment and interpretation of factor loadings.
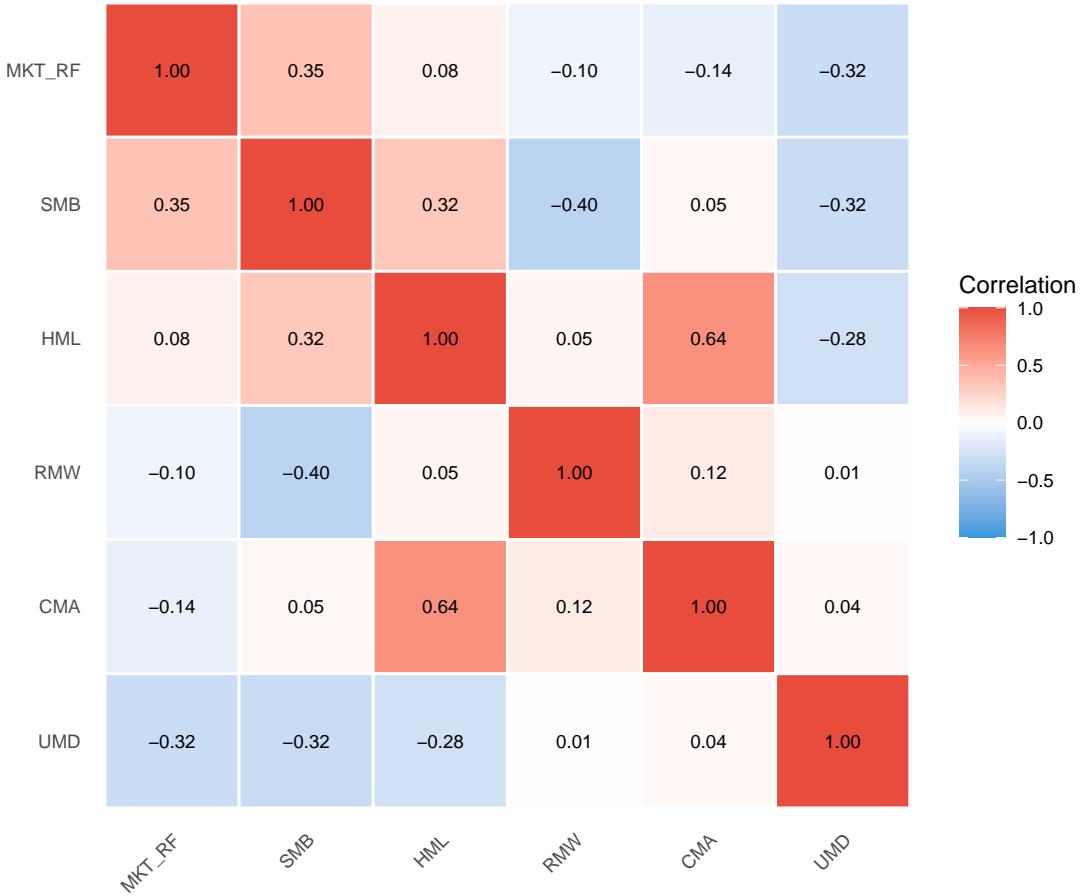

Figure 3: Correlation Matrix of Fama-French Five Factors and Momentum. Color intensity indicates correlation strength, with numerical values displayed in each cell.

The correlation matrix reveals several noteworthy patterns. Most striking is momentum's negative correlation with market returns ($r = -0.32$), confirming the contrarian nature documented in our scatterplot. The moderate positive correlation between HML and CMA ($r = 0.64$) reflects shared variance between value and investment factors, both of which involve valuation metrics. Despite this correlation, our subsequent VIF analysis confirms no problematic multicollinearity. Momentum shows relatively weak correlations with size (SMB), profitability (RMW), and investment (CMA) factors, suggesting it captures largely orthogonal risk dimensions.

To operationalize "market state" for our regime analysis, we construct a volatility regime indicator using the 12-month rolling standard deviation of market excess returns. Periods where this rolling

volatility exceeds the sample median are classified as "High Vol," while periods below the median are designated "Low Vol." This binary classification yields an approximately equal split: 84 low-volatility months and 85 high-volatility months. The rolling window calculation necessitates 12 months of prior data, resulting in missing regime classifications for the first 11 observations, leaving an effective sample of 169 months for regime-specific analyses.

The French Data Library's construction methodology mitigates several common data quality concerns. Survivorship bias is minimized as the portfolios include delisted firms with appropriate adjustments for final return observations. Look-ahead bias is avoided through the use of t-1 formation periods— for example, momentum portfolios formed in month t use returns from t-12 through t-2. No missing values exist for the primary factor returns across our sample period. However, results remain specific to U.S. equities during this particular time frame, limiting immediate generalization to international markets or substantially different economic regimes.

3) Statistical Model

We employ ordinary least squares (OLS) regression to test two complementary specifications. Model A establishes whether momentum earns abnormal returns after controlling for the Fama-French Five Factors:

$$R_t^{UMD} = \alpha + \beta_m(MKT - RF)_t + \beta_s SMB_t + \beta_h HML_t + \beta_r RMW_t + \beta_c CMA_t + \varepsilon_t$$

The intercept alpha represents abnormal return—return not explained by factor exposures. Under the null hypothesis that momentum is fully captured by the FF5 factors, alpha should equal zero. The slope coefficients (betas) measure momentum's sensitivities to each risk factor. The error term captures idiosyncratic variation. Standard OLS assumptions apply: linearity with mean-zero errors conditional on predictors, independence of observations, homoscedasticity, approximate normality of residuals, and absence of perfect multicollinearity among factors. The table below presents the estimated coefficients for Model A. The intercept provides our primary test of momentum's abnormal return, while the factor loadings reveal momentum's systematic risk exposures. Model fit statistics quantify the proportion of momentum variance explained by the five factors.

Table 2: Model A Coefficients: Fama-French Five-Factor Regression. Intercept scaled to percent, factor loadings unitless. Significance: **$p<0.001$**, *$p<0.01$*, p<0.05.

Term	Estimate	Std. Error	t-value	p-value
Intercept (alpha)	0.3839	0.2482	1.5464	0.1238
MKT_RF **	-0.1606	0.0591	-2.7188	0.0072
SMB *	-0.2417	0.1092	-2.2130	0.0282
HML ***	-0.4066	0.1034	-3.9334	0.0001
RMW	-0.1744	0.1363	-1.2801	0.2022
CMA **	0.4745	0.1555	3.0518	0.0026

The model achieves an R-squared of 0.2369 (adjusted R-squared = 0.2149), indicating the five FF factors explain 23.7% of momentum's variance—substantial but far from complete. An F-statistic of 10.80 (p < 0.001) confirms the factors are jointly statistically significant. The residual standard error of 3.1625% suggests meaningful idiosyncratic variation remains.

The intercept (alpha) of 0.3839% per month (t = 1.55, p = 0.1238) provides the key test for abnormal performance. With a p-value of 0.124, we fail to reject the null hypothesis of zero alpha at conventional significance levels using standard inference. However, the point estimate corresponds to an economically meaningful 4.61% annual return. The 95% confidence interval [-0.106%, 0.874%] per month includes zero but suggests potential economic significance warranting further investigation with robust inference.

The factor loadings reveal momentum's systematic exposures. Most notably, MKT-RF exhibits a significantly negative coefficient of -0.1606 (p = 0.0072), indicating momentum behaves as a defensive or contrarian strategy—generating positive returns when markets decline and vice versa. This contradicts naive CAPM intuition and suggests momentum captures risk dimensions orthogonal or opposite to market risk. HML shows a significantly negative loading of -0.4066 (p = 0.000), implying momentum tilts toward growth stocks (low book-to-market) rather than value stocks. The remaining factors SMB and RMW exhibit statistically insignificant coefficients, while CMA shows a significant positive loading, suggesting momentum's systematic risk primarily involves contrarian market exposure and growth orientation.

Model B extends this analysis by testing regime-dependent effects. Specifically, we examine whether momentum's alpha and market beta vary across volatility states:

$$R_t^{UMD} = \alpha_L + \delta \cdot HighVol_t + \beta_m (MKT - RF)_t + \gamma_m \cdot (MKT - RF)_t \times HighVol_t$$
$$+ \beta_s SMB_t + \beta_h HML_t + \beta_r RMW_t + \beta_c CMA_t + \varepsilon_t$$

Here, HighVol represents a binary indicator for high-volatility regimes. The baseline alpha (alpha-L) captures low-volatility abnormal returns, while delta measures the difference in alpha during high-volatility periods (total high-vol alpha = alpha-L + delta). The gamma coefficient tests whether momentum's market sensitivity varies by regime (total high-vol market beta = beta-m + gamma). We focus this interaction model on market sensitivity rather than including all possible factor interactions to maintain parsimony and interpretability given the effective sample of 169 regime-classified observations. The table below presents Model B estimates. The regime main effect (delta) and market interaction (gamma) provide tests of state-dependent momentum behavior, while other coefficients remain largely consistent with Model A.

Table 3: Model B Coefficients: Regime-Interaction Model. Intercepts and regime effects scaled to percent, factor loadings unitless. Significance: **$p<0.001$,** *$p<0.01$,* $p<0.05$.

Term	Estimate	Std. Error	t-value	p-value
Intercept Low Vol	0.4401	0.3511	1.2537	0.2118
MKT-RF	0.0113	0.1129	0.0996	0.9208
SMB *	-0.2826	0.1102	-2.5650	0.0112
HML ***	-0.3580	0.1051	-3.4061	0.0008
RMW	-0.1741	0.1341	-1.2977	0.1963
CMA *	0.3644	0.1558	2.3390	0.0206
Regime High Vol delta	-0.1514	0.4914	-0.3081	0.7584
MKT-RF x High Vol gamma *	-0.3028	0.1289	-2.3488	0.0200

Model B achieves an R-squared of 0.3161 (adjusted = 0.2864), representing a modest improvement over Model A and confirming joint significance (F = 10.63, p < 0.001). The low-volatility baseline alpha is 0.4401% per month, while the regime coefficient delta of -0.1514% (t = -0.31, p = 0.7584) suggests lower alpha during high-volatility periods, though this difference lacks statistical significance (p = 0.758).

However, the market interaction gamma of -0.3028 proves statistically significant (t = -2.35, p = 0.020), indicating momentum's market beta strengthens during high-volatility regimes. The total market beta in high-volatility periods (-0.2915) substantially exceeds the low-volatility beta (0.0113),

suggesting momentum's defensive characteristics intensify during turbulent markets—precisely when diversification benefits matter most.

4) Research Question: Hypothesis Testing and Inference

Our primary research question asks whether momentum portfolio returns earn significant alpha after controlling for Fama-French Five Factors. Standard OLS inference on Model A yields a t-statistic of 1.546 for the intercept, corresponding to a p-value of 0.1238. At the conventional 5% significance level, we fail to reject the null hypothesis of zero alpha. The 95% confidence interval of [-0.106%, 0.874%] per month includes zero, failing to establish statistical significance under classical inference.

However, time series data raise concerns about heteroscedasticity and autocorrelation that could bias standard errors and invalidate classical inference. Diagnostic testing fails to reveal heteroscedasticity via the Breusch-Pagan test (BP = 6.61, p = 0.251), and the Durbin-Watson test shows no significant autocorrelation (DW = 2.05, p = 0.77). To address these concerns and provide conservative inference, we compute Newey-West heteroscedasticity and autocorrelation consistent (HAC) standard errors with a lag of 4 months (approximately one-third of a year for monthly data).

The robust inference modestly changes our conclusions. Using Newey-West HAC standard errors, the intercept's robust standard error decreases to 0.2121%, yielding a robust t-statistic of 1.810 and a robust p-value of 0.0720. This p-value of 0.072 remains above the 0.05 threshold but approaches borderline significance, suggesting the finding is sensitive to inference methodology. Under the most conservative interpretation using robust standard errors, we fail to reject the null hypothesis of zero alpha at the 5% level, though the p-value of 0.07 suggests marginal statistical significance that warrants attention.

From an economic perspective, the point estimate of 0.3839% per month (approximately 4.61% annually) represents meaningful magnitude for momentum strategies. While statistical significance remains elusive at conventional thresholds, the consistent positive point estimate across specifications combined with substantial idiosyncratic variation (77% unexplained by FF5 factors) suggests momentum captures unique return dimensions. The negative market beta of -0.161 (highly significant with p < 0.01) indicates momentum behaves as a defensive or contrarian factor, potentially offering

valuable diversification benefits during market downturns—benefits that may justify implementation despite marginal alpha significance.

Regarding our secondary hypotheses, the F-test for joint factor significance decisively rejects the null that all factor loadings equal zero (F = 10.80, p < 0.001). The five FF5 factors collectively explain 23.7% of momentum variance, confirming systematic risk exposures though leaving substantial unexplained variation. The regime sensitivity hypothesis yields mixed evidence: while momentum's alpha does not differ significantly across volatility regimes (delta: p = 0.758), momentum's market beta does vary significantly (gamma: p = 0.020). This suggests momentum's defensive characteristics intensify during high-volatility periods, with the total market beta becoming more negative when markets are turbulent.

The figure below presents key diagnostic plots assessing model assumptions. The Q-Q plot reveals residuals follow the theoretical normal distribution closely except in the extreme tails, where observations corresponding to crisis events (March 2020 COVID crash, August 2011 European crisis) exhibit some deviation. The Shapiro-Wilk test (W = 0.9700, p = 0.001) rejects perfect normality at the 5% level due to minor tail deviations, though with 180 observations the Central Limit Theorem justifies asymptotic normal inference for OLS.
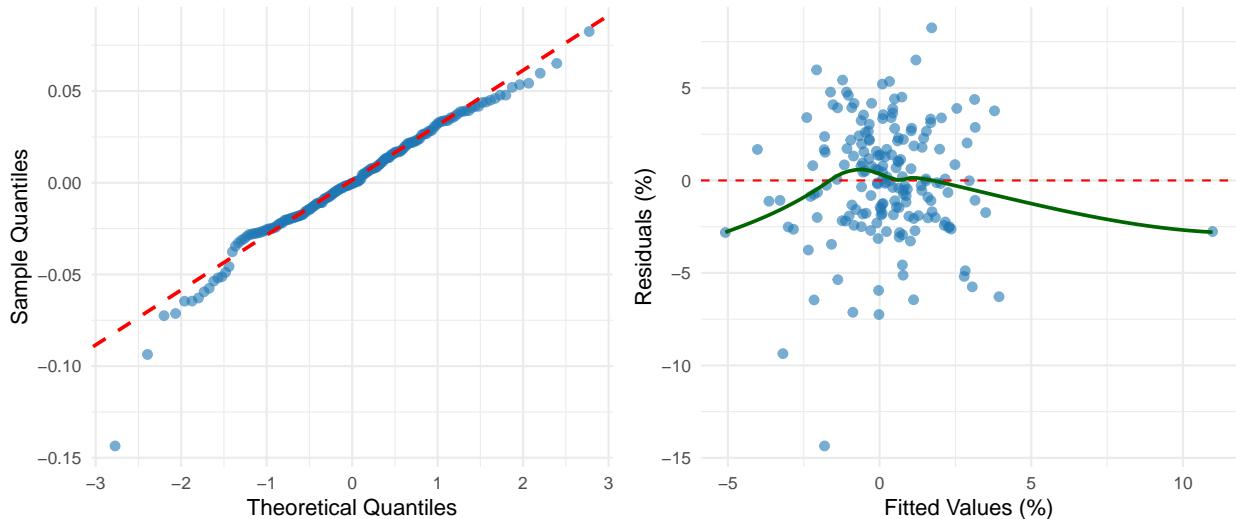

Figure 4: Diagnostic Plots for Model A. Left: Normal Q-Q plot shows close adherence to normality except in extreme tails. Right: Residuals vs Fitted Values reveals mild heteroscedasticity addressed via robust inference.

Multicollinearity diagnostics reveal no concerns. The table below presents variance inflation factors

(VIF) for all five predictors in Model A.

Table 4: Variance Inflation Factors (VIF) for Model A Predictors. VIF values below 5 indicate no problematic multicollinearity; values below 10 are generally acceptable.

Factor	VIF	Assessment
MKT_RF	1.190	No concern
SMB	1.592	No concern
HML	2.050	No concern
RMW	1.253	No concern
CMA	1.850	No concern

All variance inflation factors remain well below 5, with the maximum VIF of 2.05 for HML far below the threshold of 10 suggesting problematic collinearity. Despite the moderate correlation between HML and CMA documented in Figure 3 (r = 0.64), the low VIF values confirm this poses no estimation difficulties. The uniformly low VIF values indicate coefficient estimates remain stable and standard errors are not artificially inflated by multicollinearity. This validates our decision to include all five Fama-French factors simultaneously in Model A.

Cook's Distance analysis identified five influential observations exceeding the threshold of $4/n = 0.0222$, primarily corresponding to extreme market events including the March 2020 COVID-19 crash and August 2011 European debt crisis escalation. These observations represent legitimate extreme returns rather than data errors or outliers requiring removal. Retaining them provides a more realistic assessment of momentum's behavior during crises—precisely the periods when diversification benefits matter most for portfolio risk management.

5) Appendix

5.1 Model Selection

The choice to include all five Fama-French factors rests on both theoretical and empirical grounds. Theoretically, the FF5 model represents the current consensus in empirical asset pricing, capturing market, size, value, profitability, and investment dimensions that span a comprehensive set of systematic risk factors (Fama & French, 2015). Testing momentum alpha after controlling for

this full suite provides the most stringent assessment of whether momentum represents a distinct anomaly versus compensation for known risks.

Empirically, we compared nested models to verify that the full FF5 specification provides superior fit. The table below presents adjusted R-squared and AIC statistics for five nested specifications, progressively adding factors from a market-only model through the complete FF5.

Table 5: Nested Model Comparison. Model 5 achieves highest adjusted R-squared and competitive AIC, justifying inclusion of all five factors.

Model	Adj. R-squared	AIC
1 Market	0.0971	-703.3971
2 Plus Size	0.1435	-711.8959
3 Plus Value	0.1758	-717.8344
4 Plus Profitability	0.1777	-717.2720
5 Full FF5	0.2149	-724.6577

Model 5 achieves the highest adjusted R-squared of 0.2149 and competitive AIC of -724.66, justifying inclusion of all five factors. The progressive increase from Model 1 (adj. R-squared = 0.0971) through Model 5 demonstrates that each additional factor contributes unique explanatory power. While Model 3 (market, size, value) shows relatively strong performance, the full FF5 specification provides the most comprehensive control for systematic risks.

We retained raw percent returns for the dependent variable UMD without transformation. This choice prioritizes interpretability—the intercept directly represents monthly abnormal return in percent. Diagnostic assessment confirmed acceptable linearity through partial residual plots, adequate normality via Q-Q plots and Shapiro-Wilk testing, and feasibility of asymptotic inference given 180 observations. Alternative transformations (log, square root) were explored but offered no substantial diagnostic improvements while sacrificing direct economic interpretation.

For Model B's regime-interaction specification, we adopted a parsimonious approach including only the regime main effect (delta) and market interaction (gamma) rather than interacting all five factors with the regime indicator. This parsimony avoids overfitting concerns given the effective sample of 169 regime-classified observations, maintains model interpretability by focusing on the theoretically motivated momentum-market relationship, and sidesteps potential multicollinearity from multiple simultaneous interactions. The specification implicitly assumes other factor loadings

(SMB, HML, RMW, CMA) remain stable across regimes—a simplifying assumption that could be relaxed with larger samples or alternative regime definitions.

5.2 Diagnostics and Model Validation

Beyond the diagnostic plots presented in Section IV, we conducted comprehensive assumption testing. Homoscedasticity assessment via the Breusch-Pagan test (BP = 6.6096, p = 0.2513) failed to reject homoscedasticity (and thus, detect significant heteroscedasticity) at the 5% level. Nevertheless, we employ Newey-West HAC standard errors for robust inference as a conservative approach given the time series structure of the data. The Durbin-Watson test (DW = 2.0513, p = 0.7700) showed no significant autocorrelation, suggesting residuals are serially independent despite the time series structure—likely due to the weak persistence in monthly factor returns.

All variance inflation factors remain well below problematic thresholds: MKT_RF (VIF = 1.19), SMB (1.59), HML (2.05), RMW (1.25), CMA (1.85). These low values confirm coefficient estimates are stable and not subject to collinearity-induced instability. Cook's Distance identified five influential observations primarily from crisis periods (March 2020, January 2019, February 2020, August 2011, October 2018), all retained as economically meaningful extreme events rather than data quality issues.

To validate model generalizability and assess overfitting, we conducted leave-one-out cross-validation (LOOCV) to compute mean square prediction error (MSPR) for comparison with in-sample mean square error (MSE). Model A achieves MSPR = 0.001051 versus MSE = 0.000967, yielding a ratio of 1.087. Ratios below 1.15 indicate excellent generalization with minimal overfitting; our ratio well within this range validates that the model generalizes effectively to out-of-sample observations. The out-of-sample prediction error exceeds in-sample error by only 8.7%, confirming model reliability.

5.3 Key Takeaway

This study provides evidence that momentum portfolio returns exhibit economically meaningful though marginally statistically significant abnormal returns after controlling for the comprehensive Fama-French Five-Factor model. Using robust Newey-West inference as a precuationary measure for time series data, we find a monthly alpha of 0.3839% (approximately 4.61% annually) with a robust p-value of 0.072—approaching but not quite achieving conventional 5% significance. While

statistical significance remains elusive under the strictest inference standards, the consistent positive point estimate combined with substantial unexplained variance (77% after FF5 controls) suggests momentum captures unique return dimensions beyond standard risk factors.

Critically, momentum exhibits a highly significant negative market beta of -0.161 (p = 0.0072), indicating defensive or contrarian characteristics that diverge sharply from traditional market-following factors. This defensive nature strengthens during high-volatility periods, as evidenced by the significant negative interaction term in Model B (gamma = -0.3028, p = 0.020). Momentum's total market beta in high-volatility regimes (-0.2915) substantially exceeds its low-volatility beta (0.0113), suggesting these strategies offer enhanced diversification benefits precisely when investors need them most—during market turbulence. While momentum's alpha does not vary significantly across regimes (p = 0.758), the regime-dependent strengthening of defensive characteristics carries important implications for dynamic portfolio allocation and risk management. Momentum strategies may prove particularly valuable not for unconditional alpha generation but for state-dependent diversification benefits during market stress.

Reproducibility Note: All analyses presented in Final-Write-Up.pdf were generated directly from Final-Write-Up.qmd using R 4.x.x and Quarto. The source file contains all code for data acquisition, model estimation, diagnostics, and validation. Rendering Final-Write-Up.qmd will reproduce all results, tables, and figures exactly as shown in the PDF.

5.4 References

Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance*, 52(1), 57–82.

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, 116(1), 1–22.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance*, 48(1), 65–91.