Regime-Aware Tactical Sector Allocations

Leo Granskog

Dissertation written under the supervision of Professor Dan Tran

Dissertation submitted in partial fulfilment of requirements for the MSc in Finance, at the Universidade Católica Portuguesa, 5/1/2025.

Abstract

This thesis examines whether identifying macroeconomic regimes can improve tactical sector-level allocations relative to static benchmarks and non-regime-based strategies. Using unsupervised learning techniques applied to macroeconomic fundamentals, conditions are classified into distinct regimes that capture differences in growth, inflation, and volatility dynamics. A two-stage framework separates crisis periods from normal business-cycle conditions and uses the resulting regime information to condition systematic tactical sector allocation decisions. All strategies are evaluated in a strictly out-of-sample setting to mitigate overfitting concerns.

The empirical results show that regime-aware tactical sector allocation improves risk-adjusted performance relative to static benchmarks. Regime-aware strategies achieve higher Sharpe ratios and lower maximum drawdowns, with performance differences that are statistically significant according to Jobson–Korkie tests at the five percent level. The gains are primarily driven by improved downside protection during volatile periods, while maintaining competitive returns during expansionary periods. Importantly, these results remain robust after accounting for realistic transaction costs. More flexible machine learning-based allocation rules do not outperform constrained regime-based strategies out-of-sample.

The economic relevance of the identified regimes is supported by their ability to differentiate sector return behavior and enhance portfolio outcomes. Overall, the findings indicate that macroeconomic regime information extracted through unsupervised learning provides economically meaningful signals for tactical sector allocation and offers a disciplined, implementable framework for active portfolio management.

Title: Regime-Aware Tactical Sector Allocations

Author: Leo Granskog

Keywords: Macroeconomic Regimes; Unsupervised Learning; Tactical Sector Allocation; Portfolio Choice

Resumo

Esta dissertação analisa se a identificação de regimes macroeconómicos pode melhorar a alocação tática ao nível setorial em comparação com benchmarks estáticos e estratégias não condicionadas por regimes. Utilizando técnicas de aprendizagem não supervisionada aplicadas a fundamentos macroeconómicos, as condições económicas são classificadas em regimes distintos que capturam diferenças nas dinâmicas de crescimento, inflação e volatilidade. Um enquadramento em duas etapas distingue períodos de crise das condições normais do ciclo económico e utiliza a informação de regime resultante para condicionar decisões sistemáticas de alocação tática setorial. Todas as estratégias são avaliadas num ambiente estritamente fora da amostra, com o objetivo de mitigar riscos de sobreajustamento.

Os resultados empíricos mostram que a alocação tática setorial sensível a regimes melhora o desempenho ajustado ao risco relativamente a benchmarks estáticos. As estratégias condicionadas por regimes apresentam rácios de Sharpe mais elevados e menores drawdowns máximos, com diferenças estatisticamente significativas segundo os testes de Jobson–Korkie ao nível de cinco por cento. Os ganhos são sobretudo impulsionados por uma melhor proteção em períodos de elevada volatilidade, mantendo retornos competitivos em fases expansionistas. Estratégias de alocação baseadas em aprendizagem automática mais flexíveis não superam estratégias condicionadas por regimes e sujeitas a restrições fora da amostra.

Em conjunto, os resultados indicam que a informação sobre regimes macroeconómicos fornece sinais economicamente relevantes para a alocação tática setorial e constitui um enquadramento disciplinado e implementável para a gestão ativa de carteiras.

Título: Alocação Tática Setorial Sensível a Regimes

Autor: Leo Granskog

Palavras-chave: Regimes macroeconómicos; aprendizagem não supervisionada; alocação tática setorial; escolha de carteiras.

Contents

1. Introduction

A central challenge in portfolio management is not whether economic regimes exist, but whether regime information can be translated into allocation decisions without introducing instability or overfitting. Financial markets exhibit persistent regime dependence: periods of expansion, slowdown, and crisis differ systematically in growth, inflation, volatility, and cross-asset correlations. During adverse economic conditions, correlations tend to rise and volatility clusters, undermining the effectiveness of static diversification strategies (Ang & Bekaert, 2004). While decades of research document regime-dependent behavior in asset returns, volatilities, and correlations, translating these insights into reliable portfolio strategies remains difficult. Many regime-aware approaches rely on unconstrained optimization or highly flexible models that perform well in-sample but deteriorate out-of-sample once estimation errors, transaction costs, and turnover are considered (Bailey et al., 2014; Harvey et al., 2016).

Regime-switching models formalize these patterns and capture several stylized facts, including volatility clustering and time-varying risk premia (Hamilton, 1989; Ang & Timmermann, 2012). Practitioner-oriented studies similarly argue that conditioning portfolio exposure on macroeconomic environments can improve risk-adjusted performance (Kritzman et al., 2012; De Longis & Ellis, 2023). However, many regime-based allocation strategies struggle to deliver consistent out-of-sample performance once regime misclassification and trading frictions are considered. As a result, the practical challenge is not regime identification, but the design of allocation rules that remain stable and implementable under realistic constraints.

Recent advances in machine learning offer new tools for regime identification. Unsupervised learning methods can extract latent patterns from high-dimensional data without imposing strong parametric assumptions (Gu, Kelly, & Xiu, 2020). These methods are attractive in settings where macroeconomic relationships are complex and evolving. However, increased model flexibility also increases the risk of overfitting. In portfolio applications, small improvements in predictive accuracy often fail to translate into superior performance once turnover and transaction costs are taken into account (Snow, 2020; Harvey et al., 2016).

An emerging literature responds to these challenges by embedding economic structure and constraints into data-driven frameworks. Rather than treating machine learning as a substitute for financial theory, they emphasize disciplined portfolio construction and robustness (Giglio, Kelly, & Xiu, 2022). Regime-based allocation provides a natural setting for this perspective. Regimes offer an interpretable structure through which macroeconomic information can guide allocation decisions, while avoiding reliance on precise short-horizon return forecasts.

Most existing regime-based allocation studies focus on broad asset classes or factor portfolios (Guidolin & Timmermann, 2007; Shu & Mulvey, 2024). While informative, this leaves an important question unanswered. Sector-level allocation occupies a middle ground between aggregate market exposure and individual security selection. Sectors differ meaningfully in their sensitivity to macroeconomic conditions and are closely linked to economic intuition. They also allow straightforward implementation through liquid instruments such as sector ETFs. Whether regime information remains valuable once realistic sector-level constraints are imposed therefore remains underexplored.

This thesis addresses that gap by combining macroeconomic regime identification with tactical sector allocation. It asks two related research questions. First, can unsupervised learning methods applied to macroeconomic fundamentals identify regimes associated with economically distinct sector return behavior? Second, do regime-aware tactical sector strategies outperform passive benchmarks and traditional active approaches once realistic constraints and transaction costs are imposed? The analysis treats regime classifications as a guide to portfolio allocation decisions, not as an objective in itself.

Importantly, the proposed framework does not replace an underlying investment strategy. Instead, it operates as a tactical overlay that applies small, constrained tilts around a base allocation. In this thesis, the base strategy is the market-capitalization-weighted S&P 500, where tactical tilts are applied to the sector weights depending on regimes. This perspective aligns with practical portfolio management, where tactical allocation is typically implemented as fine-tuning rather than wholesale portfolio reallocation.

The proposed framework differs from standard regime-switching approaches in several respects. Regimes are identified using macroeconomic fundamentals rather than asset returns, reducing sensitivity to market noise (McCracken & Ng, 2016). A two-stage structure separates crisis and stress periods from normal business-cycle conditions, following the intuition of turbulence-based regime detection (Kritzman et al., 2012). Dimensionality reduction is implemented using rolling principal component analysis with sign correction, allowing the framework to adapt to evolving macroeconomic relationships while maintaining interpretability (Hirsa et al., 2024). Regime information is then mapped into constrained, long-only sector tilts designed to reflect realistic portfolio management practice.

Empirically, the analysis identifies four macroeconomic regimes, Expansion, Slowdown, Stress, and Crisis, with distinct macroeconomic profiles and sector return characteristics.

Regime-aware tactical sector strategies outperform the S&P 500 benchmark in a strictly out-of-sample evaluation. The performance gains arise primarily from improved downside protection during recessionary and high-volatility environments, while returns during expansionary periods remain competitive. We also find that economically structured and constrained regime-aware strategies outperform more flexible machine learning allocation models out-of-sample.

This thesis makes three contributions. First, it shows that macroeconomic regimes extracted from fundamental data exhibit economically meaningful and statistically distinct sector return patterns. Second, it demonstrates that regime-awareness improves tactical sector allocation primarily through risk management rather than return forecasting. Third, it provides evidence that structured, constrained regime-based allocation rules outperform more flexible machine learning approaches in out-of-sample portfolio performance.

The remainder of the thesis proceeds as follows. Section 2 reviews the relevant literature. Section 3 describes the data. Section 4 outlines the methodology for regime identification and portfolio construction. Section 5 presents the empirical results and robustness checks. Section 6 discusses the economic interpretation, practical implications, and limitations. Section 7 concludes.

2. Literature Review

The application of machine learning to portfolio management represents a natural evolution of quantitative finance, building upon decades of research into asset pricing, portfolio optimization, and business cycle analysis. This review combines three research paths, regime identification, tactical allocations, and machine learning, to position this thesis within the current literature.

2.1 Machine Learning in Asset Pricing and Portfolio Construction

Asset pricing has long operated under parametric constraints that limit what patterns can emerge from the data. Machine learning offers an alternative, one that can find patterns without those constraints. Gu, Kelly, and Xiu (2020) find that gradient boosted trees and neural networks outperform linear models for return prediction, even when restricted to standard Fama-French characteristics. This improvement does not come from discovering new anomalies but from models finding underlying relationships more flexibly.

However, this flexibility creates new issues. With hundreds of potential factors and model specifications available, the risk of data mining rises sharply. Harvey, Liu, and Zhu (2016) argue that conventional significance thresholds are far too permissive given how many specifications researchers can implicitly test, and they propose adjustments that would make many published anomalies insignificant. Bailey, Borwein, López de Prado, and Zhu (2014) had formalized this concern as "backtest overfitting," demonstrating that strategies with impressive in-sample performance frequently fail out-of-sample because selection bias inflates apparent skill. These critiques motivate the strict temporal separation used in this thesis where all model selection occurs on only training and validation data before any test period evaluation, in order to guarantee that statistical significance tests occur on genuinely unseen data.

Rather than treating models as black boxes, some researchers are imposing economic structure to combat overfitting. Giglio, Kelly, and Xiu (2022) find that constraints like no-arbitrage conditions and monotonicity improve out-of-sample results by narrowing what the model can fit. Snow (2020) raises a related point, arguing that being more accurate doesn't always mean performing better. Models with stable but slightly less precise forecasts can outperform those that are often right but sometimes spectacularly wrong. Similarly, Barroso, Saxena, and Wang (2023) show that portfolios based on characteristic-driven expected returns, including those obtained from advanced machine learning estimators, require explicit regularization of the portfolio optimization problem to avoid unstable weights and weak out-of-sample performance.

In order to avoid the gap between accuracy and performance researchers are taking on new approaches. Chevalier, Coqueret, and Raffinot (2022) train models to predict portfolio weights directly instead of returns and then weights. They report higher Sharpe ratios than traditional two step methods. However, this comes at a cost, understanding why the model chose a given allocation becomes increasingly difficult. The preceding discussion raises a natural question, if machine learning models can capture complex, nonlinear relationships, what structure should guide their application to portfolio management? One answer pursued in this thesis is that financial markets operate in distinct regimes and if regimes can be identified then making investment decisions based on regime states offers a principled way to harness machine learning's flexibility while preserving economic interpretability.

2.2 Regime Identification in Financial Markets

The empirical case for regime switching in financial markets rests on the finding that asset behavior differs across states. Ang and Bekaert (2004) find that international equity correlations increase during bear markets, precisely when diversification benefits are most needed. They also show that regime-aware strategies generate higher Sharpe ratios even with imperfect regime classification. Subsequent work by Ang and Timmermann (2012) shows that regime-switching models can explain several stylized facts such as volatility clustering, fat-tailed distributions, and time-varying risk premia, which single-regime models struggle to do. However, the question of how many regimes to specify remains contested. Ang and Bekaert assume two states, Guidolin and Timmermann (2007) find that at least four states are needed to capture the different return distributions in international equity markets. The sensitivity to regime count motivates the data driven clustering approach used in this thesis

Hidden Markov Models remain appealing because they model regime persistence through state transition probabilities. Unlike static clustering models, Hidden Markov Models capture how regimes evolve over time, diagonal-dominant transition matrices reflect that regimes often persist rather than switch randomly. Kritzman, Kulasekaran, and Turkington (2023) show that when regime classification is ambiguous, as it often is near business cycle turning points, probabilistic regime weights may outperform discrete classifications. This insight motivates testing whether transition dynamics improve upon simpler approaches in tactical allocations.

Kritzman, Page, and Turkington (2012) introduce a turbulence index based on Mahalanobis distance that detects regime shifts in real time by measuring how unusual current market conditions are relative to historical norms. Portfolios conditioned on turbulence states achieve superior risk-adjusted returns in their analysis, and their methodology directly informs the crisis detection framework employed in this thesis. This study extends their approach by combining crisis identification with separate clustering-based business cycle detection, a distinction they do not make.

Unsupervised learning approaches to regime identification have gained traction in recent years. Oliveira et al. (2025) utilize k-means clustering macroeconomic data, identifying four distinct regimes and illustrate that regime-aware portfolios achieve higher Sharpe ratios compared to static portfolios, while Mueller-Glissmann and Ferrario (2024) employed random forests models for regime detection and found reduced maximum drawdowns through regime-aware positioning. Hirsa, Xu, and Malhotra (2024) show that rolling regime detection outperforms static approaches when structural breaks occur, addressing temporal instability, a finding that motivates the rolling dimensionality reduction employed in this thesis.

2.3 Tactical Asset Allocation

The logic behind tactical allocation is straightforward, if expected returns, risk levels, and asset relationships vary predictably, adjusting portfolios accordingly should generate outperformance against static benchmarks. However, turning observable patterns into actual outperformance involves frictions, timing errors, and behavioral missteps that often erode the theoretical advantage.

Dahlquist and Harvey (2001) find that many variables used for return prediction are linked to stages of business cycles. This predictability means that active strategies can outperform passive strategies. However, complex optimization frameworks, despite their theoretical appeal, frequently underperform simple strategies out-of-sample. The culprit is estimation errors and overfitting. When an optimizer has unconstrained freedom to act on return forecasts, even modest forecast mistakes translate into extreme portfolio tilts. Simple allocation rules may sacrifice in-sample optimality but provide implicit regularization that can be valuable when forecasts are imprecise or uncertain. The strategy design in this thesis takes this lesson into account by studying constrained tactical tilts rather than unconstrained optimization.

De Longis and Ellis (2023) extend this by conditioning allocations on business cycle stages and generate Sharpe ratios higher than static portfolios. Interestingly, performance of the portfolios degrades only modestly even when regime classification is wrong 30% of the time. The explanation may lie in the directional nature of regime-conditional tilts. Expansionary periods tend to favor equities, with rising earnings and narrowing credit spreads, while recessions tend to favor bonds as equity earnings contract and central banks ease their policies. If these relationships are strong enough then getting the regime right most of the time can generate outperformance even if individual calls are sometimes wrong.

The momentum effect documented by Jegadeesh and Titman (1993) has proven remarkably persistent. However, Daniel and Moskowitz (2016) find that momentum carries substantial crash risk, with large losses occurring in panic states, following multi-year market drawdowns, and during periods of high market volatility. Barroso and Santa-Clara (2015) show that this risk is highly variable over time and predictable through realized variance, allowing risk-managed approaches to mitigate crashes. Together, these findings suggest that momentum and regime conditioning capture different information: momentum captures recent trends, while regime-awareness anticipates shifts in the environment that may reverse those trends.

Moreira and Muir (2017) introduce an alternative approach through volatility-managed portfolios that scale market exposure inversely to realized volatility. Their strategy improves Sharpe ratios by mechanically de-risking during turbulent periods. However, Cederburg, O'Doherty, Wang, and Yan (2020) challenge these findings by demonstrating that the benefits of volatility timing are period dependent.

At the same time, the economic relevance of characteristics such as value and momentum as return predictors is well established across markets and asset classes (Asness, Moskowitz, and Pedersen, 2013), motivating their continued use as inputs to portfolio construction despite these implementation challenges. Shu and Mulvey (2024) offer perhaps the closest precedent to our study by examining how factor premia vary across regimes. They find that dynamically adjusting factor exposures based on regime states improve information ratios relative to static factor allocations. The finding implies that regime-dependence extends to the cross-section of expected returns within equity markets.

2.4 Positioning This Research

The preceding sections reveal literature that has made substantial progress on regime identification and machine learning applications to finance, but also one with persistent gaps that this thesis will address.

First, most regime detection frameworks identify states using asset returns, which are noisy. This thesis instead identifies regimes from macroeconomic fundamentals, following McCracken and Ng (2016) data and Hirsa et al. (2024) methodology, but with an addition. This study will separate crisis detection from business cycle clustering because crises are ambiguous and represent a different phenomenon other than ordinary recessions. They are characterized by correlation breakdowns rather than the orderly rotation of indicators that defines normal business cycle fluctuations.

Second, recent work by Oliveira et al. (2025) applies k-means clustering to macroeconomic data for asset class rotation relies on static PCA. However, macroeconomic covariances evolve, the factors that best summarize conditions in the 2000s may differ from those that matter in the 2020s. Rolling PCA with sign correction addresses this limitation by capturing time-varying factor structures.

Third, although Shu and Mulvey (2024) demonstrate regime-dependent factor premia. Sectors offer advantages that factors do not, such as direct alignment to economic intuitions, and straightforward implementation through liquid ETFs.

Finally, Harvey et al.'s (2016) critique of multiple testing is relevant to tactical allocation research. Our study has strict temporal separation, dimensionality reduction parameters, crisis thresholds, and clustering algorithms are selected based on training and validation data. This creates an informed strategy selection applied on test data only once. Combined with the Jobson-Korkie (1981) test for Sharpe ratio differences, as corrected by Memmel (2003), this provides stronger protection against overfitting than studies relying on rolling backtests or single holdout samples.

The research design of this study is therefore aimed towards four objectives. To capture time-varying macroeconomic structures, differentiate crisis dynamics from business cycles, assess allocation decisions transparently, and validate results with statistical rigor.

3. Data

3.1 Dataset Overview

The study employs a two-stage design. Regime detection uses monthly macroeconomic data from January 1984 to December 2024. Tactical allocation evaluation spans November 2001 to December 2024, limited by GICS sector data availability following the reclassification of real estate in 2001.

The 1984 start date for regime detection avoids the pre-Volcker era and disinflation period when monetary policy regimes differed fundamentally from the modern inflation-targeting environment (Clarida et al., 2000). The evaluation period encompasses the dot-com recession, Global Financial Crisis, and COVID-19 pandemic, providing diverse market conditions across expansion, contraction, and crisis environments.

3.2 Data Sources

3.2.1 Macroeconomic Variables

The FRED-MD database provides 126 monthly U.S. macroeconomic indicators spanning eight categories in this thesis: Output and Income, Labor Market, Housing, Consumption and Orders, Money and Credit, Interest and Foreign Exchange, Prices, and the Stock Market. These categories span real economic activity, inflation dynamics, financial conditions, and asset market behavior, offering a broad representation of macroeconomic states. Collectively, the indicators capture the main dimensions of business-cycle conditions and provide the basis for regime identification in this study.

3.2.2 Sector Returns

Monthly total return indices for the eleven GICS sectors come from LSEG Workspace, covering November 2001 to December 2024. The Global Industry Classification Standard provides a standardized taxonomy widely adopted: Communication Services, Consumer Discretionary, Consumer Staples, Energy, Financials, Health Care, Industrials, Information Technology, Materials, Real Estate, and Utilities.

Log returns are computed monthly as $r_{i,t} = \ln \frac{\mathrm{RI}_{\mathrm{i,t}}}{\mathrm{RI}_{\mathrm{i,t-1}}}$, where RI denotes the total return index which incorporates dividends. Log returns ensure additivity over time and maintain alignment with continuous-time financial models. Excess returns are calculated by subtracting the 1-month U.S. Treasury bill rate obtained from the Federal Reserve (DGS1MO series).

3.2.3 Sector Market Weights

Market capitalization weights for each sector are obtained from LSEG Workspace, representing each sector's proportion of total S&P 500 market capitalization at month-end. The weights enable the implementation of constrained tactical tilts that maintain realistic tracking error bounds relative to the capitalization-weighted S&P 500 return index.

3.2.4 Fama-French-Carhart Factors

The Fama-French-Carhart four factors are gathered from the Kenneth R. French Data Library. The dataset includes the market excess return (MKT), size (SMB), value (HML), and momentum (MOM) factors at a monthly frequency. All factor returns are expressed in excess of the one-month U.S. Treasury bill rate and are aligned to the same sample period as the sector return data. These factors are used exclusively for performance attribution and factor exposure analysis and do not enter the regime identification or portfolio construction process.

3.3 Data Processing and Quality

3.3.1 Macroeconomic Data Processing

All macroeconomic data series are transformed according to FRED-MD transformation codes (McCracken & Ng, 2016) prior to analysis. This ensures stationarity, a prerequisite for principal component extraction and regime detection. Following transformation, each series is standardized to zero mean and unit variance. Critically, standardization parameters (mean and standard deviation) are computed only on training period data (1984-2008) and applied unchanged to validation and test periods. This approach prevents information leakage from future observations into the regime detection framework, ensuring out-of-sample evaluation reflects genuinely implementable strategies.

3.3.2 Stationarity Validation

We validate the transformation methodology through stationarity tests on all 126 transformed variables. The Augmented Dickey-Fuller (ADF) test examines the null hypothesis of a unit root against the alternative of stationarity, while the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test examines the null of stationarity against the alternative of a unit root. Results confirm 91.3% of variables reject the unit root null via ADF testing ($p < 0.05$), and 74.6% fail to reject the stationarity null via KPSS testing ($p > 0.05$). Agreement between both tests (83.3% of variables) validates the McCracken and Ng (2016) transformation methodology for our sample period.

3.3.3 Missing Data Treatment

The FRED-MD database maintains high data quality with minimal missing observations. Following transformation, 112 missing values remain across the panel (0.2% missingness). We handle missing data through forward-filling within the rolling PCA framework by Hirsa et al. (2024), ensuring each principal component extraction uses the most recent available information without introducing look-ahead bias.

3.3.4 Data Alignment and Panel Construction

All datasets are aligned to month-end dates. The final merged dataset for tactical allocation contains 278 monthly observations (November 2001 to December 2024) with complete coverage across sector returns, market weights, S&P 500 benchmark returns, risk-free rates, and regime labels. No observations are dropped due to missing data in the merged panel.

3.4 Descriptive Statistics

3.4.1 Sector Return Characteristics

Table 3.1 presents summary statistics for eleven GICS sectors and the S&P 500 benchmark over November 2001 to December 2024. Average annualized sector excess returns equal 8.53%, with the S&P 500 benchmark generating 9.32%. Individual sector performance exhibits substantial heterogeneity, ranging from Information Technology's strong performance to more modest returns in defensive sectors.

Table 3.1 Sector Descriptive Statistics of Excess Returns (2001–2024)

Sector	Exc. Return (%)	St. Dev (%)	Sharpe	Min	Max	Exc. Kurt	Skew
Communication Services	7.745	19.481	0.384	-15.619	32.936	3.808	0.307
Consumer Discretionary	11.537	18.996	0.577	-19.188	20.547	1.533	-0.012
Consumer Staples	9.467	11.610	0.782	-10.916	10.290	0.614	-0.448
Energy	11.460	24.951	0.437	-34.796	29.779	3.820	0.011
Financials	7.500	21.496	0.337	-26.309	22.358	2.977	-0.650
Health Care	9.673	13.550	0.684	-12.495	12.648	0.399	-0.280
Industrials	10.087	18.547	0.520	-19.182	18.095	1.811	-0.431
Information Technology	14.583	20.649	0.663	-17.789	22.297	0.788	-0.292
Materials	10.962	20.246	0.516	-22.073	17.726	1.109	-0.261
Real Estate	10.921	21.841	0.477	-31.954	34.996	6.686	-0.530
Utilities	8.669	15.097	0.553	-13.979	12.172	0.872	-0.679
S&P 500	10.087	15.102	0.639	-16.795	12.819	1.117	-0.563

Notes: This table reports descriptive statistics of annualized monthly excess returns relative to the one-month Treasury bill for sector portfolios over the period 2001–2024. Reported statistics include the mean excess return, standard deviation, Sharpe ratio, minimum, maximum, excess kurtosis, and skewness. Normality is tested using the Jarque–Bera test, and stationarity is tested using the augmented Dickey–Fuller test. All series reject normality and do not reject stationarity at the 5% significance level.

Risk-free rates average 1.54% annually over the full sample but vary substantially across subperiods, reflecting different monetary policy environments. The risk-free rate was on average 2.46% during training (2001-2008), 0.08% during validation (2009-2013), and 1.61% during test (2014-2024).

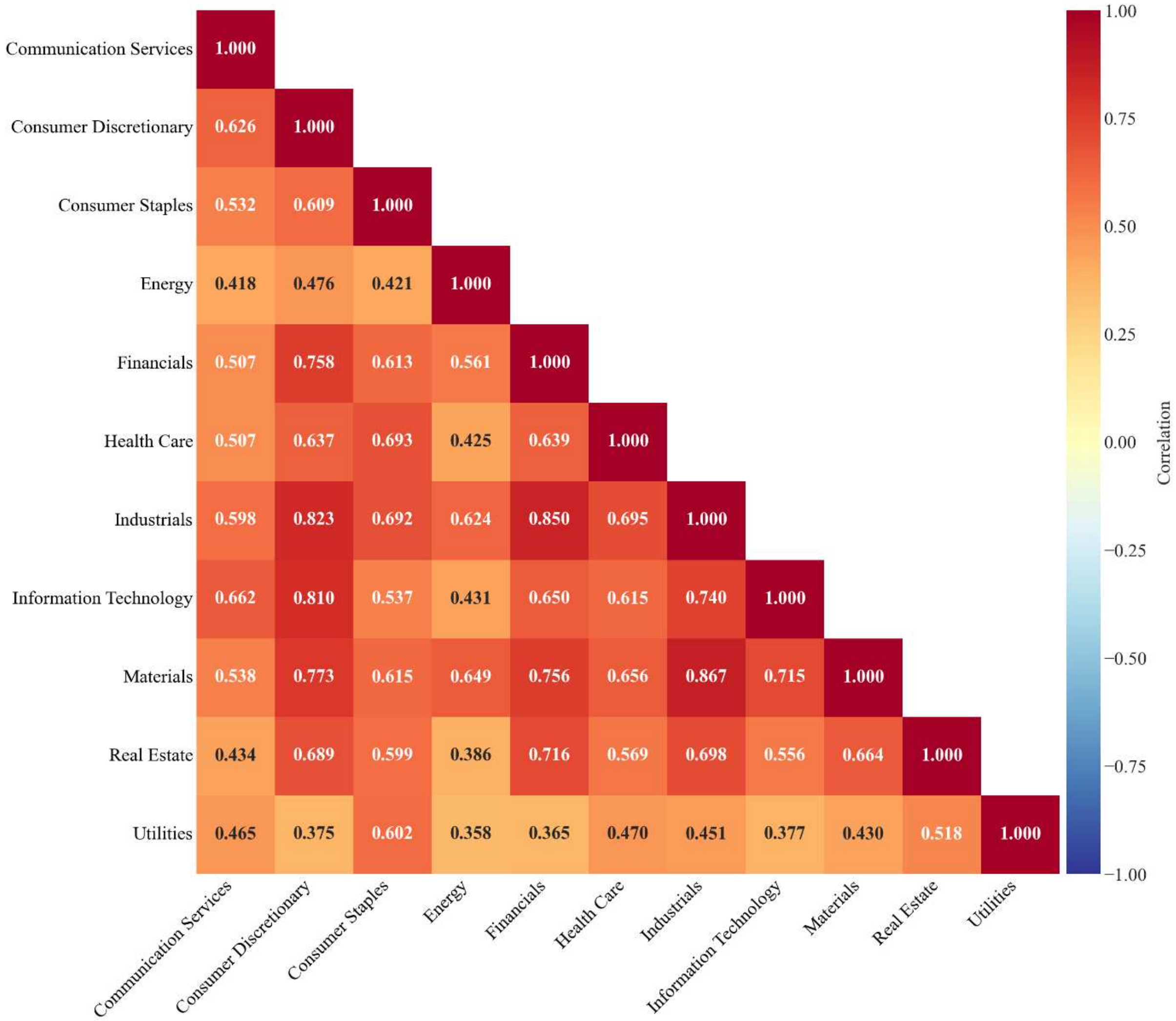
3.4.2 Sector Correlations and Market Structure

Pairwise correlations among sector excess returns average 0.59, indicating substantial common factor exposure across all sectors. However, correlation heterogeneity is economically meaningful. Cyclical sectors (Consumer Discretionary, Industrials, Materials, Financials) exhibit average correlations exceeding 0.64, while defensive sectors (Utilities, Energy) show correlations below 0.50 with other sectors. Utilities displays the lowest average correlation with other sectors (0.44), followed by Energy (0.48), reflecting commodity-specific and rate-sensitive drivers distinct from broader economic cycles.

The highest pairwise correlations emerge among economically sensitive sectors are Industrials–Materials (0.87), Financials-Industrials (0.85), and Consumer Discretionary-Industrials (0.82), reflecting their shared sensitivity to business cycle fluctuations. Conversely, the lowest correlations involve Utilities and Energy paired with other sectors, Energy-Utilities (0.36), Financials–Utilities (0.37), and Consumer Discretionary-Utilities (0.38). Notably, 73% of sector pairs exhibit correlations above 0.50, while no pairs fall below 0.30.

Information Technology maintains the largest average market-capitalization weight (20.3%), followed by Financials (15.6%), Health Care (13.3%), Consumer Discretionary (11.1%), and Industrials (10.0%). These five sectors collectively represent approximately 70% of S&P 500 market capitalization. Notably, sector weights exhibit substantial time variation: Information Technology's weight ranged from 15% in 2001 to over 28% by 2024, while Financials declined from 20% pre-2008 to 11% by 2024.

Figure 3.1 Sector Return Correlation, Full Sample



Notes: This figure reports pairwise correlations of monthly sector excess returns over the full sample. The color scale reflects the magnitude of correlation coefficients, with darker shades indicating stronger co-movement. Values on the diagonal equal one by construction.

3.4.3 Regime Distribution Preview

Section 5 presents the complete regime identification results. Preliminary descriptive analysis reveals that the 278-month allocation sample encompasses four distinct macroeconomic regimes: expansion, slowdown, stress, and crisis. The distribution exhibits concentration in expansion and slowdown conditions with sparse observations of crisis and stress periods.

4. Methodology

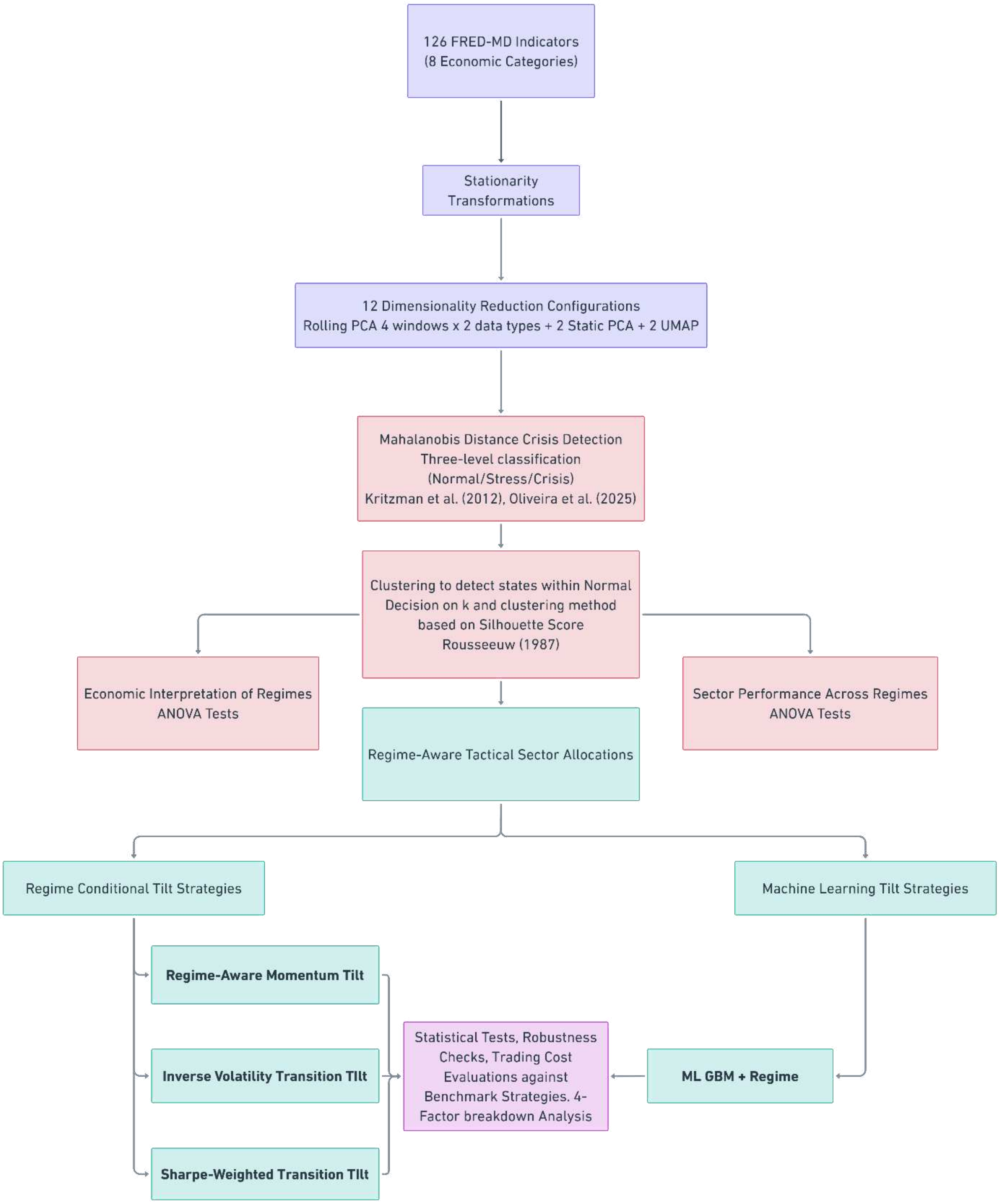
4.1 Overview

This study develops a regime-aware tactical sector allocation framework comprising two stages. Stage 1 employs unsupervised machine learning to identify regimes from macroeconomic indicators through three substages: (i) dimensionality reduction via rolling PCA, (ii) Mahalanobis distance-based crisis detection, and (iii) K-Means clustering of non-crisis periods. Stage 2 constructs tactical sector allocations conditioned on the identified regimes.

The framework prioritizes objective selection criteria at each decision point. We use F1-score against NBER recession dates for crisis detection configuration selection and Silhouette score for clustering validation. All model selection occurs exclusively on training data, ensuring genuine out-of-sample evaluation.

Figure 4.1 summarizes the full methodology. The remainder of this section follows the pipeline step-by-step. Section 4.2 describes how macroeconomic information is compressed into a small set of interpretable factors. Section 4.3 explains how extreme stress periods are identified and isolated. Section 4.4 shows how remaining observations are clustered into business-cycle regimes. Finally, Sections 4.5–4.6 describe how these regimes are mapped into constrained, implementable sector allocation rules.

Figure 4.1 Methodology Pipeline



Notes: This figure illustrates the two-stage methodology pipeline. Stage 1 processes macroeconomic indicators through stationarity transformations and dimensionality reduction (12 configurations tested), followed by Mahalanobis distance-based crisis detection (Normal/Stress/Crisis classification) and clustering of normal periods to identify distinct economic regimes. Regime distinctiveness and sector performance differences are validated via ANOVA tests. Stage 2 applies the identified regimes to tactical sector allocation through two parallel approaches: Regime-Conditional Tilt Strategies (Regime-Aware Momentum Tilt, Inverse Volatility Transition, Sharpe-Weighted Transition) and Machine Learning Tilt Strategies (GBM + Regime). All strategies are evaluated against benchmarks using statistical tests, robustness checks, transaction cost analysis, and 4-factor decomposition.

4.2 Dimensionality Reduction

Using 126 macroeconomic variables directly for regime detection faces challenges. Clustering algorithms perform poorly in high-dimensional spaces where distance metrics become less meaningful (Bellman, 1961). We address this through rolling Principal Component Analysis, which extracts linear combinations capturing maximum variance while adapting to evolving macroeconomic relationships.

Following Hirsa et al. (2024), we implement rolling PCA with 120-month windows applied separately to each of the eight FRED-MD economic categories. This category-level approach preserves economic interpretability: each extracted component corresponds to a distinct macroeconomic dimension (Output & Income, Labor Market, Housing, Consumption & Orders, Money & Credit, Interest & FX, Prices, Stock Market). The 120-month window balances stability against adaptivity to structural changes.

Standard PCA produces eigenvectors, and multiplying any eigenvector by -1 yields an equally valid solution. In rolling applications, this sign ambiguity is important as arbitrary sign flips between adjacent windows can occur, which would create discontinuities in extracted factors. Following Hirsa et al. (2024), we implement sign correction to ensure temporal consistency.

For each rolling window w and category c, we compute the first principal component as:

$$P_{1,c,t} = w_{c,t}^T \cdot x_{c,t}$$

where $w_{c,t}$ is the first eigenvector computed over window [t-w+1, t] and $x_{c,t}$ contains standardized macroeconomic variables in category *c* at time *t*. To ensure sign consistency, we apply correction after each window update:

$$w_{c,t} \leftarrow -w_{c,t} \text{ if } w_{c,t}^T \cdot w_{c,t-1} < 0$$

This correction aligns the current eigenvector with the previous period's direction, preventing spurious sign flips while preserving the mathematical validity of the principal component. This ensures that changes in the extracted principal components reflect genuine economic evolution rather than arbitrary mathematical rotations inherent to PCA.

We evaluated 12 dimensionality reduction configurations varying rolling versus static estimation, category-level versus full-variable extraction, and linear (PCA) versus non-linear (UMAP) methods. Complete configuration specifications, quality metrics, and sign correction methodology are reported in Appendix C.

4.3 Crisis Regime Identification

Crisis periods represent qualitatively distinct market states characterized by correlation spikes, volatility clustering, and breakdown of normal asset relationships (Ang & Bekaert, 2004). We identify crises separately from business cycle regimes for two reasons. (i) to enable defensive tactical positioning during extreme stress, (ii) to prevent crisis outliers from distorting subsequent clustering.

$$D_t = \sqrt{[(x_t - \mu_t)^T \cdot \Sigma_t^{-1} \cdot (x_t - \mu_t)]}$$

where x_t is the 8-dimensional vector of principal components at time t, μ_t is the mean vector computed over the expanding baseline window [1, t-1], and Σ_t is the covariance matrix over the same window. Intuitively, D_t measures how unusual the current joint configuration of macroeconomic conditions is relative to historical norms, capturing systemic stress rather than isolated indicator movements. The expanding window approach ensures the baseline adapts to structural changes while maintaining stability as more observations accumulate. A minimum of 60 months is required before computing meaningful distances, ensuring sufficient observations for stable covariance estimation.

We implement a three-level classification separating observations into normal, stress, and crisis categories based on Mahalanobis distance percentiles:

- **Normal:** Distance below the 90th percentile threshold
- **Stress:** Distance at or above 90th percentile but below 95th percentile
- **Crisis:** Distance at or above the 95th percentile

This hierarchical structure enables differentiated tactical responses, over stress and crisis periods. The thresholds are calibrated exclusively on training data, then applied unchanged to validation and test periods. The different configurations are then evaluated based on training F1-scores, which are the stress and crisis periods combined against NBER recession dates.

4.4 Regime Clustering

After separating crisis observations, we partition remaining "normal" periods into distinct business cycle phases using K-Means clustering. This captures the finding that expansions exhibit different sector leadership than slowdowns, enabling regime-conditional tactical tilts.

We select K-Means with k=2 clusters based on Silhouette score (Rousseeuw, 1987), which measures within-cluster cohesion against between-cluster separation:

$$s_i = \frac{b_i - a_i}{\max a_i, b_i}$$

where a_i is the mean distance from observation i to other observations in its cluster and b_i is mean distance to observations in the nearest neighboring cluster. The optimal configuration achieves Silhouette score of 0.180, the highest among 175 tested combinations spanning seven clustering algorithms and multiple dimension reduction approaches. All clustering method descriptions are provided in Appendix J, Table J.2.

The modest Silhouette score reflects the inherent nature of business cycles, macroeconomic states transition gradually rather than discretely (Stock & Watson, 1999). In this context, Silhouette score assesses whether macroeconomic states form meaningfully distinct environments rather than enforcing artificially sharp regime boundaries. Section 5.1 validates that despite moderate geometric separation, the identified regimes exhibit statistically significant differences in both macroeconomic characteristics and sector returns.

The final regime taxonomy combines crisis detection with clustering, yielding four macroeconomic states: Expansion, Slowdown, Stress, and Crisis.

4.5 Tactical Sector Allocation Strategies

We implement nine portfolio strategies comprising two passive benchmarks, two momentum-based strategies, four regime-aware strategies, and one volatility-managed portfolio. All regime-aware and tilt-based strategies maintain long-only constraints and maximum ±5 percentage point deviations from market-capitalization weights. These constraints ensure portfolios remain investable and tracking errors stay bounded.

Table 4.1 Portfolio Strategy Summary

#	Strategy	Core Logic	Tactical	Regime-Based
1	S&P 500	Buy-and-Hold	No	No
2	Equal Weight	1/11 per Sector	No	No
3	Cross-Sectional Momentum	Long-Short 12-1 Momentum	No	No
4	Momentum Tilt	12-1 Momentum Tilts	Yes	No
5	Volatility Managed	Risk-Based	No	Implicit
6	ML GBM + Regime	GBM Predicted Z-scores Ranked	Yes	Yes
7	Sharpe-Weighted Tilt	Expected Sharpe Ratio per Regime	Yes	Yes
8	Inverse Volatility Transition Tilt	Expected Inverse Volatility per Regime	Yes	Yes
9	Regime-Aware Momentum Tilt	Scaled Momentum based on Regime	Yes	Yes

Notes: This table summarizes the portfolio strategies considered in the study and their core investment logic. "Tactical" indicates whether the strategy actively adjusts portfolio weights through ±5 percentage point tilts from market-capitalization weights while maintaining long-only positions based on signals. "Regime-based" indicates whether the strategy explicitly incorporates regime classification or regime transition probabilities. "Yes" denotes direct regime conditioning, "Implicit" denotes strategies that respond to regime changes indirectly through realized volatility scaling, and "No" denotes strategies that do not incorporate regime information.

4.5.1 Baseline Strategies

Benchmark (S&P 500): Passive investment in the S&P 500 total return index.

Benchmark (Equal-Weight): Allocates 1/11 of portfolio capital to each GICS sector at every monthly rebalancing point. This mechanically overweighting smaller sectors while underweighting larger sectors relative to market capitalization.

4.5.2 Traditional Active Strategies

Two momentum-based strategies implement the classic cross-sectional momentum anomaly (Jegadeesh & Titman, 1993).

Cross-Sectional Momentum (Long-Short): Computes cumulative returns over months t-12 to t-1, excluding the most recent month to avoid short-term reversal and takes long positions in the top 3 sectors and short positions in the bottom 3, equally weighted at ±33.3%.

Momentum Tilt: Long-only variant applying momentum signals as tilts relative to market-capitalization weights. Using the same 12-1 momentum lookback, tilts are computed using rank-based scaling, mapping the highest momentum sector to +5% and the lowest to -5%.

Volatility-Managed Portfolio: Following Moreira and Muir (2017), dynamically scale market exposure inversely to realized volatility. At each month t, the strategy estimates portfolio volatility using a 24-month rolling window. The exposure scalar is computed as:

$$\text{Exposure}_t = \frac{\sigma_{\text{target}}}{\sigma_{\text{realized},t}}$$

Target Volatilities tested are $\sigma_{\text{target}} \in \{8, 10, 12, 14, 16\}$, with exposure clipped to [50%, 150%]. This approach is implicitly regime-aware because crisis and stress periods typically exhibit elevated volatility, triggering automatic de-risking into the risk-free asset. We evaluate all five target volatility levels in training and move forward with the portfolio with the highest Sharpe ratio.

4.5.3 Regime-Aware Strategies

ML GBM + Regime Interactions: This strategy is included to test whether flexible nonlinear prediction improves upon economically structured regime-conditional allocation rules. It uses Gradient Boosting Machine (GBM) models to predict next-month sector z-scores using an enhanced feature set incorporating regime information. The feature set is explained in appendix J.

Separate GBM models are trained for each sector using an expanding window approach with minimum 24 months of training data. Models are retrained every 12 months to incorporate new observations while avoiding excessive retraining costs. Predicted z-scores are converted to portfolio tilts using rank-based scaling with ±5 percentage point constraints.

Sharpe-Weighted Transition Tilt: Uses an expanding transition matrix to compute expected risk-adjusted returns. For each sector, expected returns and volatilities are probability-weighted across potential next-period regimes:

$$E[r_s|\text{regime}_i] = \sum_k P(\text{regime}_k \mid \text{regime}_i) \times \mu_{s,k}$$

$$E[\sigma_s \mid \text{regime}_i] = \sum_k P(\text{regime}_k \mid \text{regime}_i) \times \sigma_{s,k}$$

Expected Sharpe ratios are then computed as $E[r_s]/E[\sigma_s]$, and sectors are tilted based on rank-ordered expected Sharpe ratios.

Inverse Volatility Transition Tilt: Exploits the low-volatility anomaly in a regime-aware framework. Computes expected volatility by weighing regime-conditional volatilities by transition probabilities from an expanding transition matrix. Sectors with lower expected volatility receive positive tilts.

Regime-Aware Momentum Tilt: Addresses momentum crashes (Daniel & Moskowitz, 2016) by scaling exposure based on transition probabilities to adverse regimes:

$$\text{Momentum Exposure}_t = \begin{cases} 0.25 & \text{if } P(\text{Crisis}_t) + P(\text{Stress}_t) > 0.50 \\ 0.50 & \text{if } P(\text{Crisis}_t) + P(\text{Stress}_t) > 0.30 \\ 0.75 & \text{if } P(\text{Crisis}_t) + P(\text{Stress}_t) > 0.15 \\ 1.00 & \text{Otherwise} \end{cases}$$

Additionally, if currently in the Crisis regime, momentum exposure is further halved to account for typical momentum reversals during crisis conditions. The rule operationalizes the insight that momentum crashes tend to coincide with elevated transition risk into adverse regimes, warranting gradual exposure reduction rather than binary on/off positioning.

4.6 Implementation Details

4.6.1 Transaction Costs

Transaction costs are modeled as 10 basis points per dollar of turnover (one-way), representing combined bid-ask spreads, market impact, and execution slippage for institutional-scale sector ETF portfolios:

$$TC_t = 0.0010 \times \Sigma_i |w_{i,t} - w_{i,t-1}^{\text{drift}}|$$

where $w_{i,t}$ represents target weight and $w_{i,t-1}^{\text{drift}}$ represents weight after drift but before rebalancing. The 10 basis point assumption is conservative relative to achievable costs for liquid sector ETFs, providing implementation cushion for larger portfolios or less liquid instruments. All reported performance metrics reflect net of transaction cost returns.

4.6.2 Rebalancing Protocol

Primary results use monthly rebalancing at month-end, ensuring regime signals identified from month-end data are acted upon at the subsequent month's start. Monthly frequency balances signal responsiveness against transaction cost accumulation. Section 5.3 examines sensitivity to alternative transaction costs to assess robustness of findings to implementation choices.

4.6.3 Weight Constraint Enforcement

All tilt-based strategies enforce hard constraints through an iterative algorithm that applies tilts to base market-capitalization weights while ensuring: (i) maximum tilt of ±5 percentage points from market weights, (ii) minimum weight of 0.5% per sector, and (iii) weights sum to unity. The algorithm scales tilts iteratively until all constraints are satisfied.

4.7 Performance Evaluation and Statistical Testing

4.7.1 Risk-Adjusted Metrics

We evaluate portfolios using standard risk-adjusted metrics computed on excess net returns:

Sharpe Ratio: SR = $(\mu_p - r_x) / \sigma_p$, where μ_p is annualized portfolio return, r_x is annualized risk-free rate, and σ_p is annualized volatility. The Sharpe ratio measures excess return per unit of total risk.

Sortino Ratio: Similar to Sharpe but using downside deviation (standard deviation of negative returns only) in the denominator, focusing on harmful volatility while ignoring upside variation.

Maximum Drawdown: Largest peak-to-trough decline in cumulative portfolio wealth, measuring worst-case loss experience over the evaluation period.

Calmar Ratio: Annualized return divided by absolute maximum drawdown, combining return generation with drawdown risk in a single metric.

4.7.2 Statistical Significance Testing

Given the well-documented difficulty of distinguishing skill from luck in portfolio performance (Harvey et al., 2016), we employ multiple testing approaches to assess whether observed Sharpe ratios differ from the S&P 500's.

Jobson-Korkie Test: Tests Sharpe ratio differences using asymptotic z-statistics with correlation adjustment (Jobson and Korkie, 1981; Memmel, 2003). This parametric test complements the bootstrap approach, offering computational efficiency and theoretical tractability.

Ledoit-Wolf Bootstrap: Constructs confidence intervals for Sharpe ratio differences using circular block bootstrap with automatic block length selection (Ledoit and Wolf, 2008). This non-parametric approach accounts for serial correlation and non-normality in returns, providing robust inference without distributional assumptions.

Transaction Cost Breakdown: Tests multiple transaction costs to evaluate the stability of returns and Sharpe ratios, identifying break-even transaction costs where strategy outperformance disappears.

4.8 Out-of-Sample Evaluation Framework

Rigorous temporal separation ensures performance assessment reflects implementable strategy returns rather than in-sample overfitting. The framework employs three distinct periods:

Training Period (November 2001 – December 2008): All model estimation, regime-specific return statistics, clustering parameters, crisis thresholds, ML model training, and strategy calibration uses exclusively this data.

Validation Period (January 2009 – December 2013): We compare performance across the chosen strategies during validation to evaluate risk-adjusted returns before final out-of-sample testing. ML models are retrained on expanding windows during this period.

Test Period (January 2014 – December 2024, 132): Genuine out-of-sample evaluation where selected strategies from validation are applied, providing unbiased assessment of implementable performance.

This three-way split follows machine learning best practices (Hastie et al., 2009) and addresses the pervasive problem of overfitting in quantitative finance backtesting (Bailey et al., 2014). The key principle is that the test period must never influence any modeling choice, it exists solely to provide an honest assessment of strategy quality.

4.9 Limitations

Sample Constraints: The sample encompasses only two major crises (2008–2009 GFC, 2020 COVID), limiting statistical power for extreme event evaluation. Crisis and stress regimes contain sparse observations and might constrain regime-specific estimation precision.

ML Model Complexity: The GBM model introduces additional parameters and potential overfitting risk despite regularization. Model retraining frequency (12 months) represents a balance between adaptation and stability that may not be optimal in all market conditions.

Transaction Cost Assumptions: Uniform 10 basis points abstract from time-varying spreads that widen during stress periods precisely when regime strategies most actively rebalance.

Geographic Scope: Analysis restricts to GICS sectors within the S&P 500 universe, precluding assessment of international generalizability.

Despite these limitations, rigorous temporal separation preventing data snooping, conservative transaction cost assumptions, objective model selection criteria, and comprehensive statistical testing provide reasonable confidence that documented patterns, whether positive or null, should reflect genuine strategy characteristics rather than methodological artifacts.

5. Results

This section presents empirical findings. Section 5.1 reports regime identification results. Section 5.2 presents portfolio performance. Section 5.3 examines robustness.

5.1 Regime Identification Results

5.1.1 Configuration Selection

Following the methodology in Section 4, we evaluate 12 dimensionality reduction configurations and 175 clustering specifications. Table 5.1 provides details of the selected regime identification configuration.

Table 5.1 Optimal Regime Identification Configuration

Component	Selection	Details
Dimensionality Reduction	RollPCA_8cat_w120	Rolling PCA, 8 categories, 120-month window
Clustering Algorithm	K-Means	Highest Silhouette score
K (normal regimes)	2	2 normal + Stress + Crisis
Regime Labels	Crisis, Stress, Expansion, Slowdown	Hierarchical structure
Silhouette Score	0.180	Highest among 175 configurations

Notes: This table reports the final regime identification configuration selected from 175 tested model specifications. Rolling principal component analysis with a 120-month window is applied separately to each of the eight FRED-MD categories to capture time-varying macroeconomic relationships. The two-stage framework first identifies crisis and stress periods using Mahalanobis distance and then applies K-means clustering with two clusters to the remaining observations to distinguish expansion from slowdown regimes. This hierarchical structure mitigates the influence of extreme observations on business cycle classification.

The selected configuration achieves strong crisis detection (F1 = 0.606, Precision = 0.909) while producing the best separated clusters (Silhouette = 0.180). Rolling PCA outperforms static methods and UMAP; K-Means outperforms alternative clustering algorithms including Hidden Markov Models, Gaussian Mixture Models and more. Detailed comparisons are provided in Appendix C.

5.1.2 Final Regime Structure

The final regime taxonomy combines crisis detection (Stress, Crisis) with normal-period clustering (Expansion, Slowdown), yielding a four-regime structure. Table 5.2 presents the distribution across periods.

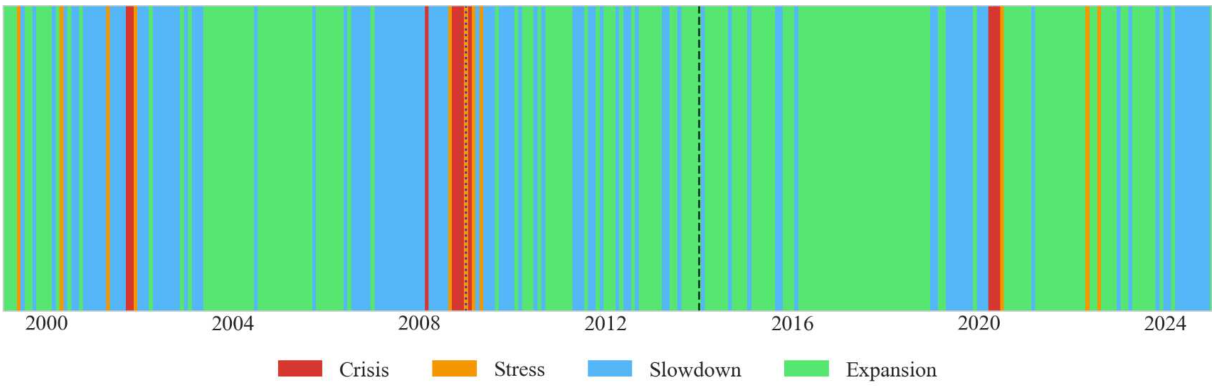
Table 5.2 Final Regime Distribution

Regime	Total Months	Percentage (%)	Train	Validation	Test
Expansion	165	59.352	39	33	93
Slowdown	97	34.892	40	24	33
Stress	8	2.878	3	2	3
Crisis	8	2.878	4	1	3
Total	278	100.000	86	60	132

Notes: This table reports the distribution of identified regimes over the full sample period from 2001 to 2024. Expansion and slowdown regimes account for the majority of observations, comprising 59.4% and 34.9% of the sample, respectively. Stress and crisis regimes each represent 2.9% of total observations. The sample is divided into training, validation, and test subsamples used for model estimation and evaluation. The training period spans 2001–2008, the validation period spans 2009–2013, and the test period spans 2014–2024.

The distribution is consistent with business cycle theory: Expansion and Slowdown dominate (94.2%), while Crisis and Stress are appropriately rare (2.9% each). The test period contains all four regimes, enabling evaluation across diverse conditions. Figure 5.1 visualizes the regime classifications over time. The framework successfully identifies major economic episodes, crisis classifications align with the 2001 dot-com recession, the 2007–2009 Global Financial Crisis, and the COVID-19 shock in early 2020.

Figure 5.1 Regime Identification Timeline



Notes: This figure plots the time series of identified macroeconomic regimes over the full sample period from 1999 to 2024. Shaded regions indicate crisis, stress, slowdown, and expansion regimes. Vertical dashed lines denote the training, validation, and test sample splits. Crisis and stress regimes align closely with major economic downturns, including the dot-com recession, the global financial crisis, and the Covid-19 shock, supporting the economic relevance of the regime identification approach.

5.1.3 Statistical Validation of Regime Distinctiveness

While Silhouette scores provide geometric validation of cluster quality, we require statistical evidence that the identified regimes represent genuinely distinct macroeconomic states. Table 5.3 presents one-way ANOVA results testing whether the extracted principal components differ significantly across regimes.

Table 5.3 ANOVA Results for Regime Distinctiveness, Full Sample

Macroeconomic Component	F-statistic	p-value	η^2 (Effect Size)	Significance
Output & Income	36.519	<0.001	**0.262**	***
Labor Market	22.397	<0.001	0.179	***
Housing	21.648	<0.001	0.174	***
Consumption & Orders	27.906	<0.001	0.214	***
Money & Credit	6.049	<0.001	0.056	***
Interest & FX	22.533	<0.001	0.180	***
Prices	4.555	0.004	**0.042**	**
Stock Market	17.387	<0.001	0.145	***

*Notes: This table reports analysis of variance (ANOVA) results testing whether macroeconomic components differ across identified regimes using the full sample. The F-statistic tests the null hypothesis of equal means across regimes. The effect size, η^2, measures the proportion of total variance explained by regime membership. Statistical significance is denoted by ***, **, and * corresponding to the 1%, 5%, and 10% levels, respectively. All components exhibit statistically significant differences across regimes at conventional levels, with an average effect size of 0.157. The largest and smallest effect sizes are highlighted in bold.*

The ANOVA results provide strong statistical evidence for regime distinctiveness. All macroeconomic components show significant differences across regimes ($p < 0.05$), with seven highly significant ($p < 0.001$). The four regimes exhibit economically interpretable profiles: Expansion features above-average growth (+1.22 pp.) and below-average volatility (VIX -2.82); Crisis captures extreme stress with VIX +23.04 above average and growth -9.21% below average. Complete Macroeconomic characterization is reported in Appendix B.

Table 5.4 presents average monthly sector returns across regimes. Ten of eleven sectors exhibit statistically significant regime effects ($p < 0.05$), with Financials and Real Estate showing the strongest sensitivity, spreads of 14.9 and 15.7 percentage points respectively between Expansion and Crisis. These findings provide empirical support for regime-based tactical sector allocation.

Table 5.4 Average Monthly Sector Returns by Regime, Full Sample

Sector	Expansion	Slowdown	Stress	Crisis	F-stat	p-value	Significance
Communication Services	0.950	0.376	-0.860	-3.816	2.195	0.089	*
Consumer Discretionary	1.568	-0.085	**5.211**	-5.658	7.848	<0.001	***
Consumer Staples	0.874	0.725	0.729	-2.509	2.653	0.049	**
Energy	1.562	-0.271	-1.271	-5.376	3.600	0.014	**
Financials	1.718	**-0.413**	0.248	**-13.175**	18.159	<0.001	***
Health Care	1.122	0.235	-0.588	**-2.279**	2.996	0.031	**
Industrials	1.541	0.024	1.592	-8.099	9.968	<0.001	***
Information Technology	**1.745**	0.144	4.320	-5.195	5.430	0.001	***
Materials	1.449	-0.053	2.266	-6.829	6.348	<0.001	***
Real Estate	1.286	0.502	3.735	-12.432	13.990	<0.001	***
Utilities	**0.826**	**1.020**	**-3.185**	-3.787	5.155	0.002	***
S&P 500	1.452	0.109	1.179	-5.478	8.148	<0.001	***

*Notes: This table reports average monthly sector excess returns for each identified regime over the full sample period. Expansion, slowdown, stress, and crisis regimes are defined based on the regime identification procedure described in the text. The F-statistic and associated p-value test the null hypothesis that mean returns are equal across regimes within each sector. Statistical significance is denoted by ***, **, and * corresponding to the 1%, 5%, and 10% levels, respectively. Sector returns vary significantly across regimes, with most sectors exhibiting statistically significant differences. Expansion regimes are associated with uniformly positive returns, while crisis* *regimes are associated with uniformly negative returns. The highest and lowest sector returns within each regime are highlighted in bold.*

5.2 Portfolio Strategy Performance

This section presents out-of-sample performance (January 2014–December 2024). All strategy parameters were fixed using training and validation data only, ensuring genuine out-of-sample evaluation.

5.2.1 Test Period Results

Table 5.5 summarizes performance metrics for all strategies during the test period.

Table 5.5 Test Period Performance Summary

Strategy	Ann. Return	Ann. Vol	Sharpe	Sortino	Max DD	Calmar	TC
Regime-Aware Momentum Tilt	**11.492**	14.364	**0.801**	0.701	-21.943	**0.524**	0.124
Momentum Tilt	11.323	14.394	0.788	**0.711**	-21.799	0.519	0.122
Inverse Volatility Transition Tilt	11.072	14.144	0.784	0.706	-24.341	0.455	0.043
Sharpe-Weighted Transition Tilt	10.981	14.506	0.758	0.680	-24.926	0.441	0.125
S&P 500	10.746	14.825	0.726	0.619	-25.413	0.423	**0.000**
Volatility Managed 10%	7.825	**11.528**	0.679	0.556	**-17.634**	0.444	0.094
ML GBM + Regime Interactions	9.986	**15.067**	0.664	0.578	-27.972	0.357	0.325
Equal Weight	8.477	14.451	0.587	0.513	-24.357	0.348	0.029
Momentum Long-Short	**-2.378**	14.716	**-0.161**	**-0.169**	**-34.722**	**-0.068**	**0.964**

Notes: This table reports test period performance metrics for all portfolio strategies in terms of excess returns. All returns are reported net of transaction costs (TC). Annualized return, volatility, Sharpe ratio, Sortino ratio, maximum drawdown (DD), Calmar ratio, and transaction costs are expressed in percentage terms, except for risk-adjusted ratios. The test period spans 132 months from January 2014 to December 2024. Transaction costs are measured as annualized turnover-based costs as a percentage of portfolio value. The highest and lowest values in each column are highlighted in bold.

The results reveal a clear performance hierarchy. Four regime-aware strategies outperform the S&P 500 benchmark in the out-of-sample evaluation, with Regime-Aware Momentum Tilt achieving the highest Sharpe ratio (0.801) and generating annualized returns of 11.49%, approximately 74 basis points above the benchmark. The Momentum Tilt (0.788), Inverse Volatility Transition Tilt (0.784), and Sharpe-Weighted Transition Tilt (0.758) complete the set of outperforming strategies.

Beyond higher returns, the top tactical strategies achieve meaningfully lower maximum drawdowns than the benchmark. Regime-Aware Momentum Tilt and Momentum Tilt experience drawdowns of -21.94% and -21.80% respectively, compared to -25.41% for the S&P 500, an improvement of approximately 3.5 percentage points.

The Inverse Volatility Transition Tilt merits particular attention for its efficiency. Despite achieving comparable risk-adjusted returns (Sharpe 0.784), it generates only 0.043% in annual transaction costs, roughly one-third of the other regime strategies. This cost efficiency stems from its 42.7% annual turnover versus 124% for Regime-Aware Momentum Tilt, making it attractive for larger portfolios where market impact concerns dominate.

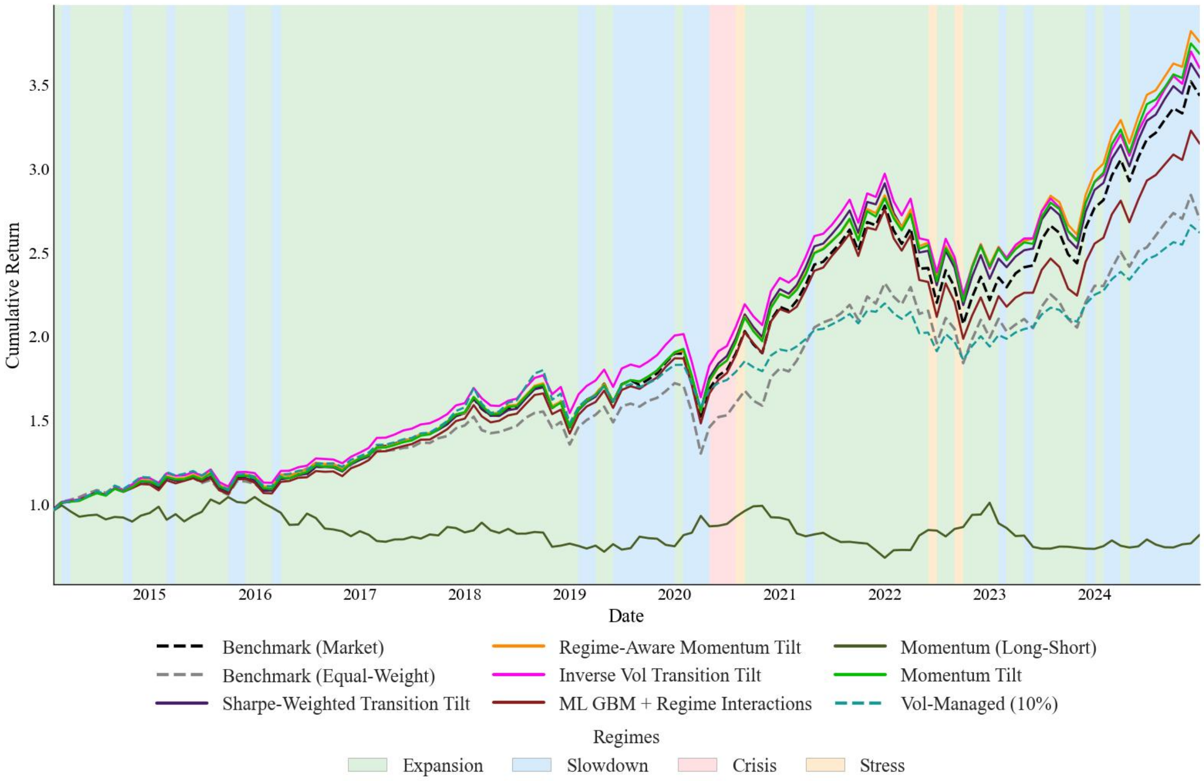
The volatility-managed strategy presents a mixed picture. It achieves the lowest maximum drawdown (-17.63%) and lowest volatility (11.53%) among all strategies, confirming that volatility scaling successfully reduces risk. However, this risk reduction comes at substantial cost to returns. The strategy's 7.83% annual return falls nearly 3 percentage points below the benchmark, resulting in a Sharpe ratio of 0.679 that underperforms passive investment. The return sacrifice from reduced market exposure during high-volatility periods outweighs the volatility reduction benefit. This finding aligns with Cederburg et al.'s (2020) critique that volatility timing benefits are period dependent.

Traditional active strategies perform poorly. The equal-weight benchmark achieves a Sharpe ratio of only 0.587, underperforming both the market-cap benchmark and all regime strategies. Moreover, the cross-sectional momentum long-short strategy, which generates deeply negative returns (-2.38% annually) with a Sharpe ratio of -0.161. This strategy also exhibits the worst maximum drawdown (-34.72%) and highest transaction costs (0.964% annually), representing a comprehensive failure across all performance dimensions.

The ML GBM + Regime strategy underperforms the benchmark despite incorporating machine learning and regime information. Its Sharpe ratio of 0.664 falls below passive investment (0.726), while generating the highest transaction costs among long-only strategies (0.325% annually) and the worst maximum drawdown (-27.97%). The added model complexity does not translate to improved risk-adjusted returns.

Figure 5.2 visualizes cumulative performance over the test period. The regime strategies maintain outperformance throughout rather than generating returns from isolated periods. During the COVID-19 crisis (March 2020), all strategies experience drawdowns, but regime strategies recover faster during the subsequent expansion. The momentum long-short strategy (green line) consistently trails all alternatives, while the volatility-managed strategy (blue line) shows a flatter trajectory reflecting its reduced market exposure.

Figure 5.2 Cumulative Returns During the Test Period



Notes: This figure plots cumulative excess returns for all portfolio strategies over the test period from January 2014 to December 2024. The market and equal-weight benchmarks are included for comparison. Background shading indicates the identified macroeconomic regimes, including expansion, slowdown, stress, and crisis periods. The figure illustrates the relative performance of regime-aware and benchmark strategies over time.

5.2.2 Statistical Significance Testing

Table 5.6 presents formal statistical tests of strategy performance. Given the well-documented difficulty of distinguishing skill from luck in portfolio returns (Harvey et al., 2016), we employ multiple testing approaches.

Table 5.6 Sharpe Ratio Significance Tests

Strategy	Sharpe	p (SR≠0)	JK p-value (vs S&P)	LW p-value (vs S&P)
Regime-Aware Momentum Tilt	0.801	**0.002*** **	**0.035** **	**0.026** **
Momentum Tilt	0.788	**<0.001*** **	0.080*	0.082*
Inverse Volatility Transition Tilt	0.784	**<0.001*** **	**0.042** **	**0.042** **
Sharpe-Weighted Transition Tilt	0.758	**<0.001*** **	0.219	0.303
S&P 500	0.726	**<0.001*** **	-	-
Volatility Managed 10%	0.679	**0.003*** **	0.656	0.645
ML GBM + Regime Interactions	0.664	**0.006*** **	**0.009*** **	**0.026** **
Equal Weight	0.587	**0.010** **	**0.042** **	0.095*
Momentum Long-Short	-0.161	0.575	0.071*	0.064*

*Notes: This table reports statistical tests of Sharpe ratios for all portfolio strategies. The column p(SR≠0) reports p-values for tests of whether the Sharpe ratio differs from zero using heteroskedasticity- and autocorrelation-consistent standard errors. JK p-values report results from the Jobson–Korkie (1981) test with the Memmel (2003) correction for comparisons against the S&P 500. LW p-values report results from the Ledoit–Wolf circular block bootstrap using 10,000 iterations. Statistical significance is denoted by ***, **, and * corresponding to the 1%, 5%, and 10% levels, respectively. Entries highlighted in bold indicate significance at the 5% level.*

Statistical tests provide strong evidence for regime strategy outperformance. Two strategies, Regime-Aware Momentum Tilt and Inverse Volatility Transition Tilt, significantly outperform the S&P 500 benchmark at the 5% level under both testing methodologies. Regime-Aware Momentum Tilt achieves p = 0.035 (Jobson-Korkie) and p = 0.026 (Ledoit-Wolf) and Inverse Volatility Transition Tilt achieves p = 0.042 under both tests. The consistency across parametric and non-parametric approaches strengthens the confidence in the outperformance.

The comparison between Momentum Tilt and Regime-Aware Momentum Tilt is particularly informative. Both strategies generate similar returns (11.32% vs. 11.49%) and Sharpe ratios (0.788 vs. 0.801), but only the regime-aware variant achieves statistical significance at the 5% level. Momentum Tilt falls just short with p-values of approximately 0.08. This marginal difference suggests that regime conditioning provides some meaningful improvement in signal reliability, the regime-aware strategy's outperformance is less likely to be attributable to chance.

The ML GBM strategy shows statistically significant underperformance relative to the benchmark (JK p = 0.009), indicating that its Sharpe ratio of 0.664 is below the benchmark's 0.726.

5.2.3 Regime Conditional Performance

Table 5.7 decomposes Sharpe ratios by regime state, to reveal where performance comes from.

Table 5.7 Sharpe Ratios by Regime, Test Period

Strategy	Overall	Expansion	Slowdown	Stress	Crisis
Regime-Aware Momentum Tilt	**0.801**	0.570	1.047	0.696	**4.434**
Momentum Tilt	0.788	0.549	1.052	0.730	4.289
Inverse Volatility Transition Tilt	0.784	**0.576**	1.026	0.204	4.053
Sharpe-Weighted Transition Tilt	0.758	0.511	**1.065**	0.415	4.348
S&P 500	0.726	0.521	0.930	0.295	4.121
Volatility Managed 10%	0.679	0.489	0.952	0.197	3.872
ML GBM + Regime Interactions	0.664	0.421	0.961	**0.093**	4.434
Equal Weight	0.587	0.456	0.587	0.696	3.326
Momentum Long-Short	**-0.161**	**-0.440**	**0.248**	**21.756**	**-1.353**

Notes: This table reports Sharpe ratios for each portfolio strategy computed separately by identified regime over the test period from January 2014 to December 2024. Regime-specific Sharpe ratios are calculated using monthly excess returns within expansion, slowdown, stress, and crisis periods. The test period contains 93 expansion months, 33 slowdown months, 3 stress months, and 3 crisis months. The overall Sharpe ratio is reported for comparison. The highest and lowest Sharpe ratios within each regime are highlighted in bold.

The decomposition reveals that regime-aware strategies outperform most notably during Slowdown periods. Regime-Aware Momentum Tilt achieves a Sharpe ratio of 1.047 during Slowdown versus the benchmark's 0.930, a 12.6% improvement. Since Slowdown comprises 25% of test observations, this consistent outperformance contributes meaningfully to overall results.

The strong Slowdown performance reflects an interesting market dynamic during the test period. The 2024 AI boom was classified as Slowdown in our framework due to weak macroeconomic data, yet equity returns, particularly in Technology, were exceptionally strong. During Expansion periods (70% of observations), regime strategies modestly outperform the benchmark. Regime-Aware Momentum Tilt achieves 0.570 versus 0.521 for the S&P 500.

The Crisis results require careful interpretation. All strategies exhibit high Sharpe ratios (3.3–4.4), reflecting the sharp V-shaped recovery following Covid-19. With only 3 crisis months in the test period, these figures are driven by a single episode and should not be over-interpreted. The more relevant observation is that regime strategies entered the crisis with reduced exposure and recovered quickly, limiting permanent capital impairment.

Comparing Momentum Tilt to Regime-Aware Momentum Tilt reveals where regime conditioning adds value. Both strategies perform very similarly, the standalone momentum signal is effective. The regime-aware variant's advantage accumulates through many small improvements in adverse conditions rather than from a single dramatic save.

5.3 Robustness Checks

This section examines whether the documented outperformance survives alternative assumptions about implementation costs and assesses the practical use of each strategy.

5.3.1 Sensitivity to Tilt Magnitude

The regime-aware strategies implemented in this thesis rely on constrained sector tilts relative to market-capitalization weights. To assess whether the documented performance improvements are sensitive to this design choice, this section evaluates the robustness of results to alternative tilt magnitudes.

Table 5.8 Sensitivity of Sharpe Ratios to Maximum Tilt Magnitude

Strategy/Max Tilt	2.5%	5%	7.5%	10%	12.5%	15%
Regime-Aware Momentum Tilt	**0.792**	**0.801**	**0.804**	**0.801**	**0.797**	**0.787**
Momentum Tilt	0.784	0.788	0.788	0.784	0.779	0.771
Inverse Volatility Transition Tilt	0.779	0.784	0.779	0.770	0.758	0.747
Sharpe-Weighted Transition Tilt	0.767	0.758	0.749	0.736	0.722	0.706
S&P 500	0.726	0.726	0.726	0.726	0.726	0.726
Volatility Managed 10%	0.679	0.679	0.679	0.679	0.679	0.679
ML GBM + Regime Interactions	0.714	0.664	0.662	0.586	0.554	0.519
Equal Weight	0.587	0.587	0.587	0.587	0.587	0.587
Momentum Long-Short	**-0.161**	**-0.161**	**-0.161**	**-0.161**	**-0.161**	**-0.161**

Notes: This table reports test period Sharpe ratios for each portfolio strategy across alternative maximum tilt magnitudes. Tilt magnitude refers to the maximum allowed deviation from benchmark sector weights. All Sharpe ratios are computed using excess returns net of transaction costs. The test period spans 2014 to 2024. The highest and lowest Sharpe ratios within each tilt magnitude are highlighted in bold.

Table 5.8 shows that regime-aware strategies exhibit stable performance across a range of moderate tilt magnitudes, with Sharpe ratios peaking between ±5% and ±10%. More aggressive tilts lead to a gradual deterioration in risk-adjusted performance, reflecting increased volatility and implementation frictions. In contrast, strategies that rely less on regime information, particularly machine-learning-based allocations, display pronounced sensitivity to tilt size. These results support the use of conservative tilt constraints and confirm that the baseline ±5% specification captures most of the attainable performance gains without incurring excessive risk.

5.3.2 Transaction Cost Sensitivity

Tactical allocation strategies must generate sufficient outperformance to overcome implementation costs. Table 5.9 presents Sharpe ratios under alternative transaction cost assumptions ranging from 0 to 50 basis points per unit turnover.

Table 5.9 Sharpe Ratios Under Alternative Transaction Cost Assumptions

Strategy	0 bps	5 bps	10 bps	20 bps	30 bps	50 bps	Break-Even (bps)
Regime-Aware Momentum Tilt	**0.810**	**0.806**	**0.801**	**0.793**	**0.784**	0.767	97.536
Momentum Tilt	0.796	0.792	0.788	0.779	0.771	0.754	82.425
Inverse Volatility Transition Tilt	0.787	0.785	0.784	0.781	0.778	**0.772**	**206.182**
Sharpe-Weighted Transition Tilt	0.767	0.762	0.758	0.750	0.741	0.724	47.548
ML GBM + Regime Interactions	0.685	0.675	0.664	0.642	0.620	0.557	-
Volatility Managed 10%	0.688	0.683	0.679	0.671	0.663	0.647	-
Equal Weight	0.589	0.588	0.587	0.585	0.583	0.579	-
Momentum Long-Short	**-0.096**	**-0.129**	**-0.161**	**-0.227**	**-0.293**	**-0.424**	-

Notes: This table reports net Sharpe ratios for each portfolio strategy under alternative transaction cost assumptions, expressed in basis points (bps) per unit of turnover. Sharpe ratios are computed using excess returns net of transaction costs. The baseline transaction cost assumption is 10 bps. The S&P 500 benchmark Sharpe ratio is 0.726 and assumes zero turnover. The break-even transaction cost is defined as the cost level at which a strategy's Sharpe ratio equals that of the benchmark. A dash indicates that the strategy does not outperform the benchmark at any tested transaction cost level. The highest Sharpe ratio for each transaction cost level is highlighted in bold.

5.3.3 Turnover Analysis

Table 5.10 decomposes the relationship between portfolio turnover, transaction costs, and outperformance across all strategies.

Table 5.10 Turnover & Transaction Cost Decomposition

Strategy	Turnover	TC	Gross SR	Net SR	Gross Return	Net Return
Regime-Aware Momentum Tilt	123.999	0.124	**0.810**	**0.801**	**11.616**	**11.492**
Momentum Tilt	121.624	0.122	0.796	0.788	11.445	11.323
Inverse Volatility Transition Tilt	42.706	0.043	0.787	0.784	11.115	11.072
Sharpe-Weighted Transition Tilt	125.448	0.125	0.767	0.758	11.107	10.981
Volatility Managed 10%	93.943	0.094	0.688	0.679	8.507	8.477
ML GBM + Regime Interactions	324.637	0.325	0.685	0.664	10.311	9.986
Equal Weight	**29.487**	**0.029**	0.589	0.588	7.919	7.825
Momentum Long-Short	**963.636**	**0.964**	**-0.096**	**-0.161**	**-1.414**	**-2.378**

Notes: This table reports portfolio turnover and the impact of transaction costs on Sharpe ratios and returns for each strategy. Turnover is measured as annual portfolio turnover. Transaction costs (TC) are computed assuming 10 basis points per unit of turnover and are expressed as a percentage of portfolio value. Gross and net Sharpe ratios and returns are reported before and after transaction costs, respectively. All returns are annualized excess returns expressed in percentage terms. The S&P 500 benchmark is excluded because it assumes zero turnover. The highest and lowest values within each column are highlighted in bold.

The outperforming strategies cluster in the 42–125% annual turnover range, losing less than 0.01 Sharpe ratio units to transaction costs. The ML GBM strategy's 325% turnover and Momentum Long-Short's 964% turnover explain much of their underperformance. Notably, Momentum Long-Short has a negative *gross* Sharpe ratio (-0.096), indicating fundamental failure beyond transaction cost drag.

6. Discussion

6.1 Summary of Findings

This study developed and evaluated regime-aware tactical sector allocation frameworks combining unsupervised machine learning for regime identification with rule-based portfolio construction. The empirical analysis addresses two research questions.

RQ1: Can machine learning-based regime detection identify distinct market states with significantly different sector return distributions?

Yes. Rolling PCA on eight FRED-MD categories combined with Mahalanobis distance-based crisis detection and K-Means clustering identifies four economically interpretable regimes. ANOVA tests confirm statistically significant differences across regimes for all eight macroeconomic factors ($p < 0.05$). Ten of eleven GICS sectors exhibit statistically significant return variation across regimes.

RQ2: Do regime-aware tactical sector allocation strategies outperform static benchmarks in terms of risk-adjusted returns?

Yes. During the test period (2014–2024), Regime-Aware Momentum Tilt achieves a Sharpe ratio of 0.801 compared to the benchmark's 0.726, with statistical significance at the 5% level under both Jobson-Korkie ($p = 0.035$) and Ledoit-Wolf bootstrap ($p = 0.026$) tests. Inverse Volatility Transition Tilt similarly achieves significance ($p = 0.042$).

Importantly, the documented performance improvements do not arise from repeated re-optimization or regime relabeling. All regime definitions, clustering choices, and strategy parameters are fixed prior to the test period and applied once. The deterioration of performance for more flexible ML strategies and larger tilts further supports the interpretation that results are not driven by overfitting but by economically meaningful regime information.

6.2 Explaining Strategy Performance

The results reveal a clear performance hierarchy that warrants explanation.

Why Cross-Sectional Momentum Long-Short Failed

Despite momentum being a well-documented anomaly, the momentum long-short strategy generates a deeply negative Sharpe ratio (−0.161). Two factors mainly explain this failure.

The first is implementation cost. The strategy generates 964% annual turnover, incurring transaction costs of approximately 0.96% per year, a figure nearly equal to the benchmark's entire excess return. Because sector rankings shift frequently, the strategy completely reconstitutes both long and short portfolios each month, creating substantial implementation drag.

The second relates to the short leg of the long–short portfolio. As shown in Appendix I, the ultimate failure of the cross-sectional momentum strategy is driven by persistently strong negative returns on the short side. These losses more than offset the gains generated by the long leg, effectively neutralizing the intended momentum premium. Consequently, this leads to the overall negative risk-adjusted performance of the strategy.

Why the Machine Learning Strategy Underperformed

The gradient boosting model achieves a Sharpe ratio of 0.664, falling short of both the benchmark (0.726) and simpler regime-based strategies. This underperformance reflects the fundamental tension between model flexibility and estimation error.

With 36 features and 162 hyperparameter combinations evaluated per sector, the model possesses sufficient capacity to fit noise rather than signal. The monthly retraining cycle compounds this problem by continuously adapting to recent patterns that may not persist, thereby generating trades based on spurious relationships. The resulting 325% annual turnover, nearly three times that of the Regime-Aware Momentum Tilt, erodes returns through transaction costs.

Taken together, these results suggest that additional nonlinear flexibility does not compensate for increased estimation error and turnover once regime information has already been incorporated through simpler means.

Why Regime-Aware Momentum Works

The Regime-Aware Momentum Tilt achieves the highest Sharpe ratio (0.801) by combining momentum's trend-following benefits with anticipatory de-risking. The key mechanism operates through the regime transition matrix.

As Table A.2 reveals, Crisis never follows Expansion directly; there is always an intervening Slowdown or Stress period. This transition structure provides early warning. When the economy moves from Expansion to Slowdown, transition probabilities shift: the likelihood of Stress and Crisis rises from near-zero to meaningful levels. The strategy exploits this pattern by preserving trend-following exposure during favorable conditions while systematically reducing it when macroeconomic fragility increases. Crucially, the strategy does not attempt to predict which sectors will outperform during crises. It simply reduces overall momentum exposure, thereby limiting losses during periods when momentum historically crashes.

Why Inverse Volatility Transition Works

The Inverse Volatility Transition Tilt achieves strong risk-adjusted returns (Sharpe ratio of 0.784) with remarkably low turnover, just 42.7% annually compared to 124% for Regime-Aware Momentum. Its success derives from two sources.

First, low-volatility sector rankings exhibit high persistence. Consumer Staples, Healthcare, and Utilities consistently rank among the lowest-volatility sectors month after month, while Technology, Energy, and Financials consistently rank among the highest. Because these rankings rarely change, the strategy trades infrequently and does not need to reconstitute positions each month.

Second, regime conditioning smooths transitions by probability-weighting expected volatilities across potential future states. Rather than reacting abruptly when regimes shift, the strategy gradually adjusts as transition probabilities evolve, reducing the turnover that typically disturbs strategies based on discrete regime classifications.

The combination of a reliable underlying anomaly, low volatility, with stable signal rankings produces the highest break-even transaction cost among all strategies: 206 basis points, more than 20 times the baseline cost assumption.

6.3 Transaction Costs Are Decisive

Transaction costs separate viable strategies from theoretical curiosities. The cross-sectional momentum long-short strategy generates deeply negative returns after costs. Even at zero transaction costs, its gross Sharpe ratio is -0.096, indicating fundamental signal failure at the sector level. The additional 0.96% annual cost drag transforms underperformance into substantial losses.

By contrast, the outperforming regime strategies exhibit break-even transaction costs of 48–206 basis points, well above typical institutional trading costs for sector ETFs. The Inverse Volatility Transition Tilt's exceptional cost resilience (206 bps break-even) derives from its persistent signal rankings and low turnover. This suggests that regime-aware strategies should prioritize signals with stable cross-sectional rankings to minimize unnecessary trading.

6.4 Factor Exposure Interpretation

The Fama-French-Carhart four-factor analysis (Appendix F) reveals that regime-aware strategies do not generate statistically significant alpha after controlling for market, size, value, and momentum factors. However, the absence of pure alpha does not diminish the strategies' value. Rather, the analysis reveals that regime-aware tactical strategies behave similarly to smart-beta funds, systematically capturing exposures across multiple risk factors beyond simple market beta.

Regime-Aware Momentum Tilt exhibits significant loadings on both the market factor (0.979, $p < 0.01$) and momentum factor (0.099, $p < 0.01$), while also showing meaningful exposure to the size factor (-0.133, $p < 0.01$) and a modest value tilt (0.035, $p < 0.05$). The regime-aware strategy effectively rotates across factor exposures as macroeconomic conditions evolve, increasing momentum exposure during favorable regimes while the negative size loading reflects a tilt toward larger, more stable companies.

The Inverse Volatility Transition Tilt displays a different factor profile, with lower market beta (0.947) and no significant momentum or value exposure, consistent with its focus on defensive, low-volatility sectors. This diversity in factor loadings across regime strategies suggests that the framework enables practitioners to harvest multiple risk premia through a single, economically motivated allocation rule.

For practitioners, this interpretation reframes the value proposition. Regime-aware strategies do not require belief in new return predictors or market inefficiencies. Instead, they offer a systematic approach to multi-factor investing that dynamically adjusts exposures based on macroeconomic conditions. They capture factor premia more efficiently than static factor allocations while also providing improved drawdown protection during regime transitions.

6.5 Contributions to Literature

This research extends the regime-aware allocation literature in several dimensions.

Sector-Level Application. While prior work focuses on asset class allocation (Ang and Bekaert, 2004; Oliveira et al., 2025), this study demonstrates that regime conditioning tactical allocations improves risk-adjusted returns at sector granularity.

Methodology Comparison. Systematic evaluation of 12 dimensionality reduction configurations and 175 clustering specifications establishes that rolling PCA on economically grouped categories outperforms alternatives for crisis detection.

Rigorous Statistical Assessment. This research provides validation using multiple testing methodologies. The consistency of significance across Jobson-Korkie and Ledoit-Wolf tests strengthens confidence in documented outperformance.

Comparison with Alternative Active Strategies. The finding that traditional momentum fails while regime-conditioned momentum succeeds provides actionable insight for strategy design.

6.6 Practical Implications

The findings offer actionable insights for institutional investors.

Strategy Selection. Regime-Aware Momentum Tilt offers the highest risk-adjusted returns with statistical significance and robust performance across transaction cost scenarios. Inverse Volatility Transition Tilt offers exceptional cost resilience, making it attractive for larger portfolios where market impact is a concern. Both strategies maintain long-only constraints and bounded deviation limits, ensuring tractable implementation.

Implementation Considerations. Monthly rebalancing captures regime dynamics effectively. Annual turnover for outperforming strategies generates transaction cost drag of only 0.04–0.13 percentage points at baseline assumptions, confirming practical viability for institutional implementation.

From a practitioner perspective, the results suggest that modest, rule-based tilts conditioned on macroeconomic regimes can meaningfully improve downside risk without materially increasing turnover or tracking error, making regime-aware allocation more suitable as a portfolio overlay than as a standalone strategy.

6.7 Limitations

Sample Constraints. The test period (2014–2024) contains only 3 crisis months and 3 stress months, limiting statistical power for regime-specific inference. The framework's crisis-period benefits rest on sparse observations dominated by a single episode (COVID-19).

Regime Classification Uncertainty. Silhouette scores of 0.180 indicate moderate cluster separation. The framework's value depends on regimes being "distinct enough" rather than perfectly separable, though ANOVA validation (all $p < 0.05$) confirms the identified regimes capture genuine heterogeneity.

Temporal Stability. Regime dynamics may shift with structural changes in monetary policy (quantitative easing, zero lower bound), market microstructure (passive investing growth), or economic structure. The expanding-window approach partially accommodates drift but cannot anticipate fundamental breaks.

Geographic Scope. Analysis restricts to GICS sectors within the S&P 500. European and emerging markets may exhibit different regime drivers not captured in FRED-MD data. International generalizability remains untested.

6.8 Future Research

Several extensions merit investigation.

Signal Enhancement. The success of Regime-Aware Momentum Tilt in scaling momentum exposure based on transition probabilities suggests that further refinement of regime-conditional signals may enhance performance. Incorporating additional predictive signals such as sentiment indicators, cross-asset momentum, or credit spreads as regime-conditional inputs represents a promising extension.

Alternative Regime Detection. While K-Means outperformed Hidden Markov Models in our setting, this finding may reflect estimation challenges with the two step implementation rather than fundamental limitations of HMMs. Probabilistic regime weights (Kritzman et al., 2023) rather than discrete classifications could reduce turnover from threshold-crossing artifacts.

Transaction Cost Dynamics. The uniform 10 basis point assumption abstracts from time-varying spreads that widen during stress periods. Future research could incorporate regime-conditional transaction cost models to assess whether the documented outperformance persists.

International Application. Extending the framework to European sectors using ECB-equivalent macroeconomic databases, or to emerging markets with commodity-cycle overlays, would test cross-market generalizability.

7. Conclusion

This thesis develops and evaluates a regime-aware tactical sector allocation framework that combines unsupervised machine learning for macroeconomic regime identification with systematic portfolio construction. The research demonstrates that conditioning tactical sector allocation decisions on macroeconomic states generates statistically significant improvements in risk-adjusted returns.

The regime identification framework employs rolling PCA on 126 macroeconomic indicators, Mahalanobis distance-based crisis detection, and K-Means clustering to identify four economically interpretable regimes: Expansion, Slowdown, Stress, and Crisis. These regimes are not merely statistical artifacts. They exhibit statistically significant differences across all eight macroeconomic factor categories and generate meaningfully different sector return distributions, with ten of eleven GICS sectors showing significant regime effects. The alignment of crisis classifications with major economic episodes validates the framework's economic relevance.

Over the 2014–2024 test period, regime-aware strategies deliver consistent outperformance. Regime-Aware Momentum Tilt achieves a Sharpe ratio of 0.801 compared to the S&P 500 benchmark's 0.726, representing approximately 74 basis points of annual outperformance. This outperformance achieves statistical significance at the 5% level under both Jobson-Korkie ($p = 0.035$) and Ledoit-Wolf bootstrap ($p = 0.026$) tests. The Inverse Volatility Transition Tilt similarly achieves significance ($p = 0.042$) while demonstrating exceptional robustness to transaction costs with a break-even of 206 basis points.

The regime transition matrix reveals that crises never follow expansions directly; there is always an intervening slowdown or stress period. This structure provides early warning, enabling strategies to reduce exposure before momentum crashes materialize. The Regime-Aware Momentum Tilt exploits this pattern through anticipatory de-risking, while the Inverse Volatility Transition Tilt combines a reliable low-volatility anomaly with persistent sector rankings that minimize turnover.

Traditional active strategies fail to match this performance. Cross-sectional momentum generates deeply negative returns with extreme turnover costs, while the machine learning approach underperforms due to overfitting and excessive trading. These findings establish that regime conditioning captures information beyond established return predictors, and that simpler economically structured signals outperform complex model-driven approaches in this setting.

The practical implications are actionable. Regime-aware strategies maintain moderate turnover (42–125% annually), generate robust performance against transaction cost drag, and exhibit break-even costs well above typical institutional trading expenses. The framework requires no proprietary data or exotic instruments; it can be implemented using publicly available macroeconomic indicators and liquid sector ETFs.

Future research integrating additional predictive signals, employing probabilistic regime weights, and extending to international markets may further strengthen the case for regime-aware tactical sector allocation. The methodology developed here, combining rolling dimensionality reduction, systematic crisis detection, and objective clustering selection, provides a replicable framework for such extensions. The evidence supports a clear conclusion: macroeconomic regimes identified through unsupervised learning contain economically valuable information for tactical sector allocation, enabling investors to improve risk-adjusted returns while maintaining transparency and implementability.

Bibliography

Ang, A., & Bekaert, G. (2004). How regimes affect asset allocation. *Financial Analysts Journal*, 60(2), 86–99. https://doi.org/10.2469/faj.v60.n2.2612

Ang, A., & Timmermann, A. (2012). Regime changes and financial markets. *Annual Review of Financial Economics*, 4(1), 313–337. https://doi.org/10.1146/annurev-financial-110311-101808

Arthur, D., & Vassilvitskii, S. (2007). k-means++: The advantages of careful seeding. *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, 1027–1035.

Asness, C. S., Moskowitz, T. J., & Pedersen, L. H. (2013). Value and Momentum Everywhere. *The Journal of Finance*, *68*(3), 929–985. http://www.jstor.org/stable/42002613

Bailey, D. H., Borwein, J. M., López de Prado, M., & Zhu, Q. J. (2014). Pseudomathematics and financial charlatanism: The effects of backtest overfitting on out-of-sample performance. *Notices of the American Mathematical Society*, 61(5), 458–471.

Barroso, P. & Santa-Clara, P., (2015). Momentum Has Its Moments. *Journal of Financial Economics* 116(1), 111-120, https://doi.org/10.1016/j.jfineco.2014.11.010

Barroso, P., Saxena, K., & Wang, H. (2023). Calm your portfolio: The importance of disciplining intelligent but fickle forecasts in portfolio optimization. *Working paper, Católica Lisbon School of Business and Economics, ESCP Business School, and National University of Singapore.* https://ssrn.com/abstract=5046369

Bellman, R. (1961). Adaptive Control Processes: A Guided Tour. *Princeton University Press.* https://www.jstor.org/stable/j.ctt183ph6v

Bernanke, B. S., & Gertler, M. (1995). Inside the black box: The credit channel of monetary policy transmission. *Journal of Economic Perspectives*, 9(4), 27–48. https://doi.org/10.1257/JEP.9.4.27

Burns, A. F., & Mitchell, W. C. (1946). *Measuring Business Cycles*. National Bureau of Economic Research. https://www.nber.org/books-and-chapters/measuring-business-cycles

Campbell, J. Y., & Cochrane, J. H. (1999). By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy*, 107(2), 205–251. https://ssrn.com/abstract=226593

Campbell, J. Y., Lo, A. W., & MacKinlay, A. C. (2008). *The Econometrics of Financial Markets*., 2(4), 559-562, https://doi.org/10.1017/S1365100598009092

Cederburg, S., O'Doherty, M. S., Wang, F., & Yan, X. S. (2020). On the performance of volatility-managed portfolios. *Journal of Financial Economics*, 138(1), 95–117. https://ssrn.com/abstract=3357038

Chevalier, G., Coqueret, G., & Raffinot, T. (2022). Supervised portfolios. *Quantitative Finance*, 22(12), 2275–2295. https://doi.org/10.1080/14697688.2022.2122543

Clarida, R., Galí, J., & Gertler, M. (2000). Monetary policy rules and macroeconomic stability: Evidence and some theory. *Quarterly Journal of Economics*, 115(1), 147–180. https://www.jstor.org/stable/2586937

Dahlquist, M., & Harvey, C. R. (2001). Global tactical asset allocation. *Journal of Global Capital Markets*. https://doi.org/10.2139/ssrn.795376

Daniel, K., & Moskowitz, T. J. (2016). Momentum crashes. *Journal of Financial Economics*, 122(1), 221–247. https://doi.org/10.1016/j.jfineco.2015.12.002

De Longis, A., & Ellis, D. (2023). Tactical asset allocation, risk premia, and the business cycle: a macro regime approach. *The Journal of Portfolio Management*, 49(4), 103–126. https://doi.org/10.3905/jpm.2022.1.456

Fama, E. F., & French, K. R. (1989). Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics*, 25(1), 23–49. https://doi.org/10.1016/0304-405X(89)90095-0

Ferson, W. E., & Harvey, C. R. (1991). The variation of economic risk premiums. *Journal of Political Economy*, 99(2), 385–415. https://ssrn.com/abstract=3281992

Giglio, S., Kelly, B., & Xiu, D. (2022). Factor models, machine learning, and asset pricing. *Annual Review of Financial Economics*, 14(1), 337–368. https://doi.org/10.1146/annurev-financial-101521-104735

Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. *Review of Financial Studies*, 33(5), 2223–2273. https://doi.org/10.1093/rfs/hhaa009

Guidolin, M., & Timmermann, A. (2007). Asset allocation under multivariate regime switching. *Journal of Economic Dynamics and Control*, 31(11), 3503–3544. https://doi.org/10.1016/j.jedc.2006.12.004

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2), 357–384. https://www.jstor.org/stable/1912559

Harvey, C. R., Liu, Y., & Zhu, H. (2016). ... and the cross-section of expected returns. *Review of Financial Studies*, 29(1), 5–68. https://doi.org/10.1093/rfs/hhv059

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd ed.). Springer. https://link.springer.com/book/10.1007/978-0-387-84858-7

Hirsa, A., Xu, S., & Malhotra, S. (2024). Robust Rolling Regime Detection (R2-RD): A Data-Driven Perspective of Financial Markets. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4729435

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance*, 48(1), 65–91. https://doi.org/10.1111/j.1540-6261.1993.tb04702.x

Jobson, J. D., & Korkie, B. M. (1981). Performance hypothesis testing with the Sharpe and Treynor measures. *Journal of Finance*, 36(4), 889–908. https://doi.org/10.1111/j.1540-6261.1981.tb04891.x

Kiyotaki, N., & Moore, J. (1997). Credit cycles. *Journal of Political Economy*, 105(2), 211–248. https://www.jstor.org/stable/10.1086/262072

Kritzman, M., Kulasekaran, C., & Turkington, D. (2023). Portfolio Construction When Regimes Are Ambiguous. *The Journal of Portfolio Management*, 50(1), 8–18. https://doi.org/10.3905/jpm.2023.1.538

Kritzman, M., Page, S., & Turkington, D. (2012). Regime shifts: Implications for dynamic strategies. *Financial Analysts Journal*, 68(3), 22–39. https://doi.org/10.2469/faj.v68.n3.3

Ledoit, O., & Wolf, M. (2008). Robust performance hypothesis testing with the Sharpe ratio. *Journal of Empirical Finance*, 15(5), 850–859. https://doi.org/10.1016/j.jempfin.2008.03.002.

McCracken, M. W., & Ng, S. (2016). FRED-MD: A monthly database for macroeconomic research. *Journal of Business & Economic Statistics*, 34(4), 574–589. https://www.jstor.org/stable/44166591

McInnes, L., Healy, J., & Melville, J. (2018). UMAP: Uniform Manifold Approximation and Projection for dimension reduction. *The Journal of Open Source Software, 3(29), 861,* https://joss.theoj.org/papers/10.21105/joss.00861

Memmel, C. (2003). Performance Hypothesis Testing with the Sharpe Ratio. *Finance Letters, 1, 21-23.* https://ssrn.com/abstract=412588

Moreira, A., & Muir, T. (2017). Volatility-managed portfolios. *Journal of Finance*, 72(4), 1611–1644. https://doi.org/10.1111/jofi.12513

Mueller-Glissmann, C., & Ferrario, A. (2024). Dynamic asset allocation using machine learning: Seeing the forest for the trees. *The Journal of Portfolio Management*, 50(5), 132–157. https://doi.org/10.3905/jpm.2024.1.582

Oliveira, D. C., Sandfelder, D., Fujita, A., Dong, X., & Cucuringu, M. (2025). Tactical Asset Allocation with Macroeconomic Regime Detection. https://ssrn.com/abstract=5183762

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20, 53–65. https://doi.org/10.1016/0377-0427(87)90125-7

Shu, Y., & Mulvey, J. M. (2024). Dynamic factor allocation leveraging Regime-Switching signals. *The Journal of Portfolio Management*, jpm.2024.1.649. https://doi.org/10.3905/jpm.2024.1.649

Snow, D. (2020). Machine Learning in Asset Management-Part 2: Portfolio Construction-Weight Optimization. *The Journal of Financial Data Science*, 2(2), 17–24. https://doi.org/10.3905/jfds.2020.1.029

Stock, J. H., & Watson, M. W. (1999). Business cycle fluctuations in US macroeconomic time series. *Handbook of Macroeconomics*, 1, 3–64. https://doi.org/10.1016/S1574-0048(99)01004-6

Appendices

Appendix A: Regime Dynamics and Transition Analysis

Table A.1 Regime Duration Statistics, Full Sample

Regime	Mean	Median	Min	Max	Episodes
Expansion	4.140	2.000	1	34	43
Slowdown	2.511	1.000	1	13	45
Stress	1.000	1.000	1	1	11
Crisis	2.000	2.000	1	3	5

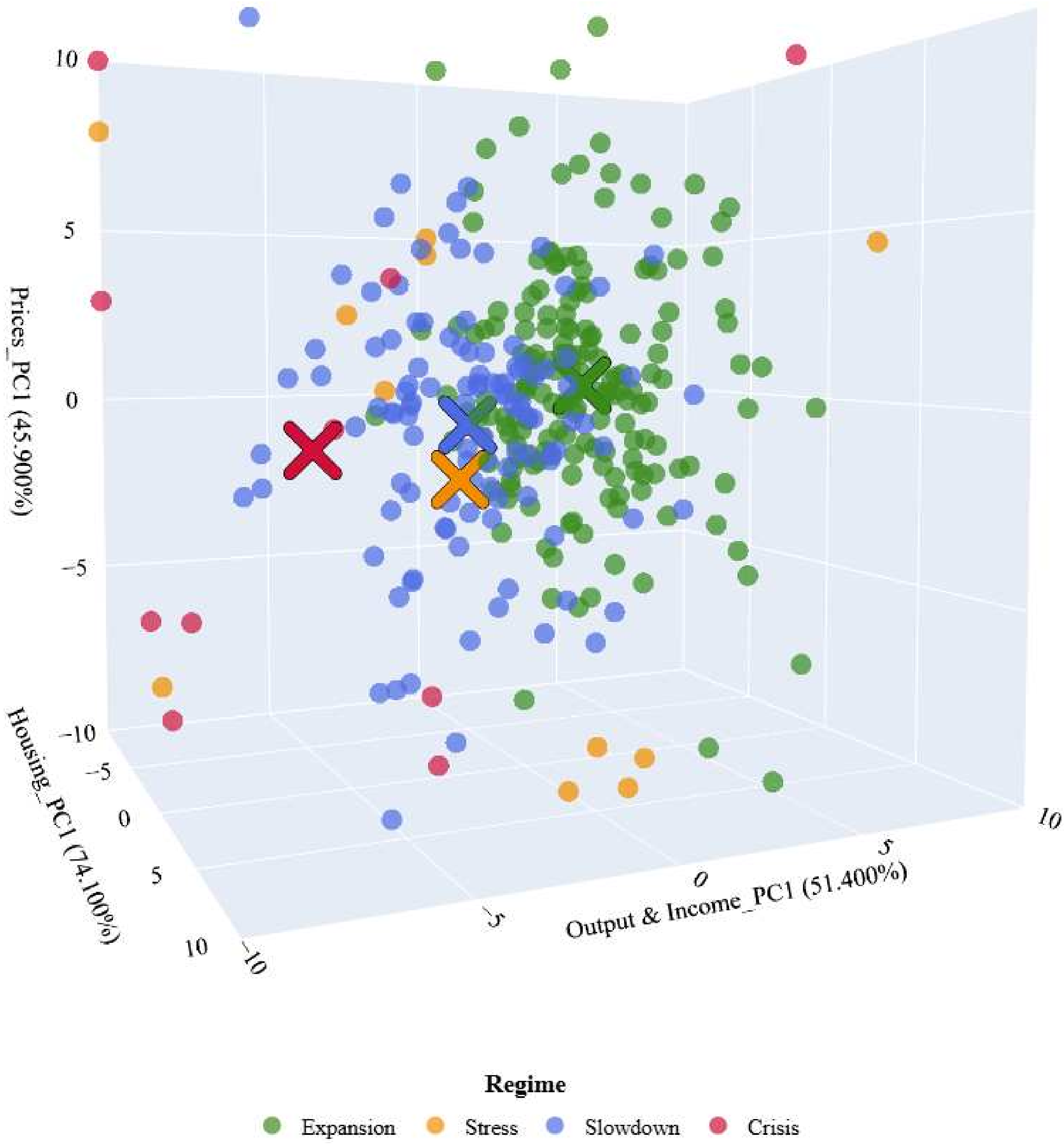
Notes: This table reports descriptive statistics for regime durations measured in months over the full sample period. Reported statistics include the mean, median, minimum, maximum, and total number of regime episodes for each identified regime.

Table A.2 Regime Transition Probability Matrix, Full Sample

From/To	Expansion	Slowdown	Stress	Crisis
Expansion	0.763	0.215	0.023	0.000
Slowdown	0.345	0.602	0.027	0.027
Stress	0.273	0.545	0.000	0.182
Crisis	0.000	0.100	0.400	0.500

Notes: This table reports the estimated one-month transition probability matrix for the identified regimes over the full sample period. Each row corresponds to the regime in the previous month, and each column corresponds to the regime in the subsequent month. Diagonal elements represent regime persistence probabilities, while off-diagonal elements capture transitions across regimes.

Figure A.1 Three-Dimensional Regime Clustering Plot



Notes: This figure presents a three-dimensional visualization of the regime clustering results. Monthly observations are projected onto the first principal components of the Output and Income, Housing, and Prices macroeconomic categories which has the highest explained variances. The percentages in axis labels indicate the proportion of variance explained by each principal component. Cluster centroids are marked with an "X." To improve visualization, values are capped at ±10. The figure illustrates the relative dispersion of observations across regimes in the reduced-dimensional space.

Appendix B: Macroeconomic Characterization (Training & Validation)

Table B.1 ANOVA Results for Regime Distinctiveness, Training and Validation

Factor	F-statistic	p-value	η^2 (Effect Size)	Significant
Output & Income	29.408	<0.001	0.334	***
Labor Market	47.224	<0.001	**0.446**	***
Housing	13.654	<0.001	0.189	***
Consumption & Orders	22.893	<0.001	0.281	***
Money & Credit	8.151	<0.001	0.122	***
Interest & FX	10.077	<0.001	0.147	***
Prices	1.965	0.121	**0.032**	
Stock Market	11.037	<0.001	0.158	***

*Notes: This table reports analysis of variance (ANOVA) results testing whether macroeconomic factor realizations differ across identified regimes using the training and validation samples. The F-statistic tests the null hypothesis of equal means across regimes. The effect size, η^2, measures the proportion of total variance explained by regime membership. Statistical significance is denoted by ***, **, and * corresponding to the 1%, 5%, and 10% levels, respectively. The largest and smallest effect sizes are highlighted in bold.*

Table B.2 Macroeconomic Deviations from Overall Average by Regime, Training and Validation

Indicator	Overall Average	Expansion	Slowdown	Stress	Crisis
Industrial Growth	1.127	+1.840	-0.651	-6.340	-7.649
Unemployment	6.241	+0.159	-0.137	-0.166	-0.183
Inflation	2.417	-0.031	+0.068	-0.550	+0.231
VIX	22.423	-3.980	+1.182	+11.811	+21.315
10Y Treasury	4.025	-0.148	+0.162	+0.154	-0.226

Notes: This table reports deviations of selected economic indicators from their training and validation sample averages across identified regimes. Reported values represent differences relative to the overall sample mean computed using the training and validation observations. Positive (negative) values indicate realizations above (below) the overall average for a given indicator and regime.

Table B.3 Macroeconomic Deviations from Overall Average by Regime, Full Sample

Indicator	Overall Average	Expansion	Slowdown	Stress	Crisis
Industrial Growth	0.809	+1.217	-0.628	-4.781	-9.206
Unemployment	5.639	-0.072	-0.092	+0.434	+1.841
Inflation	2.564	+0.086	-0.127	+0.375	-0.510
VIX	20.488	-2.822	+1.268	+11.702	+23.037
10Y Treasury	3.374	-0.275	+0.454	+0.245	-0.495

Notes: This table reports deviations of selected economic indicators from their full-sample averages across identified regimes. Reported values represent differences relative to the overall full-sample mean for each indicator. Positive (negative) values indicate realizations above (below) the overall average within each regime.

Appendix C: Complete Methodology Configurations and Results

Table C.1 Dimensionality Reduction Configurations and Properties

Method	Data Input	Window	Configurations	Linear	Aggregated	Static
Rolling PCA	8 Categories	60, 90, 120, 150 months	4	Yes	Yes	No
Rolling PCA	126 Variables	60, 90, 120, 150 months	4	Yes	No	No
Static PCA	8 Categories/126 Variables	Full Training Period	2	Yes	Yes/No	Yes
UMAP	8 Categories/126 Variables	Full Training Period	2	No	Yes/No	Yes

Notes: This table summarizes the dimensionality reduction configurations evaluated in the study. Configurations vary by method, data input, rolling versus static implementation, aggregation level, and linearity. Rolling specifications differ by window length, while static specifications are estimated over the full training period. These configurations are used to assess the robustness of regime identification to alternative dimensionality reduction approaches.

Table C.2 Variance Explained per Dimensionality Reduction Configuration

Method	Data Type	Window	Variance Explained	Interpretation
Rolling PCA	8 Categories	60 months	45.983%	Avg. PC1 per Category
Rolling PCA	8 Categories	90 months	46.201%	Avg. PC1 per Category
Rolling PCA	8 Categories	120 months	46.032%	Avg. PC1 per Category
Rolling PCA	8 Categories	150 months	46.142%	Avg. PC1 per Category
Rolling PCA	126 Variables	60 months	58.300%	Cumulative Variance
Rolling PCA	126 Variables	90 months	56.791%	Cumulative Variance
Rolling PCA	126 Variables	120 months	54.844%	Cumulative Variance
Rolling PCA	126 Variables	150 months	52.811%	Cumulative Variance
Static PCA	8 Categories	Static	45.590%	Avg. PC1 per Category
Static PCA	126 Variables	Static	52.363%	Cumulative Variance
UMAP	8 Categories	Static	N/A	Manifold Method
UMAP	126 Variables	Static	N/A	Manifold Method

Notes: This table reports the variance explained by each dimensionality reduction configuration. For category-level principal component analysis, reported values correspond to the average variance explained by the first principal component within each category. For full-data principal component analysis, reported values correspond to cumulative variance explained by the retained principal components. UMAP is a nonlinear manifold learning method and does not produce variance-explained measures; therefore, variance explained is not applicable for these configurations.

Table C.3 Quality Metrics per Dimensionality Reduction Configuration

Method	Data Type	Window	Stability	Variance	Stationarity (%)	Outliers (%)
Rolling PCA	8 Categories	60 months	0.421	11.473	50.000	0.260
Rolling PCA	8 Categories	90 months	0.412	12.088	50.000	0.238
Rolling PCA	8 Categories	120 months	0.397	14.276	50.000	0.278
Rolling PCA	8 Categories	150 months	0.384	16.810	37.500	0.417
Rolling PCA	126 Variables	60 months	0.248	11.212	**87.500**	0.208
Rolling PCA	126 Variables	90 months	0.223	11.139	50.000	0.298
Rolling PCA	126 Variables	120 months	0.225	12.538	75.000	0.556
Rolling PCA	126 Variables	150 months	**0.206**	13.963	75.000	**0.750**
Static PCA	8 Categories	Static	0.425	6.376	75.000	0.708
Static PCA	126 Variables	Static	0.392	7.682	25.000	0.458
UMAP	8 Categories	Static	0.317	**40.632**	75.000	**0.000**
UMAP	126 Variables	Static	**0.679**	**0.469**	**0.000**	**0.000**

Notes: This table reports quality metrics evaluated on the training sample. Stability measures the first-order autocorrelation of changes in principal component scores, with higher values indicating smoother component evolution. Variance reports the average variance of the extracted principal component scores. Stationarity reports the percentage of components that reject a unit root based on augmented Dickey–Fuller tests at the 5% level. Outliers report the percentage of observations exceeding four standard deviations. The highest and lowest values within each metric are highlighted in bold.

Table C.4 Crisis Detection Performance by Dimensionality Reduction Configuration

Rank	Method	Data Type	Window	F1-Score	Precision	Recall
1	Rolling PCA	8 Categories	150 months	**0.667**	**1.000**	**0.500**
2	Rolling PCA	8 Categories	120 months	0.606	0.909	0.455
3	Rolling PCA	126 Variables	150 months	0.593	0.889	0.444
4	Rolling PCA	8 Categories	60 months	0.485	0.727	0.364
5	Rolling PCA	8 Categories	90 months	0.485	0.727	0.364
6	Rolling PCA	126 Variables	120 months	0.485	0.727	0.364
7	Static PCA	8 Categories	-	0.485	0.727	0.364
8	Static PCA	126 Variables	-	0.485	0.727	0.364
9	UMAP	126 Variables	-	0.424	0.636	0.318
10	Rolling PCA	126 Variables	60 months	0.424	0.636	0.318
11	Rolling PCA	126 Variables	90 months	0.364	0.545	0.273
12	UMAP	8 Categories	-	**0.303**	**0.455**	**0.227**

Notes: This table reports crisis detection performance for each dimensionality reduction configuration. F1-scores, precision, and recall are computed by comparing predicted stress and crisis periods to National Bureau of Economic Research recession months. The evaluation sample includes recession episodes associated with the dot-com recession and the onset of the global financial crisis. The highest and lowest values within each performance metric are highlighted in bold.

Table C.5 Clustering Performance by Configuration (Top 10)

Rank	Method	Data Type	Window	Clustering Method	k	Silhouette
1	**Rolling PCA**	**8 Categories**	**120 months**	**K-Means**	**2**	**0.180**
2	Rolling PCA	8 Categories	120 months	K-Means	3	0.178
3	Rolling PCA	8 Categories	150 months	K-Means	6	0.175
4	Rolling PCA	8 Categories	60 months	K-Means	2	0.174
5	Rolling PCA	126 Variables	150 months	K-Means	6	0.172
6	Rolling PCA	126 Variables	150 months	K-Means	5	0.171
7	Rolling PCA	8 Categories	120 months	Spectral	2	0.171
8	Rolling PCA	126 Variables	150 months	GMM	5	0.170
9	Rolling PCA	8 Categories	90 months	K-Means	2	0.168
10	Rolling PCA	8 Categories	60 months	Spectral	2	0.168

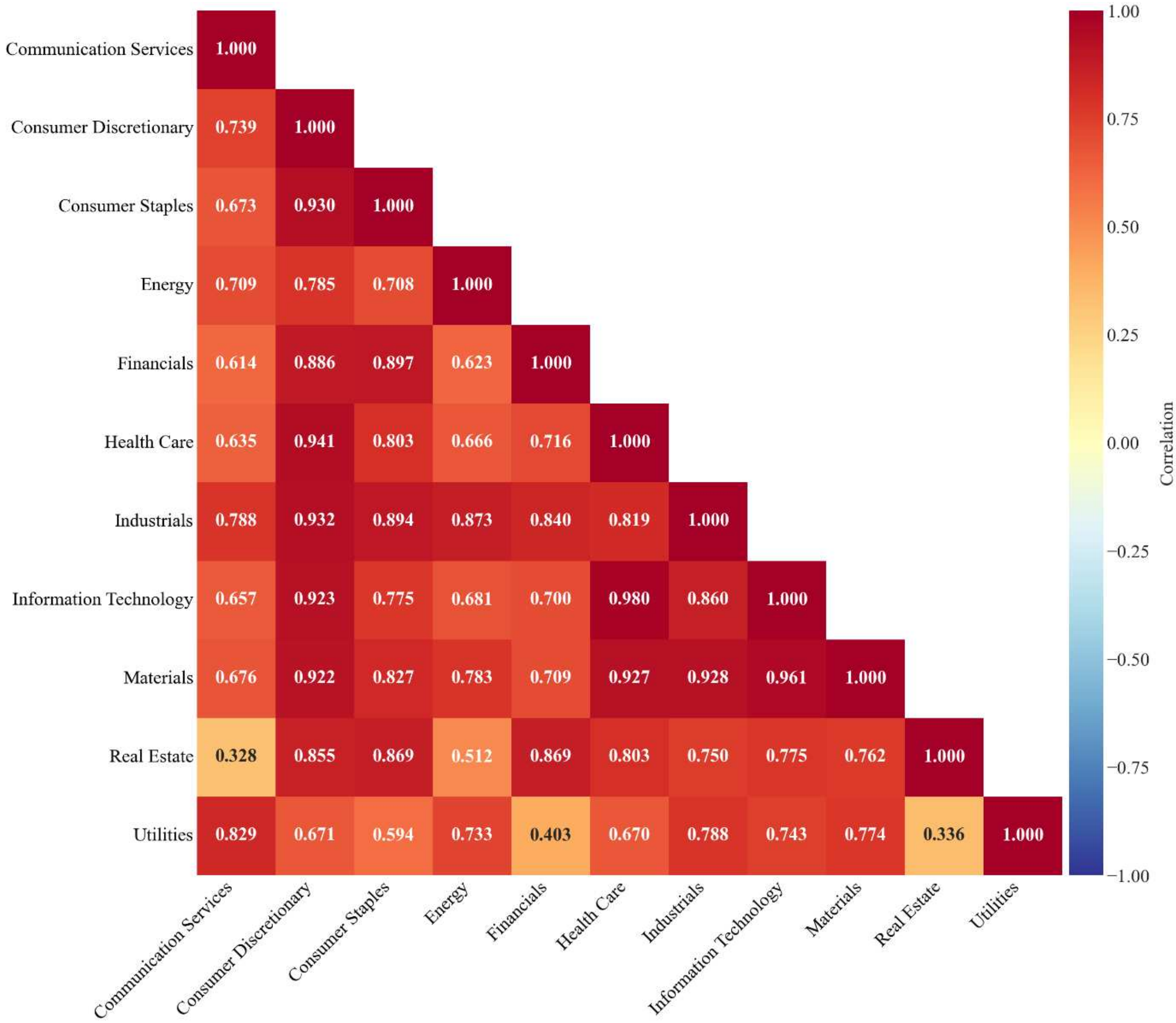
Notes: This table reports the ten highest-performing clustering configurations ranked by silhouette score. Silhouette scores are computed using training sample observations classified as normal periods, excluding stress and crisis regimes. Higher silhouette values indicate better-defined cluster separation. All configurations are estimated using training data through December 2008. The configuration highlighted in bold corresponds to the selected clustering specification used in the main analysis.

Table C.6 Clustering Method Performance Summary

Clustering Method	Max Silhouette	Mean Silhouette
K-Means	**0.180**	**0.160**
Spectral Clustering	0.171	0.146
Gaussian Mixture	0.170	0.139
Hierarchical (Average)	0.167	0.138
Hierarchical (Ward)	0.164	0.136
Hierarchical (Complete)	0.160	0.110
Hidden Markov Model	**0.153**	**0.047**

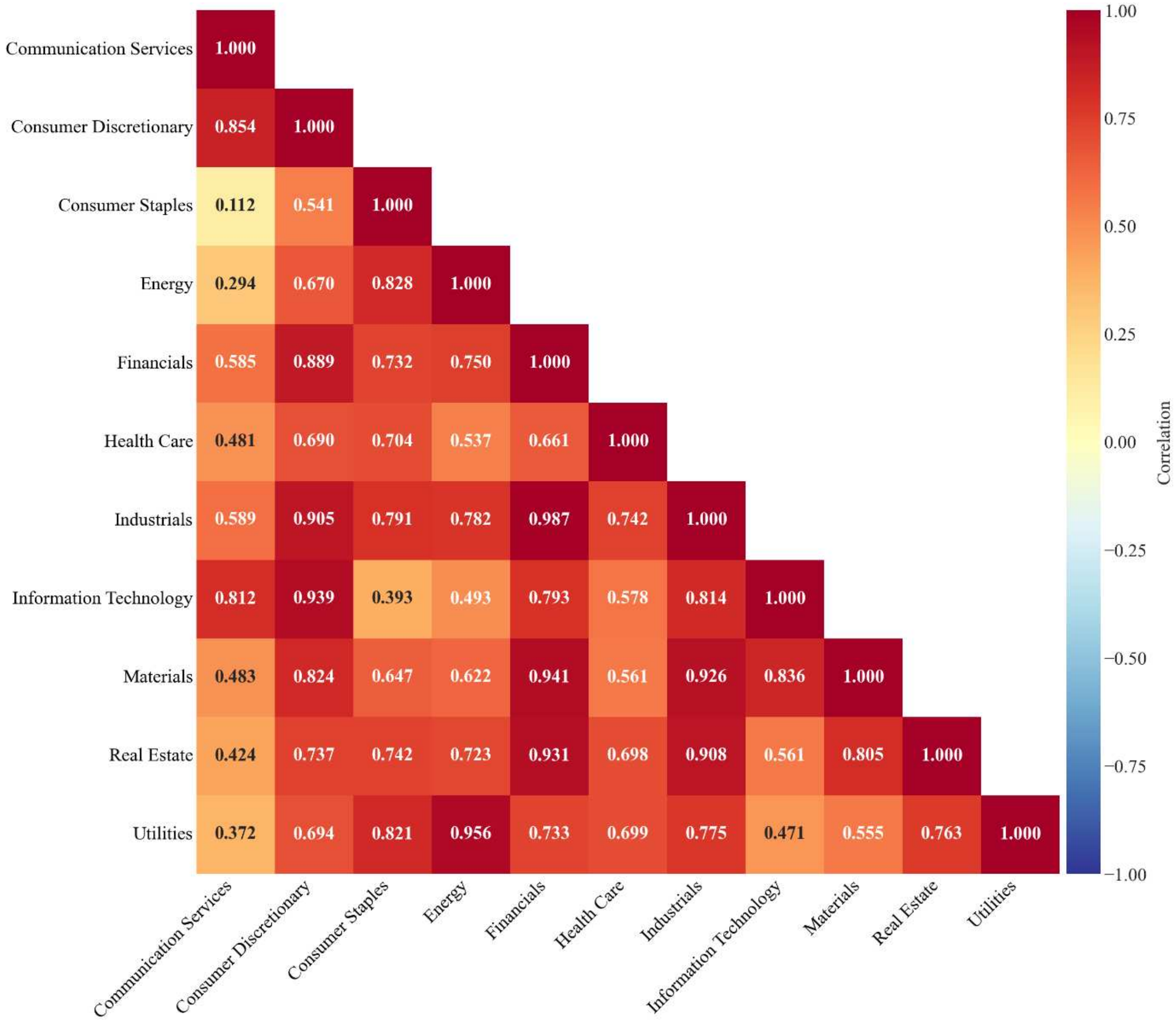
Notes: This table reports summary statistics of clustering performance across all tested configurations grouped by clustering method. Reported statistics include the maximum and mean silhouette scores achieved by each method. Silhouette scores are computed using training sample observations classified as normal periods. The highest and lowest values within each column are highlighted in bold.

Figure C.1 Sector Return Correlation During Crisis Periods



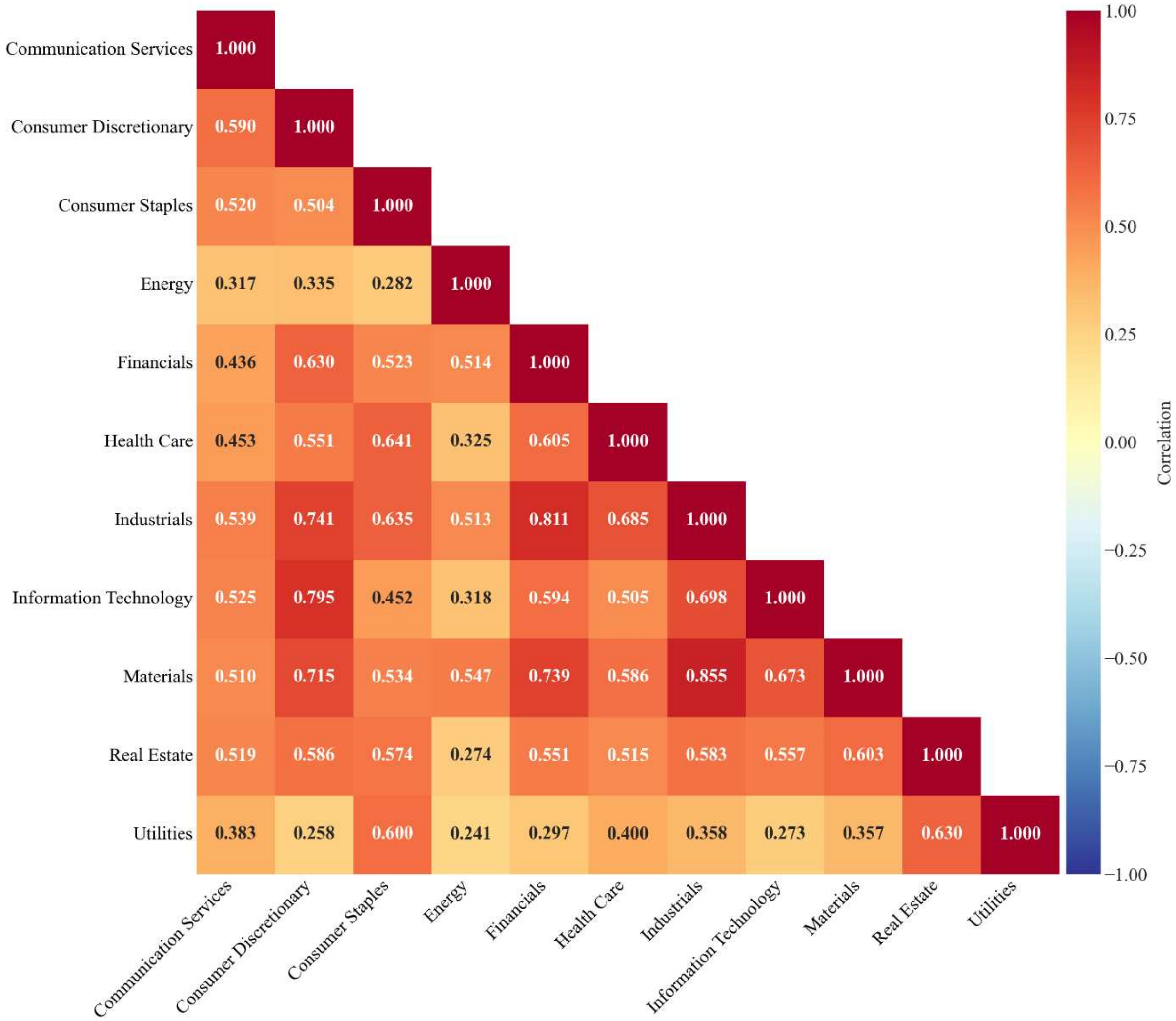
Notes: This figure reports pairwise correlations of monthly sector excess returns during identified crisis periods. Correlations are computed using observations classified as crisis regimes only. The color scale reflects the magnitude of correlation coefficients, with darker shades indicating stronger co-movement. Values on the diagonal equal one by construction.

Figure C.2 Sector Return Correlation During Stress Periods



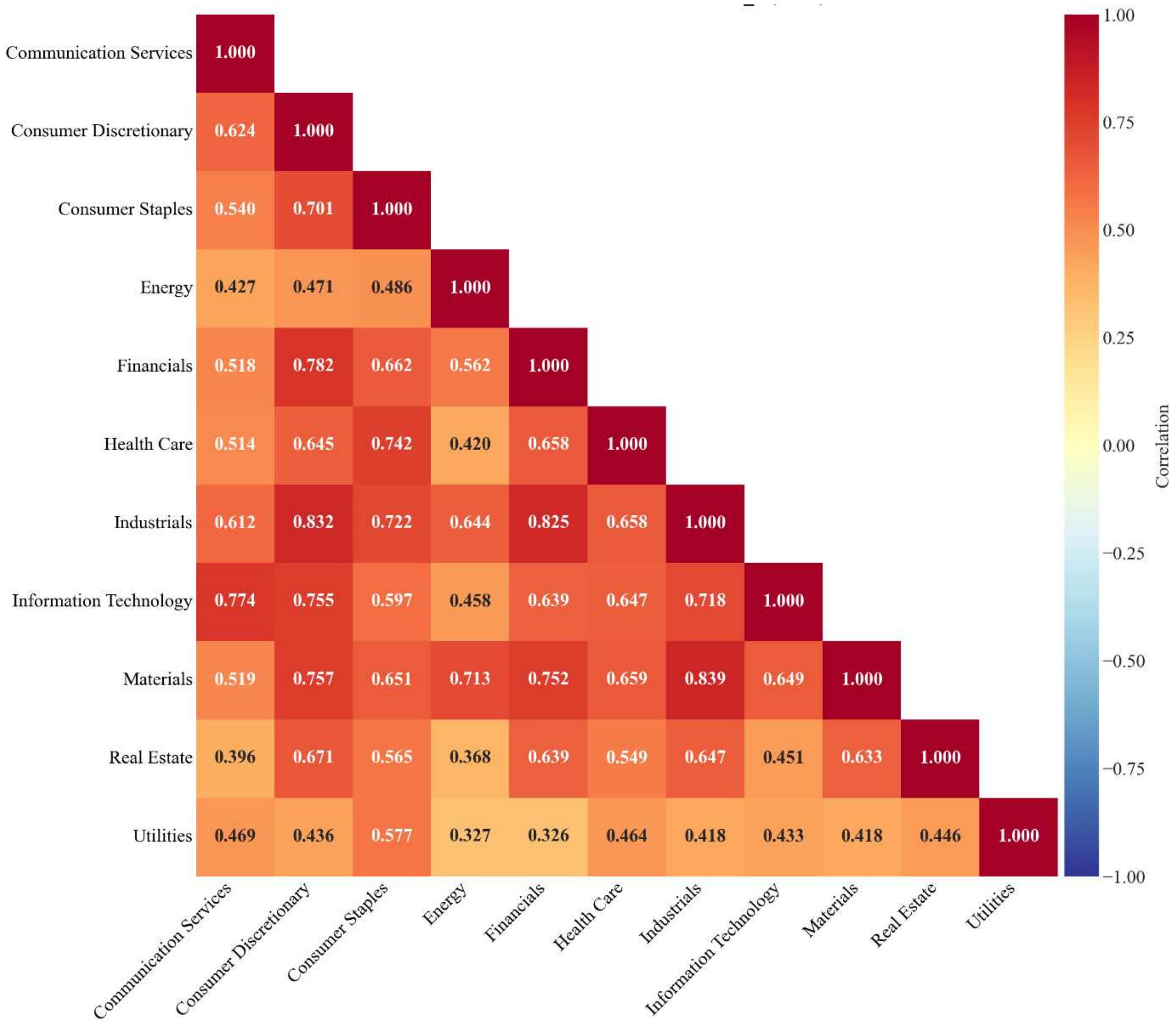
Notes: This figure reports pairwise correlations of monthly sector excess returns during identified stress periods. Correlations are computed using observations classified as stress regimes only. The color scale reflects the magnitude of correlation coefficients, with darker shades indicating stronger co-movement. Values on the diagonal equal one by construction.

Figure C.3 Sector Return Correlation During Expansion Periods



Notes: This figure reports pairwise correlations of monthly sector excess returns during identified expansion periods. Correlations are computed using observations classified as expansion regimes only. The color scale reflects the magnitude of correlation coefficients, with darker shades indicating stronger co-movement. Values on the diagonal equal one by construction.

Figure C.4 Sector Return Correlation During Slowdown Periods



Notes: This figure reports pairwise correlations of monthly sector excess returns during identified slowdown periods. Correlations are computed using observations classified as slowdown regimes only. The color scale reflects the magnitude of correlation coefficients, with darker shades indicating stronger co-movement. Values on the diagonal equal one by construction.

Appendix D: Strategy Results (Training & Validation)

Table D.1 Training Period Performance Summary

Strategy	Ann. Return	Ann. Vol	Sharpe	Sortino	Max DD	Calmar	TC
Sharpe-Weighted Transition Tilt	**-0.289**	14.307	**-0.020**	**-0.016**	-38.549	**-0.007**	0.072
Volatility Managed 10%	-0.704	14.869	-0.048	-0.040	-39.531	-0.018	0.209
Volatility Managed 8%	-0.726	**13.275**	-0.055	-0.047	-32.873	-0.022	0.080
Volatility Managed 12%	-1.143	16.165	-0.071	-0.058	-45.003	-0.025	0.293
ML GBM + Regime Interactions	-1.037	14.479	-0.072	-0.058	-41.217	-0.025	0.302
Volatility Managed 14%	-1.816	17.242	-0.106	-0.085	-49.124	-0.037	0.331
Equal Weight	-1.822	15.041	-0.122	-0.091	-43.193	-0.042	0.032
Volatility Managed 16%	-2.239	**18.082**	-0.124	-0.098	**-52.292**	-0.043	0.346
Regime-Aware Momentum Tilt	-2.159	14.561	-0.149	-0.117	-39.697	-0.054	0.136
Momentum Tilt	-2.243	14.729	-0.153	-0.122	-40.598	-0.055	0.133
Inverse Volatility Transition Tilt	-2.426	14.015	-0.174	-0.135	-40.203	-0.060	0.043
S&P 500	-1.997	14.925	-0.202	-0.161	-42.770	-0.070	**0.000**
Momentum Long-Short	**-3.039**	14.877	**-0.204**	**-0.202**	**-30.737**	**-0.099**	**0.966**

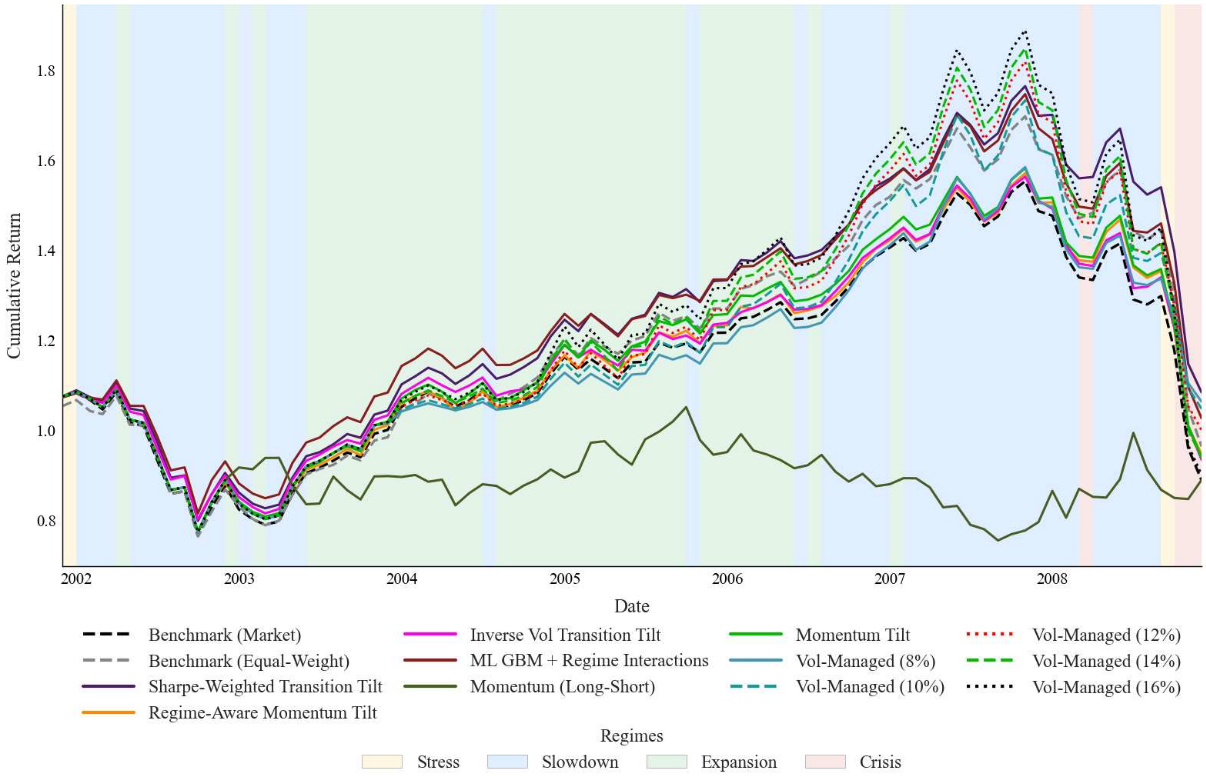
Notes: This table reports training period performance metrics for all portfolio strategies in terms of excess returns. All returns are reported net of transaction costs (TC). Annualized return, volatility, Sharpe ratio, Sortino ratio, maximum drawdown (DD), Calmar ratio, and transaction costs are expressed in percentage terms, except for risk-adjusted ratios. Transaction costs are computed assuming 10 basis points per unit of turnover and are expressed as a percentage of portfolio value. The training period spans 86 months from November 2001 to December 2008. The highest and lowest values within each column are highlighted in bold.

Table D.2 Sharpe Ratios by Regime, Training Period

Strategy	Overall	Expansion	Slowdown	Stress	Crisis
Sharpe-Weighted Transition Tilt	**-0.020**	**0.477**	-0.125	0.910	-1.278
Volatility Managed 10%	-0.048	0.457	-0.130	0.968	-1.533
Volatility Managed 8%	-0.055	0.454	-0.149	0.908	**-1.579**
Volatility Managed 12%	-0.071	0.449	-0.128	1.033	-1.534
ML GBM + Regime Interactions	-0.072	0.447	**-0.115**	1.007	-1.523
Volatility Managed 14%	-0.106	0.440	-0.128	1.105	-1.507
Equal Weight	-0.122	0.471	-0.141	1.148	-1.486
Volatility Managed 16%	-0.124	0.445	-0.127	**1.183**	-1.449
Regime-Aware Momentum Tilt	-0.149	0.407	-0.158	0.926	-1.314
Momentum Tilt	-0.153	0.428	-0.169	0.901	-1.378
Inverse Volatility Transition Tilt	-0.174	0.422	-0.163	1.047	-1.482
S&P 500	-0.202	0.424	-0.175	1.014	-1.521
Momentum Long-Short	**-0.204**	**0.045**	**-0.210**	**0.142**	**0.554**

Notes: This table reports training period performance metrics for all portfolio strategies in terms of excess returns. All returns are reported net of transaction costs. Annualized return, volatility, Sharpe ratio, Sortino ratio, maximum drawdown, Calmar ratio, and transaction costs are expressed in percentage terms, except for risk-adjusted ratios. Transaction costs are computed assuming 10 basis points per unit of turnover and are expressed as a percentage of portfolio value. The training period spans 86 months from November 2001 to December 2008. The highest and lowest values within each column are highlighted in bold.

Figure D.1 Cumulative Returns During the Training Period



Notes: This figure plots cumulative excess returns for selected portfolio strategies over the training period from November 2001 to December 2008. The market and equal-weight benchmarks are included for comparison. Background shading indicates the identified macroeconomic regimes, including expansion, slowdown, stress, and crisis periods.

Table D.3 Validation Period Performance Summary

Strategy	Ann. Return	Ann. Vol	Sharpe	Sortino	Max DD	Calmar	TC
Inverse Volatility Transition Tilt	**16.674**	15.240	**1.094**	**1.010**	-15.200	**1.097**	0.080
Volatility Managed 10%	10.003	**9.200**	1.087	0.976	**-11.405**	0.877	0.026
Regime-Aware Momentum Tilt	16.295	15.447	1.055	0.929	-17.023	0.957	0.133
Sharpe-Weighted Transition Tilt	16.158	15.577	1.037	0.978	-16.921	0.955	0.156
Momentum Tilt	16.157	15.577	1.037	0.917	-17.570	0.920	0.130
S&P 500	16.194	15.918	1.017	0.923	-16.666	0.972	**0.000**
ML GBM + Regime Interactions	16.023	15.817	1.013	0.947	-17.621	0.909	0.301
Equal Weight	15.601	15.979	0.976	0.904	-15.610	0.999	0.028
Momentum Long-Short	**-1.774**	**16.946**	**-0.105**	**-0.102**	**-39.852**	**-0.045**	**1.017**

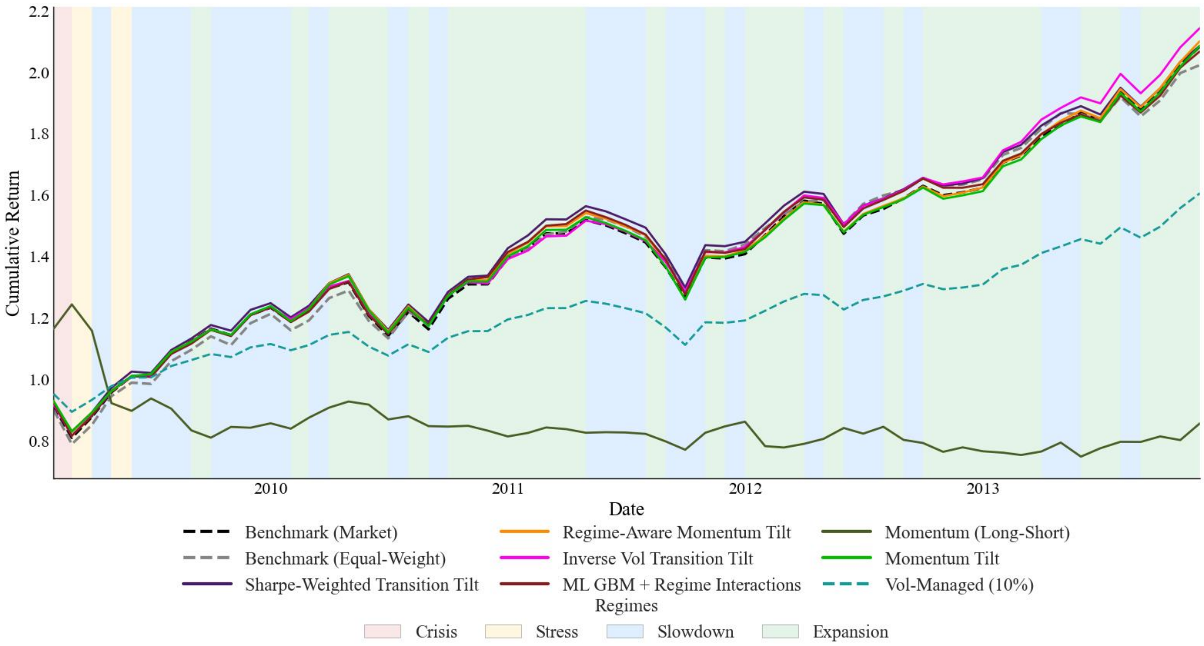
Notes: This table reports validation period performance metrics for all portfolio strategies in terms of excess returns. All returns are reported net of transaction costs (TC). Annualized return, volatility, Sharpe ratio, Sortino ratio, maximum drawdown (DD), Calmar ratio, and transaction costs are expressed in percentage terms, except for risk-adjusted ratios. Transaction costs are computed assuming 10 basis points per unit of turnover and are expressed as a percentage of portfolio value. The validation period spans 60 months from January 2009 to December 2013. The highest and lowest values within each column are highlighted in bold.

Table D.4 Sharpe Ratios by Regime, Validation Period

Strategy	Overall	Expansion	Slowdown	Stress	Crisis
Inverse Volatility Transition Tilt	**1.094**	0.561	**0.279**	-0.080	N/A
Volatility Managed 10%	1.087	**0.589**	0.222	-0.087	N/A
Regime-Aware Momentum Tilt	1.055	0.557	0.229	-0.111	N/A
Sharpe-Weighted Transition Tilt	1.037	0.551	0.210	-0.050	N/A
Momentum Tilt	1.037	0.542	0.224	-0.109	N/A
S&P 500	1.017	0.552	0.232	-0.074	N/A
ML GBM + Regime Interactions	1.013	0.545	0.221	-0.079	N/A
Equal Weight	0.976	0.520	0.274	**-0.043**	N/A
Momentum Long-Short	**-0.105**	**-0.023**	**-0.109**	**-0.352**	N/A

Notes: This table reports Sharpe ratios for each portfolio strategy computed separately by identified regime during the validation period. Regime-specific Sharpe ratios are calculated using monthly excess returns within expansion, slowdown, stress, and crisis regimes. The validation period contains 33 expansion months, 24 slowdown months, 2 stress months, and 1 crisis month. The overall Sharpe ratio is reported for comparison. Entries marked as N/A indicate that the Sharpe ratio cannot be computed due to insufficient observations. The highest and lowest Sharpe ratios within each regime are highlighted in bold.

Figure D.2 Cumulative Returns During the Validation Period



Notes: This figure plots cumulative excess returns for selected portfolio strategies over the validation period from January 2009 to December 2013. The market and equal-weight benchmarks are included for comparison. Background shading indicates the identified macroeconomic regimes, including expansion, slowdown, stress, and crisis periods.

Appendix E: Tactical Sector Tilt Analysis

Table E.1 Average Monthly Sector Tilt per Regime, Regime-Aware Momentum Tilt

Sector	Overall	Expansion	Slowdown	Stress	Crisis
Communication Services	+0.082	-0.264	+1.149	-2.453	+1.581
Consumer Discretionary	+0.108	0.120	+0.255	-1.778	-0.001
Consumer Staples	-0.656	**-0.858**	-0.534	+1.264	+2.353
Energy	**-1.029**	-0.650	**-2.301**	**+2.554**	-2.386
Financials	-0.236	-0.018	-0.412	-2.053	-3.255
Health Care	-0.018	-0.006	-0.389	+2.172	+1.467
Industrials	-0.006	+0.262	-0.169	**-3.043**	**-3.488**
Information Technology	**+1.979**	**+1.836**	**+2.250**	+1.385	**+4.016**
Materials	-0.260	+0.099	-1.078	-0.689	-1.991
Real Estate	-0.105	-0.388	+0.593	+0.993	-0.109
Utilities	0.143	-0.135	+0.637	+1.659	+1.812

Notes: This table reports average monthly sector weight tilts for the regime-aware momentum tilt strategy over the test period from January 2014 to December 2024. Sector tilts are expressed in percentage points relative to market-capitalization benchmark weights. Positive (negative) values indicate sector overweights (underweights). Reported values are averages computed within each identified regime.

Table E.2 Average Monthly Sector Tilt per Regime, Inverse Volatility Transition Tilt

Sector	Overall	Expansion	Slowdown	Stress	Crisis
Communication Services	-1.765	-1.823	-1.918	-0.932	+0.884
Consumer Discretionary	+1.364	+1.436	+1.324	+0.674	+0.279
Consumer Staples	**+4.767**	**+4.765**	**+4.801**	**+4.675**	**+4.557**
Energy	-3.826	-3.934	-3.699	**-3.305**	**-2.388**
Financials	-1.027	-0.663	-1.830	-1.245	-3.273
Health Care	+3.665	+3.683	+3.751	+3.175	+2.669
Industrials	+1.382	+1.490	+1.373	+0.442	-0.923
Information Technology	**-4.305**	**-4.335**	**-4.583**	-3.048	-1.580
Materials	-1.936	-2.101	-1.506	-2.110	-1.373
Real Estate	-0.970	-1.061	-0.610	-1.481	-1.626
Utilities	+2.650	+2.542	+2.897	+3.155	+2.773

Notes: This table reports average monthly sector weight tilts for the inverse volatility transition tilt strategy over the test period from January 2014 to December 2024. Sector tilts are expressed in percentage points relative to market-capitalization benchmark weights. Positive (negative) values indicate sector overweights (underweights). Reported values are averages computed within each identified regime.

Table E.3 Average Monthly Sector Tilt per Regime, Sharpe-Weighted Transition Tilt

Sector	Overall	Expansion	Slowdown	Stress	Crisis
Communication Services	**-2.822**	**-2.808**	**-2.992**	-2.806	-1.417
Consumer Discretionary	+2.100	+2.321	+1.226	**+3.662**	**+3.329**
Consumer Staples	**+3.350**	**+3.257**	**+3.754**	+2.864	+2.273
Energy	-2.220	-2.055	-2.689	**-2.822**	-1.592
Financials	-1.139	-0.525	-2.605	-0.910	**-4.270**
Health Care	+1.102	+0.817	+1.689	+2.713	+1.890
Industrials	+1.845	+2.501	+0.426	+1.022	-2.047
Information Technology	-1.685	-2.297	-0.477	+0.235	2.069
Materials	-1.190	-1.259	-1.027	-1.354	-0.667
Real Estate	+1.445	+1.442	+1.638	-0.285	+1.131
Utilities	-0.786	-1.394	+1.058	-2.320	-0.700

Notes: This table reports average monthly sector weight tilts for the Sharpe-weighted transition tilt strategy over the test period from January 2014 to December 2024. Sector tilts are expressed in percentage points relative to market-capitalization benchmark weights. Positive (negative) values indicate sector overweights (underweights). Reported values are averages computed within each identified regime.

Table E.4 Average Monthly Sector Tilt per Regime, ML GBM + Regime Interactions

Sector	Overall	Expansion	Slowdown	Stress	Crisis
Communication Services	**-1.105**	-0.834	**-2.345**	-0.057	+3.076
Consumer Discretionary	+0.569	+0.468	+0.585	+2.918	+1.162
Consumer Staples	-0.348	-0.185	-0.478	**-2.676**	-1.629
Energy	+0.129	+0.142	+0.072	+0.978	-0.520
Financials	-0.994	**-0.902**	-1.204	+1.958	**-4.512**
Health Care	-0.529	-0.340	-0.801	-3.081	-0.823
Industrials	+0.175	+0.280	+0.312	-1.392	-2.997
Information Technology	-0.087	-0.706	+0.906	**+3.495**	**+4.582**
Materials	**+0.831**	**+0.924**	+0.842	-1.004	-0.356
Real Estate	+0.740	+0.602	**+1.431**	-1.010	-0.816
Utilities	+0.619	+0.550	+0.679	-0.127	+2.834

Notes: This table reports average monthly sector weight tilts for the ML GBM with regime interactions strategy over the test period from January 2014 to December 2024. Sector tilts are expressed in percentage points relative to market-capitalization benchmark weights. Positive (negative) values indicate sector overweights (underweights). Reported values are averages computed within each identified regime.

Table E.5 Average Monthly Sector Tilt per Regime, Momentum Tilt

Sector	Overall	Expansion	Slowdown	Stress	Crisis
Communication Services	+0.047	-0.325	+1.192	**-2.778**	+1.822
Consumer Discretionary	+0.102	+0.076	+0.508	-2.761	-0.708
Consumer Staples	-0.527	-0.801	-0.355	**+3.247**	+2.285
Energy	**-1.221**	**-0.959**	**-2.198**	+2.554	-2.389
Financials	-0.246	+0.060	-0.844	-1.386	-2.023
Health Care	+0.127	+0.156	+0.013	+1.505	-0.884
Industrials	-0.011	+0.294	-0.306	-3.043	**-3.201**
Information Technology	**+1.914**	**+1.827**	**+2.034**	+1.385	**+3.824**
Materials	+0.264	+0.128	-1.082	-1.671	-1.994
Real Estate	+0.103	-0.396	+0.566	+0.984	+0.519
Utilities	+0.183	-0.060	+0.472	+1.965	+2.749

Notes: This table reports average monthly sector weight tilts for the Momentum tilt strategy over the test period from January 2014 to December 2024. Sector tilts are expressed in percentage points relative to market-capitalization benchmark weights. Positive (negative) values indicate sector overweights (underweights). Reported values are averages computed within each identified regime.

Appendix F: Fama-French-Carhart 4-Factor Analysis

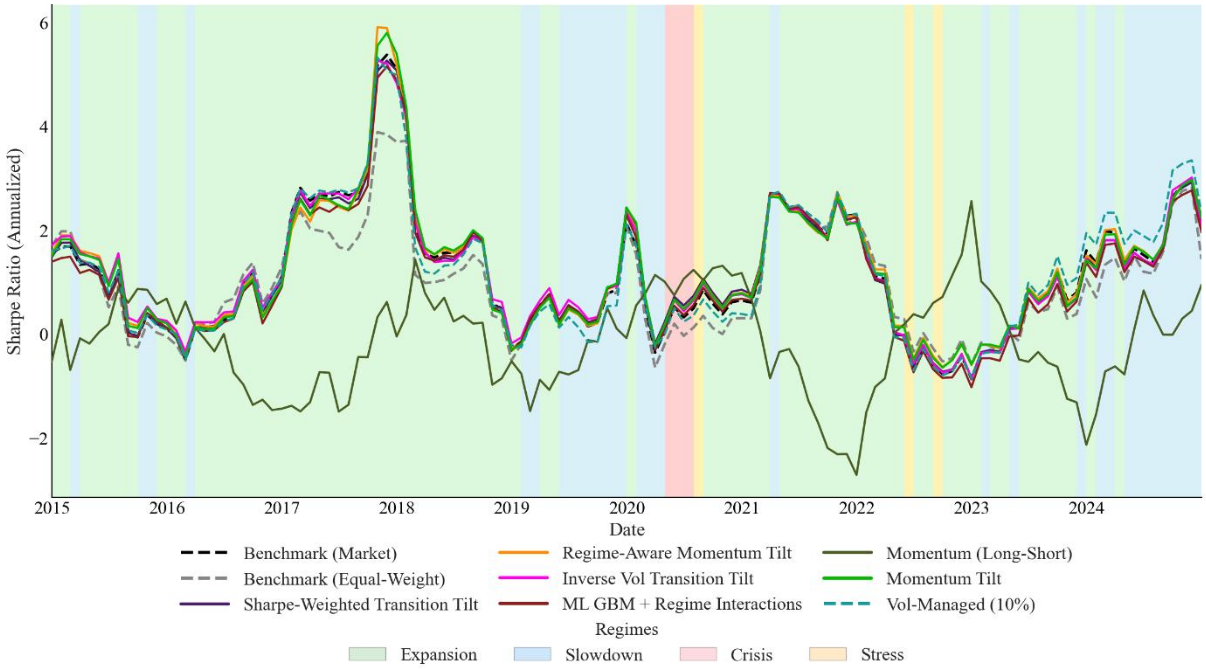
Table F.1 Fama-French-Carhart 4-Factor Regression Results

Strategy	α Annualized (%)	Mkt-Rf	SMB	HML	MOM	R^2
Regime-Aware Momentum Tilt	**-0.333**	0.979***	-0.133***	0.035**	0.099***	0.988
Inverse Volatility Transition Tilt	-0.346	0.947***	**-0.146*****	-0.007	0.033**	0.986
Momentum Tilt	-0.513	0.981***	-0.136***	0.041***	0.097***	0.989
Sharpe-Weighted Transition Tilt	-0.699*	0.971***	-0.137***	-0.005	0.034**	0.989
Volatility Managed 10%	-0.947	0.732***	-0.090**	0.010	0.046	0.879
S&P 500	-0.977**	0.984***	-0.142***	0.015	0.001	**0.995**
ML GBM + Regime Interactions	-1.913***	**0.997*****	-0.125***	**-0.008**	**-0.006**	0.988
Equal Weight	-2.172**	0.924***	-0.125***	**0.155*****	-0.012	0.956
Momentum Long-Short	**-3.377**	**-0.039**	**0.013**	0.129	**0.814*****	**0.534**

*Notes: This table reports Fama–French–Carhart four-factor regression results for portfolio excess returns over the test period from January 2014 to December 2024. The dependent variable is the strategy excess return. α denotes the annualized intercept. Mkt–Rf is the market excess return factor, SMB is the size factor (small minus big), HML is the value factor (high minus low), and MOM is the momentum factor. R^2 denotes the coefficient of determination. Statistical significance is denoted by ***, **, and * corresponding to the 1%, 5%, and 10% levels, respectively.*

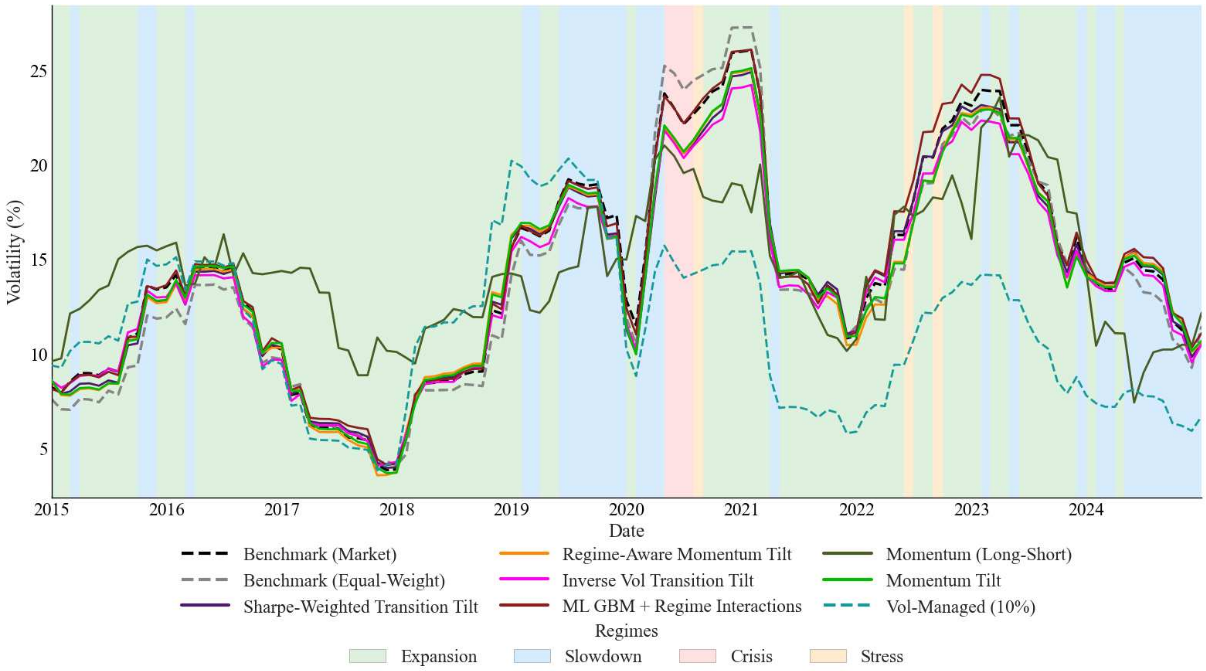
Appendix G: Rolling Volatility, Sharpe Ratio, & Monthly Return Dynamics

Figure G.1 Rolling 12-Month Sharpe Ratios



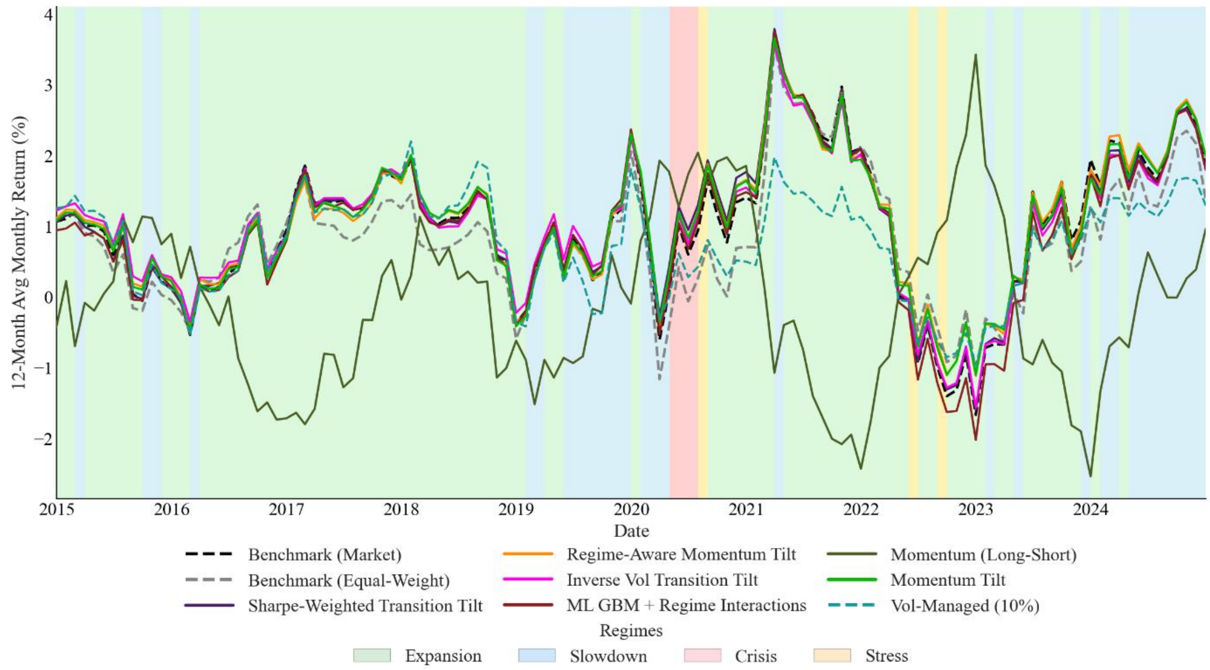
Notes: This figure plots rolling 12-month annualized Sharpe ratios for all portfolio strategies over the test period from January 2014 to December 2024. Sharpe ratios are computed using monthly excess returns. The market and equal-weight benchmarks are included for comparison. Background shading indicates the identified macroeconomic regimes, including expansion, slowdown, stress, and crisis periods.

Figure G.2 Rolling 12-Month Volatility



Notes: This figure plots rolling 12-month annualized volatility for all portfolio strategies over the test period from January 2014 to December 2024. Volatility is computed using monthly excess returns and expressed in percentage terms. The market and equal-weight benchmarks are included for comparison. Background shading indicates the identified macroeconomic regimes.

Figure G.3 Rolling 12-Month Average Monthly Excess Returns



Notes: This figure plots rolling 12-month average monthly excess returns for all portfolio strategies over the test period from January 2014 to December 2024. Volatility is computed using monthly excess returns and expressed in percentage terms. The market and equal-weight benchmarks are included for comparison. Background shading indicates the identified macroeconomic regimes.

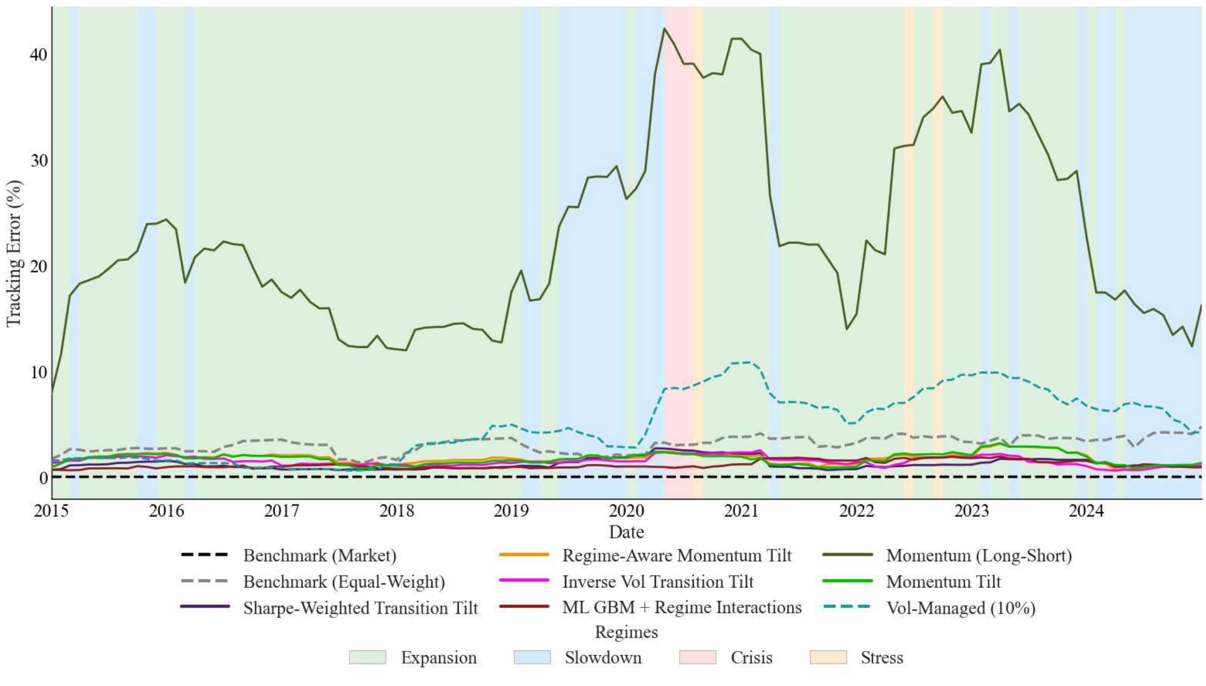
Appendix H: Tracking Errors and Information Ratio

Table H.1 Tracking Error and Information Ratio Relative to the S&P 500

Strategy	Tracking Error (%)	Information Ratio
Regime-Aware Momentum Tilt	1.755	**0.425**
Momentum Tilt	1.732	0.333
Inverse Volatility Transition Tilt	1.494	0.218
Sharpe-Weighted Transition Tilt	1.286	0.183
S&P 500	0.000	N/A
ML GBM + Regime Interactions	**1.180**	-0.644
Volatility Managed 10%	5.578	-0.525
Equal Weight	3.282	**-0.691**
Momentum Long-Short	**23.967**	-0.548

Notes: This table reports tracking error and information ratio for each strategy relative to the S&P 500 benchmark over the test period from January 2014 to December 2024. Tracking error is defined as the annualized standard deviation of the difference between strategy returns and benchmark returns. The information ratio is defined as the annualized average active excess return divided by tracking error. All statistics are computed using monthly excess returns. The S&P 500 benchmark has zero tracking error by construction and therefore no defined information ratio.

Figure H.1 Rolling Tracking Error Relative to the S&P 500



Notes: This figure plots rolling 12-month tracking error for each strategy relative to the S&P 500 benchmark over the test period from January 2014 to December 2024. Tracking error is computed as the annualized standard deviation of rolling 12-month return differentials. The market and equal-weight benchmarks are included for reference. Background shading indicates the identified macroeconomic regimes, including expansion, slowdown, stress, and crisis periods.

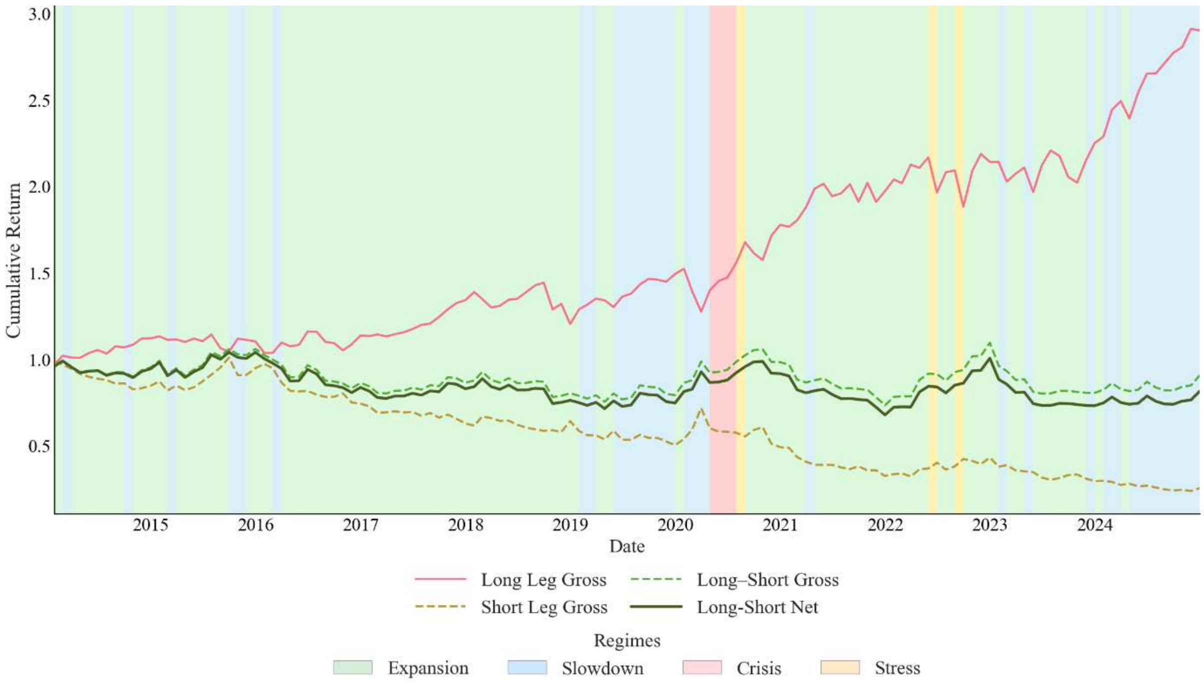
Appendix I: Long-Short Momentum Performance Decomposition

Table I.1 Performance Decomposition of Long and Short Momentum Legs

Strategy	Ann. Return	Ann. Vol	Sharpe	Sortino	Max DD	TC
Momentum Long Leg Gross	**9.123**	**14.198**	**0.643**	**0.601**	**-16.565**	-
Momentum Short Leg Gross	**-12.150**	**18.874**	**-0.644**	**-0.648**	**-76.319**	-
Momentum Long-Short Gross	-1.414	14.727	-0.096	-0.100	-30.753	-
Momentum Long-Short Net	-2.378	14.716	-0.161	-0.169	-34.722	**0.964**

Notes: This table reports test-period performance metrics for the long leg, short leg, and combined long–short momentum strategy over the period from January 2014 to December 2024. Returns are measured as excess returns. Gross returns are reported before transaction costs (TC), while net returns incorporate transaction costs. Reported metrics include annualized return, annualized volatility, Sharpe ratio, Sortino ratio, maximum drawdown (DD), and annualized transaction costs expressed as a percentage of portfolio value. Transaction costs are based on realized turnover. Risk-adjusted ratios are unitless. The highest and lowest values within each column are highlighted in bold.

Figure I.1 Cumulative Returns of Long and Short Momentum Legs



Notes: This figure plots cumulative returns of the long and short legs of the long–short momentum strategy over the test period from January 2014 to December 2024. The figure also reports gross and net cumulative returns of the combined long–short portfolio. Gross returns are computed before transaction costs, while net returns incorporate transaction costs. Background shading indicates the identified macroeconomic regimes, including expansion, slowdown, stress, and crisis periods.

Appendix J: Parameter Breakdown & Implementation Details

Table J.1 Crisis Detection Parameters

Strategy	Value	Description
Distance Metric	Mahalanobis	$d_M = \sqrt{(x-\mu)^T \Sigma^{-1} (x-\mu)}$
Stress Threshold	90th Percentile	Based on training period distances
Crisis Threshold	95th Percentile	Based on training period distances
Selection Metric	F1 Score	Stress + Crisis vs. NBER recessions
Covariance Estimation	Training Only	Avoid look-ahead bias

Notes: This table reports the parameters used in the crisis detection component of the two-stage regime identification framework. Crisis and stress thresholds are defined as percentiles of the Mahalanobis distance distribution estimated using training-period data only to avoid look-ahead bias. The Mahalanobis distance measures deviations from normal macroeconomic conditions using the inverse covariance matrix of indicators. Threshold values are selected to maximize the F1 score when classifying stress and crisis months relative to NBER recession dates.

Table J.2 Clustering Methods and Selection

Panel A Clustering Methods Evaluated

Method	Cluster Shape	Key Assumption	Parameters
K-Means	Spherical	Equal variance, convex clusters	k, n_init = 10, init = k-means++
Gaussian Mixture Model	Elliptical	Gaussian distributed clusters	k, Covariance = Full
Hierarchical (Ward)	Compact	Variance-minimizing merges	k, Linkage = Ward
Hierarchical (Complete)	Flexible	Maximum pairwise distances	k, Linkage = Complete
Hierarchical (Average)	Compact	Average pairwise distances	k, Linkage = Average
Spectral	Non-convex	Graph connectivity	k, Affinity = rbf
Hidden Markov Model	Sequential	Markov state transitions	k, n_iter = 100, init = random

Panel B Selection Parameters

Parameter	Value
K Values	{2, 3, 4, 5, 6}
Selection Criterion	Maximum Silhouette Score
Evaluation Period	60 Month Training Period (November 2001 – December 2008)
Configurations Tested	5 × 7 × 5 = 175
Random State	42

Notes: This table summarizes the clustering algorithms evaluated for regime identification and the criteria used for model selection. Panel A reports key assumptions and parameterizations for each clustering method. Panel B reports the grid of cluster counts and the evaluation criterion used for selection. All configurations are evaluated using training-period data only. The final clustering specification is selected based on the maximum silhouette score across all tested configurations.

Table J.3 Strategy Constraint Parameters

Constraint	Value	Applied to
Maximum Tilt	±5 percentage points	Tilt strategies
Minimum Weight	0.5%	All strategies
Momentum Lookback	12 months	Momentum strategies
Momentum Skip	1 month	Momentum strategies
Transaction Cost	10 bps (one-way)	All Strategies
Vol-Scale Maximum	1.5	Volatility-managed strategies
Vol-Scale Minimum	0.5	Volatility-managed strategies
Vol-Scaling Lookback	24 months	Volatility-managed strategies

Notes: This table reports the implementation constraints applied to all portfolio strategies. Maximum sector tilts are imposed to limit concentration risk while allowing meaningful active positions. Minimum sector weights prevent corner solutions. Momentum strategies use a 12-month lookback with a one-month skip. Volatility-managed strategies rescale exposure based on realized volatility subject to upper and lower bounds. Transaction costs are modeled as 10 basis points per one-way trade to ensure implementability.

Table J.4 ML GBM Hyperparameter Configuration

Panel A Hyperparameter Grid

Hyperparameter	Grid Values	Selection Model
n_estimators	{50, 100, 150}	TSCV MSE
max_depth	{2, 3, 4}	TSCV MSE
learning_rate	{0.03, 0.04, 0.05}	TSCV MSE
min_samples_leaf	{3, 5, 8}	TSCV MSE
subsample	{0.7, 0.8}	TSCV MSE

Panel B Training Configuration

Parameter	Value
Algorithm	GradientBoostingRegressor
Target Variable	Next-month sector excess return (z-score)
Features	8 PCA + 4 regime dummies + 4 transition probs + 5 momentum + 3 volatility + 2 ranks + 8 interactions + 2 expected returns (36 total)
Grid Search	$3 \times 3 \times 3 \times 3 \times 2 = 162$ combinations per sector
Cross-Validation	Expanding window TSCV
Retrain Frequency	Every 12 Months

Panel C Optimized Parameters by Sector

Sector	Learning Rate	Max Depth	Min Samples Leaf	N Estimators	Subsample
Communication Services	0.05	3	3	50	0.7
Consumer Discretionary	0.03	2	5	50	0.8
Consumer Staples	0.03	2	8	50	0.7
Energy	0.03	4	8	50	0.8
Financials	0.10	4	8	100	0.8
Health Care	0.03	2	3	50	0.7
Industrials	0.05	4	3	50	0.8
Information Technology	0.03	2	8	50	0.7
Materials	0.05	2	8	50	0.7
Real Estate	0.03	2	8	50	0.7
Utilities	0.03	2	8	50	0.7

Notes: This table reports the hyperparameter configuration for the ML GBM with regime interactions strategy. Panel A reports the hyperparameter grid searched for each sector. Panel B describes the training setup, including the target variable, feature set, cross-validation scheme, and retraining frequency. Panel C reports the optimized hyperparameters selected separately for each sector using time-series cross-validation with an expanding window. All models are trained using information available at the time of estimation.

Table J.5 Portfolio Backtest Configuration

Parameter	Setting
Backtest Type	Walk-Forward Expanding Window
Training Period	November 2001 – December 2008 (86 months)
Validation Period	January 2009 – December 2013 (60 months)
Test Period	January 2014 – December 2024 (132 months)
Rebalancing Frequency	Monthly (end-of-month)
Return Calculation	Excess Return

Notes: This table summarizes the backtesting framework used to evaluate portfolio strategies. The analysis employs a walk-forward expanding-window design with strict separation between training, validation, and test periods. Portfolios are rebalanced monthly using end-of-month data, and returns are computed as excess returns relative to the risk-free rate. All performance metrics reported in the main text refer to the test period unless otherwise noted.