

# Regime-Aware Asset Allocation: a Statistical Jump Model Approach

Yizhan Shu\*, Chenyu Yu\*, John M. Mulvey\*†

February 9, 2024

## Abstract

This article investigates the impact of regime switching on asset allocation decisions, with a primary focus on comparing different regime identification models. In contrast to traditional Markov-switching models, we adopt the *statistical jump model*, a recently proposed robust model known for its ability to capture persistent market regimes by applying an explicit jump penalty. The feature set of our jump model comprises return and volatility features derived solely from the price series. We introduce a data-driven approach for selecting the jump penalty within a time-series cross-validation framework, which directly optimizes the performance metric of the regime-aware asset allocation strategy constructed following a comprehensive multi-step process. Through empirical analysis using daily return series from major US equity indices, we highlight the outperformance of employing jump models in comparison to both buy-and-hold strategies and Markov-switching asset allocation approaches. These results underline the enhanced robustness, interpretability, and realism inherent in asset allocation strategies guided by jump models, offering insights for portfolio managers.

**Keywords:** Dynamic Asset Allocation; Statistical Jump Models; Regime Switching; Temporal Clustering

---

\*Department of Operations Research and Financial Engineering, Princeton University, Princeton, USA

†Yizhan Shu (corresponding author), [yizhans@princeton.edu](mailto:yizhans@princeton.edu), Chenyu Yu, [chenyu@princeton.edu](mailto:chenyu@princeton.edu), John M. Mulvey, [mulvey@princeton.edu](mailto:mulvey@princeton.edu)

# 1 Introduction

Throughout history, traded markets have exhibited changing patterns of performance, such as significant spikes in volatility and correlation during market crashes. These dynamic market conditions are driven by a complex interplay of behavioral, political, and economic factors. A natural approach to understanding these variations is the identification of *financial regimes* – distinct time periods characterized by relatively homogeneous behavior that persists over extended durations. The study of regime-switching models originates from the work of Hamilton (1989), which primarily focused on macroeconomics and the study of business cycles. It was further extended into the realm of finance by Turner et al. (1989) and Rydén et al. (1998), with a specific focus on stock market returns. Building upon these seminal works, various regime-switching models have been developed and applied to financial data (Hardy, 2001; Ang and Bekaert, 2002; Choi and Hammoudeh, 2010).

Following the initial development of regime models, researchers began applying them to a range of downstream tasks, including portfolio construction. Many studies (Ang and Bekaert, 2004; Guidolin and Timmermann, 2007; Tu, 2010) have highlighted the advantages of integrating regime information into portfolio management strategies. Particularly after the 2008 financial crisis, the limitations of traditional strategic asset allocation, which aimed to create static “all-weather” portfolios, became evident. This led to a growing interest in *dynamic asset allocation*, which adapts portfolio weights based on market regime information to leverage favorable conditions and mitigate potential drawdown (Sheikh and Sun, 2012). Subsequently, various regime-switching asset allocation models have been introduced (van Vliet and Blitz, 2011; Bae et al., 2014; Nystrup et al., 2015; Kim and Kwon, 2023).

To incorporate regime information into portfolio decisions, we propose an interpretable four-step framework for the development of a regime-aware asset allocation strategy that includes 1. regime identification, 2. regime forecast, 3. regime-based portfolio model, and 4. out-of-sample testing. The first step constitutes an unsupervised learning problem, as the market regime remains a latent variable necessitating labeling for each period. The second step turns into a supervised problem focused on forecasting the probability of future regime states. Subsequently, regime forecasts are integrated into a portfolio optimization model, such as risk parity, in diverse ways. The final step involves the assessment of strategy performance over an independent test period. Numerous

pieces of literature discussed in this article align with this overarching four-step process, and we find this framework valuable for its interpretability and flexibility. This multi-step approach draws from foundations in financial and behavioral economics and holds promise for achieving strong empirical performance. In contrast, an alternative methodology in the form of conditional portfolio optimization (Tuck et al., 2022; Chan et al., 2023) bases portfolio decisions on observable market conditions, with regime characteristics implicitly embedded into the strategy. While this method integrates regime dynamics indirectly, our direct, step-wise approach aims to provide enhanced strategy interpretability.

Among the studies on regime-switching asset allocation, the work of Bulla et al. (2011) holds particular relevance to our research. Their study explores a basic two-asset scenario, switching between an all-cash and all-index position based on forecasts of whether the next period belongs to the high-volatility regime. In this article, we follow their principles because their straightforward strategy framework forms a solid foundation for further study. We integrate their strategy seamlessly into our four-step process, introducing a key advancement in the first step of regime identification. Here, we replace hidden Markov models (HMMs) with *statistical jump models*<sup>1</sup> (JMs), a recently introduced robust model known for its ability to identify persistent market regimes. This modification has the potential to address some of the limitations of HMMs in financial applications, including issues related to statistical accuracy (Nystrup et al., 2020b,a) and the persistence of inferred hidden state sequences (Bulla, 2011).

## 1.1 Contributions

This article employs jump models to construct a practical regime-aware asset allocation strategy. This strategy, developed following our proposed four-step framework, is rigorously evaluated in a realistic out-of-sample testing that considers transaction costs. To ensure a fair comparison with HMMs, we construct our jump model using a straightforward feature set comprising return and volatility features derived solely from the price series itself. While prior research has demonstrated the improved statistical accuracy achieved by JMs in identifying hidden state sequences, our main contribution lies in showcasing that this can have tangible financial implications to generate excess

---

<sup>1</sup>It is worth noting that the statistical jump models discussed herein are not related to *jump-diffusion models*, a common class of stochastic processes.

profits.

Our second contribution centers on the selection of the jump penalty, a critical hyperparameter in the application of jump models. We introduce a time-series cross-validation methodology, which directly optimizes the performance metric of downstream financial tasks linked to JMs. More precisely, we select the jump penalty that maximizes the Sharpe ratio of the JM-guided strategy over the validation period. In contrast to prior approaches that primarily rely on statistical criteria, our innovation lies in selecting a jump penalty that maximizes the practical benefits for financial applications associated with JMs.

The empirical results, based on the daily return series from three major US indices spanning over a 50-year period, suggest that during the validation period, the impact of the jump penalty on strategy performance is favorable, facilitating the selection of the optimal jump penalty at the point where performance metrics peak. In the subsequent test period, our strategy, guided by the chosen jump penalty, outperforms both HMM-based and buy-and-hold strategies across various performance metrics, including return, Sharpe ratio, and maximum drawdown. Visualization reveals that the regimes identified by JMs are more persistent and interpretable compared to those identified by HMMs, leading to higher returns.

The paper proceeds as follows: Section 2 delves into the methodology, explaining the details of JMs and the asset allocation strategy framework within the context of our proposed four-step process. Section 3 covers the data and features used in our analysis. In Section 4, we present empirical results, comparing strategy performance and visualizing identified regimes. Finally, our paper concludes with key insights and implications in Section 5.

## 2 Methodology

### 2.1 Statistical Jump Models

Hidden Markov models (HMMs) have been a foundational regime-switching model in financial data analysis, owing to their natural representation of hidden state variables as regimes. Nevertheless, recent studies have highlighted that inherent characteristics in financial time series, including state persistence & imbalance, time-varying parameters, and limited sample sizes, have led to questions about HMM’s statistical accuracy for regime identification (Nystrup et al., 2020b,a),

potentially affecting downstream performance<sup>2</sup>. Additionally, HMMs, as parametric models, are susceptible to model mis-specification in both the conditional distributions and the sojourn time distribution (Bulla and Bulla, 2006), producing inferred state sequences that lack genuine persistence (Bulla, 2011). The excess regime shifts can potentially lead to elevated turnover and transaction costs. To address these challenges, various non-parametric, data-driven alternatives like trend filtering (Mulvey and Liu, 2016) and spectral clustering HMMs (Zheng et al., 2021) have emerged. This article focuses on the application of jump models as one such alternative.

In a recent work, Bemporad et al. (2018) introduced statistical jump models (JMs), an unsupervised learning model that fits multiple model parameters to a time series while incorporating temporal information. Specifically, given an observation sequence of  $D$  (standardized) features  $\mathbf{Y} := \{\mathbf{y}_0, \dots, \mathbf{y}_{T-1}\}$ , where  $\mathbf{y}_t \in \mathbb{R}^D$  for all  $t$ , we estimate a JM with  $K$  states by solving the optimization problem

$$\arg \min_{\Theta, \mathbf{S}} \sum_{t=0}^{T-1} l(\mathbf{y}_t, \boldsymbol{\theta}_{s_t}) + \lambda \sum_{t=1}^{T-1} \mathbb{1}_{\{s_{t-1} \neq s_t\}}, \quad (1)$$

where  $\Theta := \{\boldsymbol{\theta}_k \in \mathbb{R}^D : k = 0, \dots, K - 1\}$  represents the collection of  $K$  model parameters,  $\mathbf{S} := \{s_0, \dots, s_{T-1}\}$  denotes the unobserved state sequence,  $l(\cdot, \cdot)$  is a loss function between observations and model parameters, and  $\lambda \in \mathbb{R}_+$  is a hyperparameter called the *jump penalty*. Each hidden state  $s_t$  takes on values in  $\{0, \dots, K - 1\}$  and is associated with a state-specific model parameter  $\boldsymbol{\theta}_{s_t}$ . We adhere to the standard practice of employing the scaled squared  $\ell_2$ -distance as the loss function, defined by  $l(\mathbf{y}, \boldsymbol{\theta}) := \frac{1}{2} \|\mathbf{y} - \boldsymbol{\theta}\|_2^2$ . After estimating a jump model, we can infer the transition probability matrix based on the optimal hidden state sequence and compute metrics such as mean and volatility within each cluster. In this article, the optimal hidden state  $\hat{s}_t$  serves as the regime label for period  $t$ , a crucial role in our analysis.

The objective function (1) represents a tradeoff between fitting the data with multiple model parameters and incorporating our prior beliefs about the persistence of the state sequence. At its core, the hyperparameter  $\lambda$  acts as a regulator for the fixed-cost regularization term, which is incurred whenever there is a transition between states. With a jump penalty of zero, the model simplifies to a  $k$ -means clustering algorithm, disregarding temporal information. As  $\lambda$  increases, state transitions become less frequent, until eventually all data points are classified into a single

---

<sup>2</sup>As highlighted by Hess (2006), “a wrong regime forecast not only may lead to a non-optimal but to a detrimental allocation in the contrary direction relative to the ‘neutral’ single state”.

cluster. Previously, the optimal selection of the hyperparameter  $\lambda$  has relied on statistical criteria such as information criteria (Cortese et al., 2023b) or in-sample accuracy (Nystrup et al., 2020b). This article, however, introduces a novel approach utilizing a cross-validation technique specifically tailored to the downstream financial applications, as elaborated in the Step four of the next subsection.

Minimizing objective function (1) jointly over model parameters  $\Theta$  and the hidden state sequence  $S$  typically involves a coordinate descent algorithm with two stages: 1. optimizing model parameters given a fixed state sequence, and 2. optimizing the state sequence given fixed model parameters. The second stage is crucial for online state inference in a real-time live setting, as discussed in Step one of the next subsection. Due to the non-convex nature of the objective function, we run the coordinate descent algorithm ten times, each starting from a different initial value generated by the  $k$ -means++ algorithm, and retain the best result<sup>3</sup>.

Bemporad et al. (2018) demonstrate that, under certain assumptions, JMs nest HMMs in the probabilistic setting. Subsequently, Nystrup et al. (2020b) introduced JMs to financial applications, utilizing the feature set initially proposed in Zheng et al. (2021)<sup>4</sup>. Their work showcases that JMs can achieve improved statistical accuracy compared to HMMs, when applied to financial data, and can yield a more persistent hidden state sequence.

Several extensions to JMs have emerged. Nystrup et al. (2021) introduced a feature selection procedure on top of the jump model for handling high-dimensional feature sets, which gives JMs an advantage over HMMs in such scenarios. Additionally, Aydinhan et al. (2023) expanded JMs into the *continuous statistical jump model* (CJM) by generalizing the discrete hidden state variable into a probability vector over all states. This extension establishes a probabilistic interpretation where the hidden state vector represents the probability of each period belonging to each regime. While these probability values can be useful for the Step three of portfolio optimization models, our studies employ a straightforward 0/1 weight in the portfolio model, making no significant difference between using JM and CJM in terms of final strategy performance. Consequently, we present results for JM alone for simplicity.

---

<sup>3</sup>Our HMM implementation using the Baum-Welch algorithm is also initialized ten times for improved performance.

<sup>4</sup>A detailed explanation of the feature set we use can be found in Section 3.2.

## 2.2 Regime-Aware Asset Allocation Strategy

In this subsection, we will detail our asset allocation strategy. Before explaining each of the four steps, we want to emphasize three key principles drawn from Bulla et al. (2011). First, we use daily data instead of monthly data to reduce the impact of incorrect regime forecasts, which now typically span only a few days instead of a whole month. Second, we employ a two-state regime model, with regimes distinguished primarily by their volatility. In terms of notation, we assume the inferred hidden state  $\hat{s}_t \in \{0, 1\}$ , with  $\hat{s}_t = 1$  indicating the high-volatility regime. Throughout our discussion, we will occasionally refer to the high-volatility regime as the market crash or bear market. Additionally, we impose a conservative transaction cost of 10 basis points for each one-way trade throughout the study.

**Step One: Regime Identification** The first step involves real-time regime identification, where we use JMs to infer each day’s hidden state  $\hat{s}_t$  as the regime label, immediately after the market return data  $r_t$  for that day becomes available, without using any data beyond day  $t+1$ <sup>5</sup>. To achieve this, we initially fit a full two-state JM every six months, using a 2000-day lookback window covering a typical bull/bear market cycle. In essence, we update the model parameters biannually. Then within each six-month interval, we perform online state inference to obtain the regime label  $\hat{s}_t$  in a live setting. Specifically, for each day  $t$ , we run the second stage of the coordinate descent algorithm, minimizing the objective function (1) over the hidden state sequence while keeping model parameters fixed at their most recently fitted values, using data from the beginning of the most recent fitting window up to day  $t$ . After obtaining the optimal hidden state sequence, we extract the final day’s state as the online state inference  $\hat{s}_t$  for day  $t$ , ensuring that the regime label  $\hat{s}_t$  relies solely on information available up to that day. Our approach differs from Bulla’s, which applied an HMM to a rolling window of the same length but fitted a full model every day without performing online inference. We believe that the regime model parameters  $\Theta$  should remain consistent rather than changing daily. The jump penalty selection process is discussed in Step four.

---

<sup>5</sup>It is worth emphasizing that, due to the unsupervised nature of regime identification, we can infer the regime for day  $t$  only at the end of that day when market data for day  $t$  becomes available, a task referred to as “nowcasting” by James et al. (2019). Regime forecasting, i.e. predicting the regime for day  $t+1$  at the end of day  $t$ , is addressed as a separate task in Step two of our four-step process.

**Step Two: Regime Forecast** We follow the same approach as Bulla et al. (2011) by using the regime label  $\hat{s}_t$  as the forecast  $f_{t+1}$  for tomorrow. This assumption relies on the belief that daily market regimes exhibit high persistence and do not change frequently. Some studies have treated the regime forecasting task as a separate classification algorithm, using information up to day  $t$  to predict the regime label  $\hat{s}_{t+1}$  for day  $t + 1$ . These labels can either be learned from Step one, as demonstrated in Uysal and Mulvey (2021), or obtained from external sources such as the NBER recession indicators, as shown in James et al. (2019). We leave the comparison between a simple forecast, as presented here, and a separate machine learning classification algorithm for future research to determine which approach is superior.

**Step Three: Portfolio Model** We maintain a two-asset world, switching between a weight of 0.0 or 1.0 in the index based on the regime forecast  $f_{t+1}$ . If the next day is predicted to be in the high-volatility regime, we allocate the entire investment to the risk-free asset; otherwise, it all goes into the index. As demonstrated by Bulla et al. (2011), this volatility-averse portfolio strategy effectively reduces volatility while modestly improving the overall strategy’s Sharpe ratio after accounting for transaction costs. While other portfolio models in a two-asset world, such as weighting as a linear function of probability forecast, exist, Nystrup et al. (2016) has shown that they fall short of the seemingly simple 0/1 strategy in terms of performance. Alternative possibilities, such as short selling the index, fail to achieve a reduction in volatility and consequently do not yield satisfactory results. We leave the exploration of integrating regime forecasts into advanced portfolio models like the Black–Litterman model (Kolm et al., 2021) within our four-step process to future research.

It’s worth noting that selling all index positions shortly after buying them is impractical for portfolio managers. Hence, the persistence of the inferred regime label sequence  $\{\hat{s}_t\}$  is crucial. Bulla et al. (2011) address this by applying a median-filtering procedure on the label sequence, which, in essence, means that the investor only switches asset allocation positions when a majority of days within a waiting period indicate a regime shift. Nystrup et al. (2015), utilizing an adaptively estimated HMM, adopts a similar procedure, shifting positions only when “the confidence in the (regime shift) prediction is above 95%”. These procedures enhance performance by mitigating the lack of persistence seen in HMM’s inferred state sequence. In our approach, thanks to the inherent

persistence offered by JMs, we can avoid such filtering altogether, eliminating the introduction of additional hyperparameters related to the number of waiting days or confidence levels.

**Step Four: Out-of-sample Testing** We employ a single time-series data split, resulting in training, validation, and test periods<sup>6</sup>. We identify the optimal jump penalty  $\lambda$  based on the strategy’s performance during the validation period and apply this  $\lambda$  throughout the test period to assess its effectiveness. To achieve this, for each candidate value of  $\lambda$ , we aggregate its online forecasts over the validation period and calculate various strategy performance metrics<sup>7</sup>. The jump penalty that leads to the highest Sharpe ratio achieved during the validation period is selected. Our analysis indicates that this chosen  $\lambda$  also excels in other performance metrics. To ensure the robustness of this selection, we employ a long enough validation period in the empirical studies. We then apply this selected penalty on the test period to evaluate and report the final performance metrics.

### 3 Data and Features

#### 3.1 Data

The data used in this article includes daily total return series for three major US equity indices: S&P 500, DJIA, and Nasdaq-100, sourced from the Bloomberg Terminal. Table 1 provides details on their tickers, start dates, and the validation/test period durations, all extending until the end of October 2023. We clarify that the strategy is constructed independently for each of the three indices. For the US risk-free rate, we employ the US Treasury bills’ 3-month constant maturity yield from Global Financial Data, following the guidance of Fairbanksa et al. (2021).

Equity Index	Ticker	Start Date	Validation Period	Test Period
S&P 500	SPX	1960-1	1968 – 1990	1990 – 2023
DJIA	INDU	1960-1	1968 – 1990	1990 – 2023
Nasdaq-100	NDX	1985-7	1993 – 2002	2002 – 2023

Table 1: Bloomberg Tickers, Start Dates, and Validation/Test Period Durations for the Three US Equity Indices.

---

<sup>6</sup>While a nested cross-validation process, which would update the optimal value of  $\lambda$  periodically (e.g., annually), could potentially enhance performance, we prioritize robustness by seeking a unique optimal jump penalty that is effective over the entire test period.

<sup>7</sup>A detailed list of performance metrics can be found in Section 4.

### 3.2 Features

The quality of the feature set plays a pivotal role in enhancing the effectiveness of jump models. In this article, our goal is to develop a feature set for JMs that strikes a balance between informativeness and dimensionality, while solely relying on the log-return series itself. This approach allows a fair comparison with HMMs, our baseline method, which also focuses only on return series<sup>8</sup>. We prioritize lower dimensionality in the feature set to mitigate the sensitivity of general clustering algorithms to the challenges associated with high-dimensional data, such as collinearity. In clustering, unlike regression where a coefficient of zero can completely exclude a feature, every feature equally contributes to the squared  $\ell_2$ -distance in the loss function, unless explicitly weighted otherwise. While the ideal process involves creating an extensive feature set and refining it via a feature selection mechanism such as sparse jump models (SJMs) used in cryptocurrency markets (Cortese et al., 2023a), we leave it for future research due to the added complexity and extra hyperparameters.

In our analysis, we focus on return and volatility features derived from index return series, as these are crucial components for understanding index performance. We utilize a set of exponentially weighted moving (EWM) measures with varying half-lives, a widely accepted method in both industry and academia. The selection of half-life for these EWM measures is crucial, balancing the consideration of effects over long and short horizons. Short-horizon EWM measures respond quickly to market fluctuations but often include substantial noise prevalent in index returns. In contrast, EWM measures with long half-lives filter out noise but may be less sensitive to recent price changes. This concept is directly relevant to jump models, where features based on short half-lives exhibit high variability and thus yield less persistent regimes even when heavily penalized, while features with longer half-lives produce smoother, more persistent regimes. Due to the substantial amount of noise in daily index returns, based on our experience, we prefer features with longer half-lives, which provide a stable basis for regime identification in jump models.

The composition of our final feature set is detailed in Table 2. As highlighted in previous research (Schwert, 1989; Bulla, 2011), market regimes are predominantly characterized by their volatility profiles. Consequently, our feature set comprises three volatility-related features and one

---

<sup>8</sup>While incorporating exogenous variables like the VIX index (Nystrup et al., 2016) and economic indicators (Kim and Kwon, 2023) could enhance regime identification, their exploration in JMs is reserved for future research.

return feature, with a primary focus on the volatility aspects. This feature set is applied uniformly across all three indices, without further fine-tuning the half-lives for individual indices.

No.	Category	Feature
1	Vol	DD (hl=20)
2	Vol	DD (hl=20) - DD (hl=60)
3	Vol	DD (hl=60) - DD (hl=120)
4	Ret	Return (hl=120)

Table 2: Feature Set Utilized for Jump Models in this Article. Exponential Weighted Smoothing Applied to Both DD (Downside Deviation) and Returns with the Specified Half-lives.

For the three volatility features, we choose exponentially weighted moving downside deviation (EWM-DD) over standard deviation (STD). DD focuses on downside risks by considering only negative returns and excluding potential upside in its calculation (Grootveld and Hallerbach, 1999), reducing the likelihood of mis-classifying periods with high volatility while positive returns as bear markets, as seen during periods of 1997-1999 preceding the dot-com bubble burst. Another advantage we've observed with DD is its tendency to revert to the long-term average faster than STD, especially during market rebound phases that typically follow the end of a market crash. This characteristic allows our strategy to more promptly transition back to an all-index position, capitalizing on the market rebounds. Thus, EWM-DD better aligns with our goal of accurately identifying and profiting from different market regimes.

We explore the incorporation of downside deviation across various horizons, specifically considering half-lives of 20, 60, and 120 days. Figure 1 illustrates the three DD measures alongside the S&P 500 price presented on a logarithmic scale, over the period of 1960-2023. We observe that spikes in DD primarily correspond to periods of falling prices, highlighting the efficacy of using DD features in regime identification. However, directly utilizing these three measures reveals high correlation, leading to inefficiency in clustering algorithms. To overcome this, we utilize the differences between DD measures across different half-lives. This approach not only mitigates collinearity but also introduces a momentum aspect by contrasting long-horizon and short-horizon effects, efficiently synthesizing DD measures across various horizons.

In our analysis of return features, we initially considered exponentially weighted moving average returns (EWMA) with half-lives of both 60 and 120 days, as depicted in Figure 2. However, as anticipated, the EWMA proves to be notably noisy even with a half-life as long as 60 days.

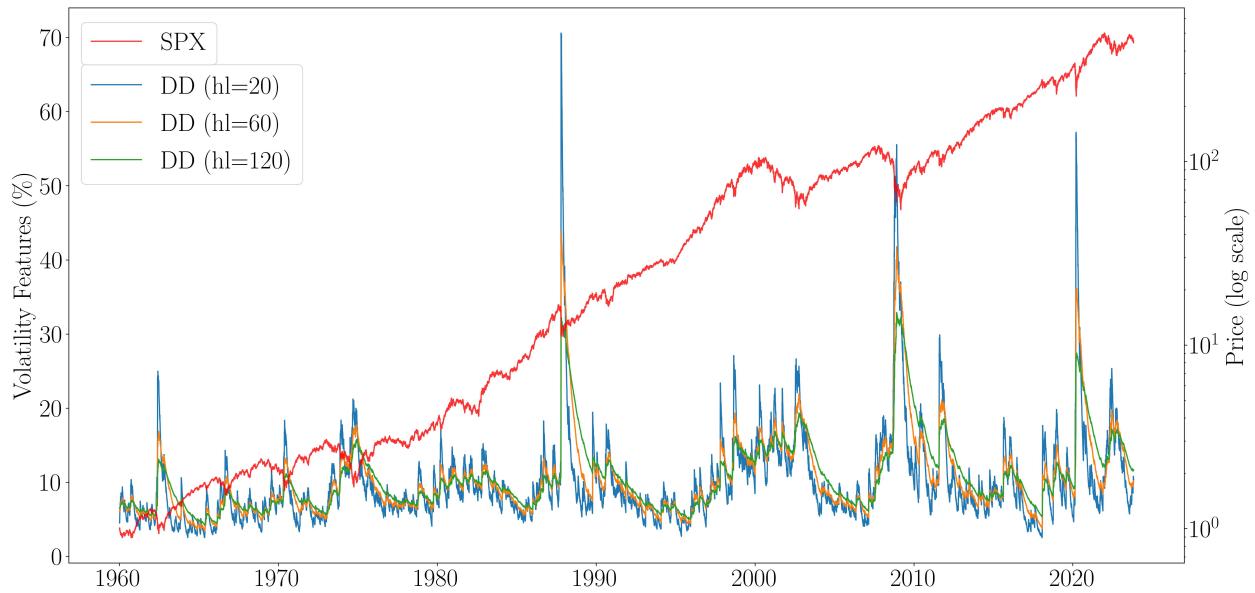


Figure 1: Volatility Features for S&P 500 over the 1960-2023 Period: Exponentially Weighted Moving Downside Deviation (DD) with Half-lives of 20, 60, and 120 Days, Alongside the Index Price in Log Scale.

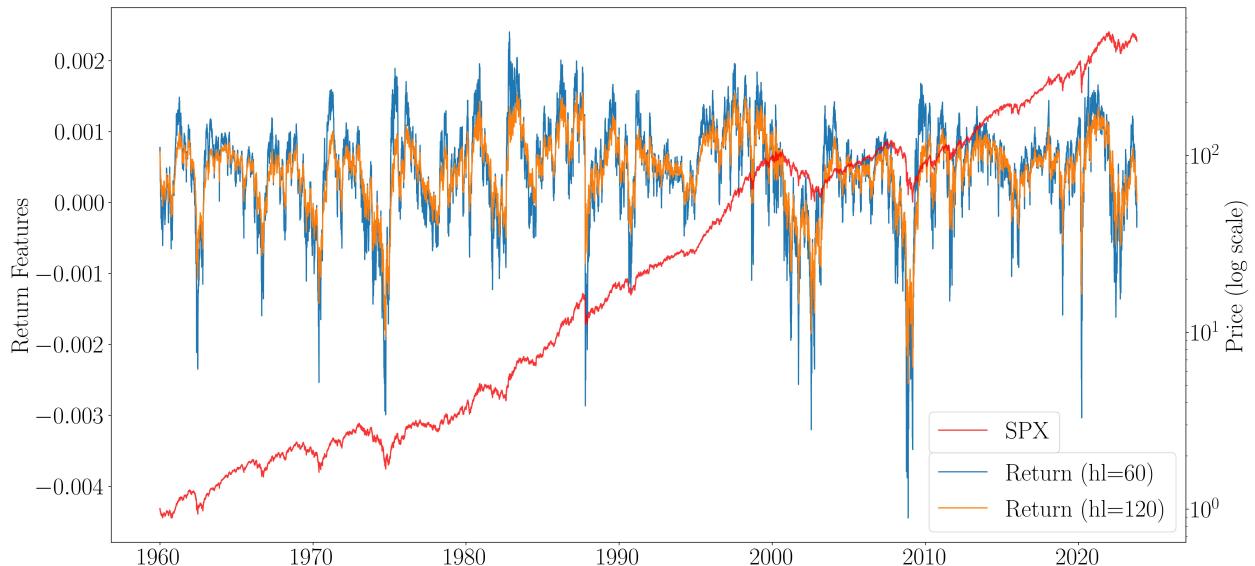


Figure 2: Return Features for S&P 500 Over the 1960-2023 Period: Exponentially Weighted Moving Average Returns with Half-lives of 60 and 120 days, Alongside the Index Price in Log Scale.

Consequently, we include only the EWMA with a half-life of 120 days in our feature set.

## 4 Empirical Results

### 4.1 Validation Period

We start our empirical studies by analyzing the strategy’s performance during the validation period, focusing exclusively on the results for S&P 500 for brevity. The out-of-sample performance for the other two indices will be presented in the next subsection. Recall that our dataset begins in 1960, with a rolling window length of 8 years, and we desire a sufficiently long validation period to ensure a robust selection of the optimal jump penalty. Therefore, for S&P 500, we select a validation period spanning from 1968 to 1990. We consider a range of candidate jump penalty values and evaluate the strategy’s performance during the validation period across various metrics, including annualized return & volatility, Sharpe ratio, downside deviation (DD), Sortino ratio, maximum drawdown (MDD), Calmar ratio, and average daily turnover. Table 3 presents these performance metrics of the strategy across all candidate  $\lambda$  values, with the performance of S&P 500 included in the last column as a reference. It is worth emphasizing that we impose a transaction cost of 10 basis points for each one-way trade throughout our empirical study.

	$\lambda = 10.$	22.	50.	100.	220.	500.	1000.	S&P 500
return	0.0993	0.1055	0.1083	0.1156	0.1089	0.0878	0.1057	0.1057
vol	0.1305	0.1316	0.1348	0.1358	0.1386	0.1422	0.1530	0.1530
Sharpe	0.2393	0.2816	0.2972	0.3445	0.2962	0.1554	0.2625	0.2625
DD	0.0948	0.0949	0.0971	0.0975	0.0993	0.1024	0.1093	0.1093
Sortino	0.3246	0.3854	0.4075	0.4739	0.4086	0.2133	0.3641	0.3641
MDD	-0.3293	-0.3293	-0.3293	-0.3293	-0.3293	-0.3550	-0.4480	-0.4480
Calmar	0.0935	0.1110	0.1201	0.1403	0.1233	0.0615	0.0888	0.0888
turnover	0.0069	0.0043	0.0029	0.0018	0.0014	0.0011	0.0000	0.0000

Table 3: Performance Metrics of JM-based Strategy with Candidate Jump Penalty Values During Validation Period (1968-1990), Alongside S&P 500 Performance in the Last Column. “DD” Represents Downside Deviation, and “MDD” Represents Maximum Drawdown. 10 bps Transaction Cost Applied.

In Table 3, we make several observations regarding how different  $\lambda$  values affect the strategy’s performance. As  $\lambda$  decreases, turnover increases, indicating more frequent regime shifts, as expected due to the smaller jump penalty. Simultaneously, both volatility and downside deviation

decrease with smaller  $\lambda$  values, resulting from the reduced market exposure during high-volatility periods. Another risk measure, maximum drawdown, also improves when regime switching is incorporated into the strategy. Additionally, concerning return and risk-adjusted return metrics including Sharpe, Sortino, and Calmar ratios, a clear trend is evident that they initially increase and then decrease with smaller  $\lambda$  values, forming an inverted U shape. Four  $\lambda$  values between 22.0 and 220.0 outperform the index across various metrics. Furthermore, nearly all the metrics reach their optimal values at  $\lambda = 100.0$ , indicating it as the best-performing jump penalty during the validation period, and we will apply this value throughout the test period in the next subsection. These observed patterns remain consistent for the other two indices, underscoring the robustness of our approach in selecting the optimal jump penalty.

The U-shape pattern arises because when  $\lambda$  is set too large, regime switches become prohibitively costly, causing the model to overlook prolonged market crashes and exhibit delayed reactions to market conditions. Conversely, a too-small  $\lambda$  leads to an excessive focus on separating daily returns into two clusters, akin to a  $k$ -means algorithm, disregarding temporal consistency. As a result, the model reacts to minor, noise-driven fluctuations in input features that do not persist into the following day. This lack of persistence, combined with high turnover, results in sub-optimal performance, mirroring the shortcomings observed in HMMs, as visualized in the next subsection.

For a visual representation of regime identification, in Figure 3 we plot the performance curve of our strategy with the optimally selected  $\lambda = 100.0$ , shown in orange, over the validation period from 1968 to 1990, together with the index price in blue color, both in a logarithmic scale. The shaded areas highlighted in red indicate the forecasted high-volatility regime, during which we allocate all assets into the risk-free asset. Notably, our jump model effectively identifies two major crashes around 1970 and 1974, during which our strategy outperforms the index. In the 1970 crash, our strategy avoids a significant drawdown with high volatility, while benefiting from a considerably high risk-free rate during the stagflation period. The crash in 1974 witnesses a substantial outperformance of our strategy in terms of return compared to the index. However, it's worth noting that we were unable to detect the 1987 Black Monday due to its rapid and drastic drop, which our features typically do not respond to that quickly. Furthermore, our identified regimes exhibit high persistence, requiring only 10 asset re-allocations over a 22-year period. Although this may not align with the trading frequency of a real portfolio manager, our model provides valuable

signal with reduced false alarms.

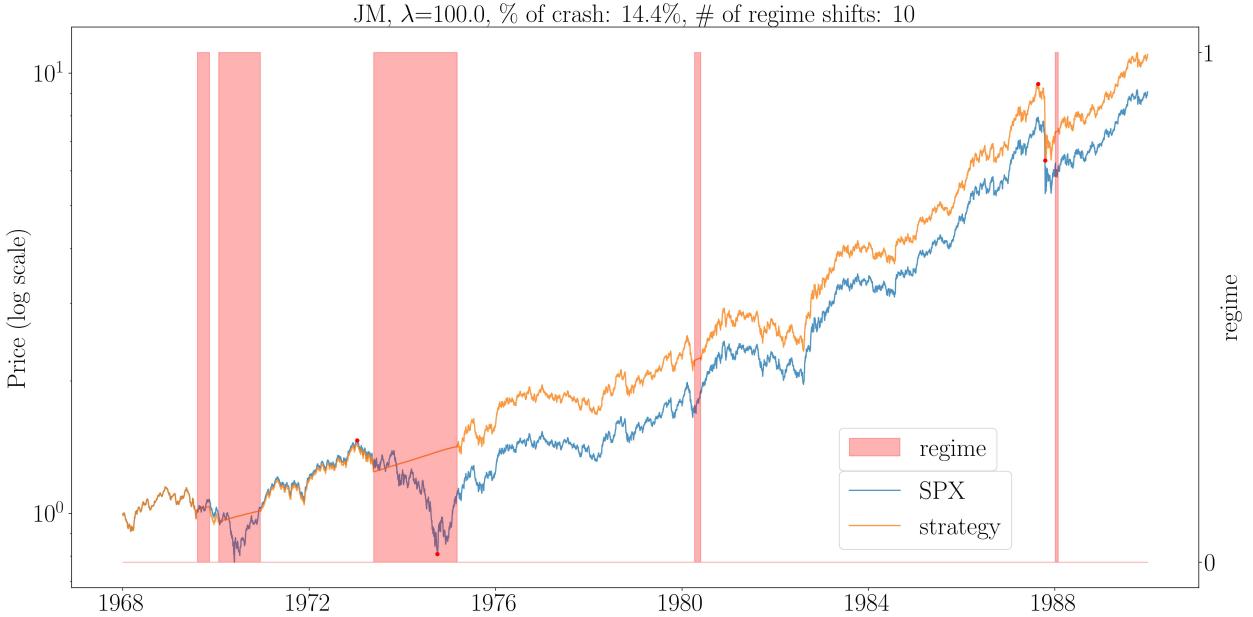


Figure 3: Performance Curve of JM-based Strategy (Orange) with Optimal  $\lambda = 100.0$  During Validation Period (1968-1990), Alongside S&P 500 Price (Blue) in Log Scale. Red Shaded Areas Indicate Forecasted High-Volatility Regimes with Full Risk-Free Allocation. Red Dots Mark Maximum Drawdown Peak/Trough. 10 bps Transaction Cost Applied.

## 4.2 Test Period

After selecting the jump penalty that leads to the best performance during the validation period, we apply it consistently throughout the test period. Table 4 presents a comparison of out-of-sample performance metrics between the JM-based regime-switching strategy (with an optimal  $\lambda = 100.0$ ), the HMM-based strategy, and the S&P 500 index. It is worth noting that the hidden state sequence generated by the HMM has undergone the median filtering procedure as implemented in Bulla et al. (2011); without this filtering, the performance is worse. Comparing the HMM-based strategy with the index itself, we observe that the former achieves higher risk-adjusted return metrics, such as Sharpe and Sortino ratios, primarily by significantly reducing volatility at the expense of return. Notably, it achieves the most substantial improvement in controlling drawdown.

In contrast, when comparing the HMM- and JM-based strategies, our approach realizes a boost in annualized return, increasing from 9.82% of the market to 12.55%. This substantial improvement translates into higher risk-adjusted return, with our Sharpe ratio reaching 0.78, compared to the

HMM’s modest increase from 0.46 to 0.51. The outperformance seen in Sortino and Calmar ratios echoes the same story. Furthermore, our turnover is notably lower than that of the HMM, highlighting that the persistence inherent in JM not only reduces transaction costs but, more importantly, enhances returns by avoiding excessive trading.

	JM ( $\lambda = 100.0$ )	HMM	S&P 500
ret	0.1255	0.0751	0.0982
vol	0.1292	0.0998	0.1822
Sharpe	0.7784	0.5149	0.4628
DD	0.0902	0.0721	0.1289
Sortino	1.1137	0.7118	0.6532
MDD	-0.2450	-0.1905	-0.5522
Calmar	0.4102	0.2693	0.1525
turnover	0.0016	0.0135	0.0000

Table 4: Performance Metrics of JM- and HMM-based Strategy During Test Period (1990-2023), Alongside S&P 500 Performance in the Last Column. “DD” Represents Downside Deviation, and “MDD” Represents Maximum Drawdown. 10 bps Transaction Cost Applied.

In Figures 4 and 5, we provide visual representations of the regimes forecasted by JM and HMM, respectively, along with the strategy’s performance curve in orange, over the test period. We also include the index price in blue, both displayed in a logarithmic scale. The high-volatility regimes identified by JM exhibit notable persistence and align with major downturns in the U.S. stock market, particularly during the dot-com bubble, the 2008 financial crisis, and the 2020 COVID-19 crash. These periods account for the majority of the outperformance achieved by our model. It is worth noting that JM is not able to identify the 2022 market downturn caused by interest rate hikes, coinciding with the maximum drawdown period of our strategy. Additionally, while JM promptly identifies the conclusion of the 2008 financial crisis, it exhibits some delay in recognizing the end of other market downturns, such as the dot-com bubble and the 2020 COVID-19 crash, missing some of the subsequent market rebound opportunities.

In contrast, HMM exhibits a notable frequency of regime shifts, totaling 115 times compared to JM’s 14, resulting in increased transaction costs for the HMM-guided strategy. Furthermore, HMM identifies numerous short-lived regimes primarily driven by the noise in daily market returns. While reducing market exposure based on these false alarms can lower volatility, it does so at the expense of potential market returns. HMM spends more time in the risk-free asset, with 34.4% compared

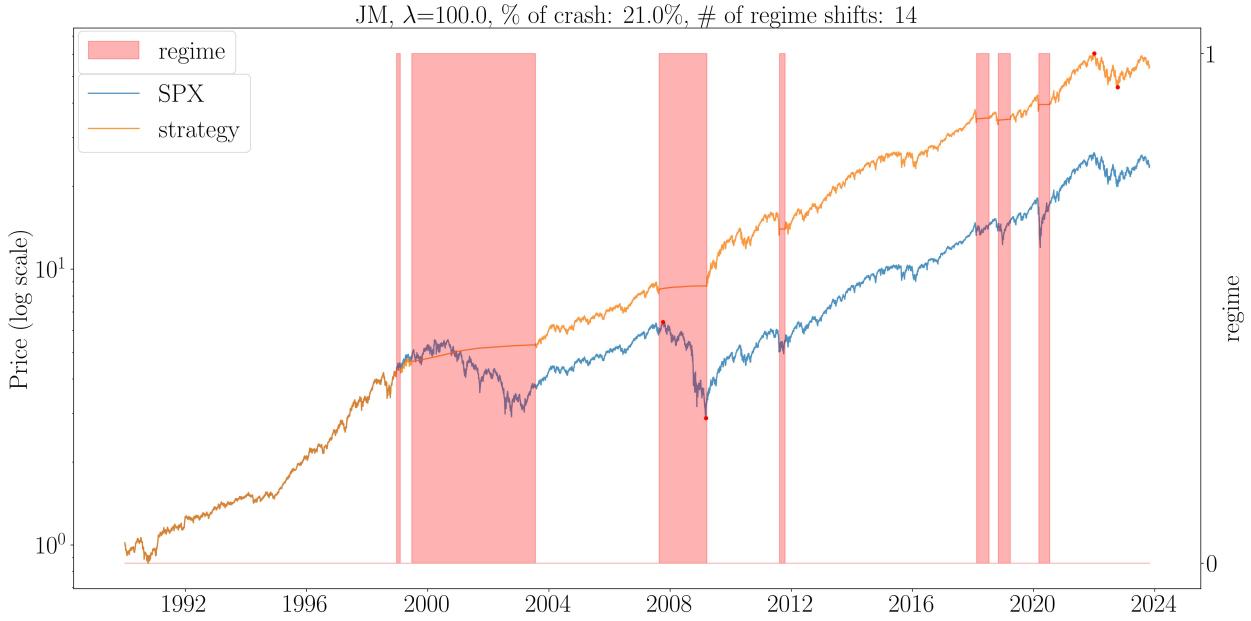


Figure 4: Performance Curve of JM-based Strategy (Orange) with Optimal  $\lambda = 100.0$  During Test Period (1990-2023), Alongside S&P 500 Price (Blue) in Log Scale. Red Shaded Areas Indicate Forecasted High-Volatility Regimes with Full Risk-Free Allocation. Red Dots Mark Maximum Drawdown Peak/Trough. 10 bps Transaction Cost Applied.

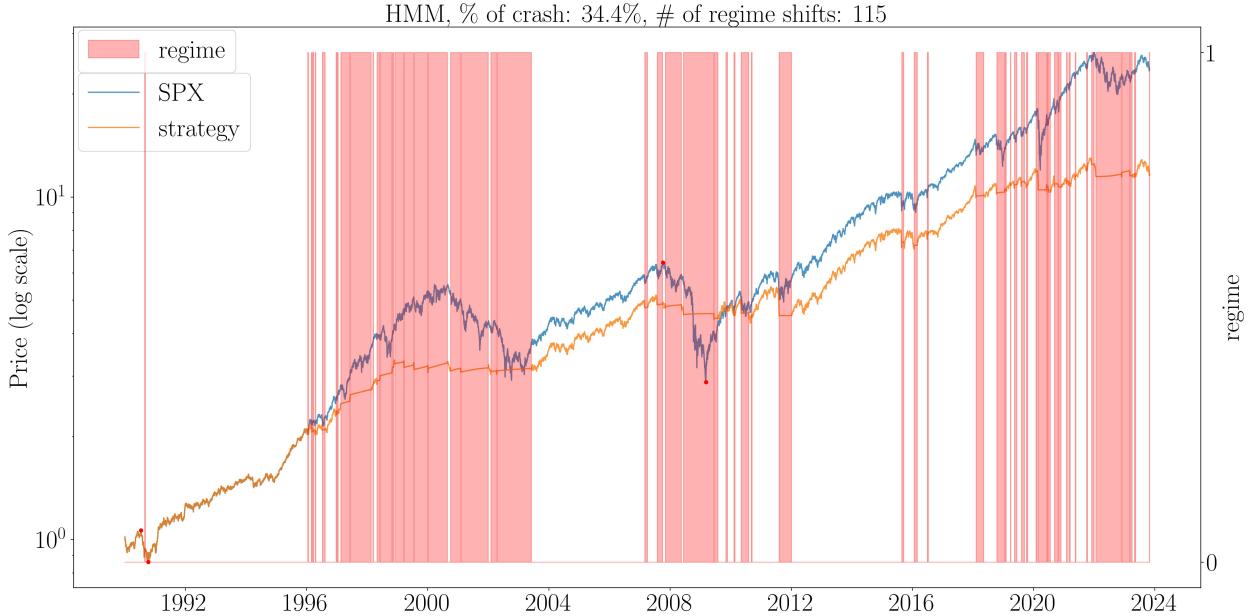


Figure 5: Performance Curve of HMM-based Strategy (Orange) During Test Period (1990-2023), Alongside S&P 500 Price (Blue) in Log Scale. Red Shaded Areas Indicate Forecasted High-Volatility Regimes with Full Risk-Free Allocation. Red Dots Mark Maximum Drawdown Peak/Trough. 10 bps Transaction Cost Applied.

to JM’s 21.0%, explaining the HMM-based strategy’s lower volatility. These characteristics make interpreting and trading HMM signals challenging. When considering the broader regime shape, the main distinctions between HMM and JM lie in two periods that are the pre-dot-com bubble from 1997 to 1999 and the 2022 interest rate hike. HMM identifies both of these periods, missing growth opportunities in the former while avoiding the downturn in the latter.

	JM ( $\lambda = 220.0$ )	HMM	DJIA
ret	0.1102	0.0717	0.1025
vol	0.1245	0.1052	0.1752
Sharpe	0.6932	0.4639	0.4964
DD	0.0878	0.0756	0.1237
Sortino	0.9816	0.6444	0.7025
MDD	-0.2073	-0.3277	-0.5178
Calmar	0.4158	0.1488	0.1678
turnover	0.0028	0.0157	0.0000

Table 5: Performance Metrics of JM- and HMM-based Strategy During Test Period (1990-2023), Alongside DJIA Performance in the Last Column. “DD” Represents Downside Deviation, and “MDD” Represents Maximum Drawdown. 10 bps Transaction Cost Applied.

	JM ( $\lambda = 10.0$ )	HMM	Nasdaq-100
ret	0.1357	0.0980	0.1156
vol	0.1643	0.1507	0.2363
Sharpe	0.7734	0.6048	0.5225
DD	0.1158	0.1074	0.1656
Sortino	1.0963	0.8469	0.7455
MDD	-0.2777	-0.3698	-0.5345
Calmar	0.4571	0.2461	0.2310
turnover	0.0024	0.0151	0.0000

Table 6: Performance Metrics of JM- and HMM-based Strategy During Test Period (2002-2023), Alongside Nasdaq-100 Performance in the Last Column. “DD” Represents Downside Deviation, and “MDD” Represents Maximum Drawdown. 10 bps Transaction Cost Applied.

The comparative performance of our JM-based strategy against the HMM-based strategy, applied to DJIA and Nasdaq-100, is highlighted in Tables 5 and 6, respectively. This comparison aligns with the results for S&P 500: the JM-guided strategy achieves an increase in annualized return by 1-2%, resulting in a more substantial improvement in the Sharpe and Sortino ratios. Additionally, there is an enhancement in the maximum drawdown for both indices compared to both the index itself and the HMM-based strategy, leading to a higher Calmar ratio. Furthermore,

the turnover is significantly reduced by approximately a factor of 5. It's important to note that the optimally selected jump penalties vary for different indices.

Finally, Figure 6 presents the visualization of the regimes identified by JM for Nasdaq-100, which is considered distinct from the other two indices. The notable persistence observed in these regimes remains consistent with previous observations, and JM effectively identifies the 2022 market downturn caused by interest rate hikes. These achievements by JM are promising, given the simplicity of our approach, which doesn't rely on advanced portfolio optimization techniques and utilizes only a minimal set of features. We demonstrate that the improved statistical accuracy of JMs in identifying market regimes has significant financial implications, enabling the generation of excess profits.

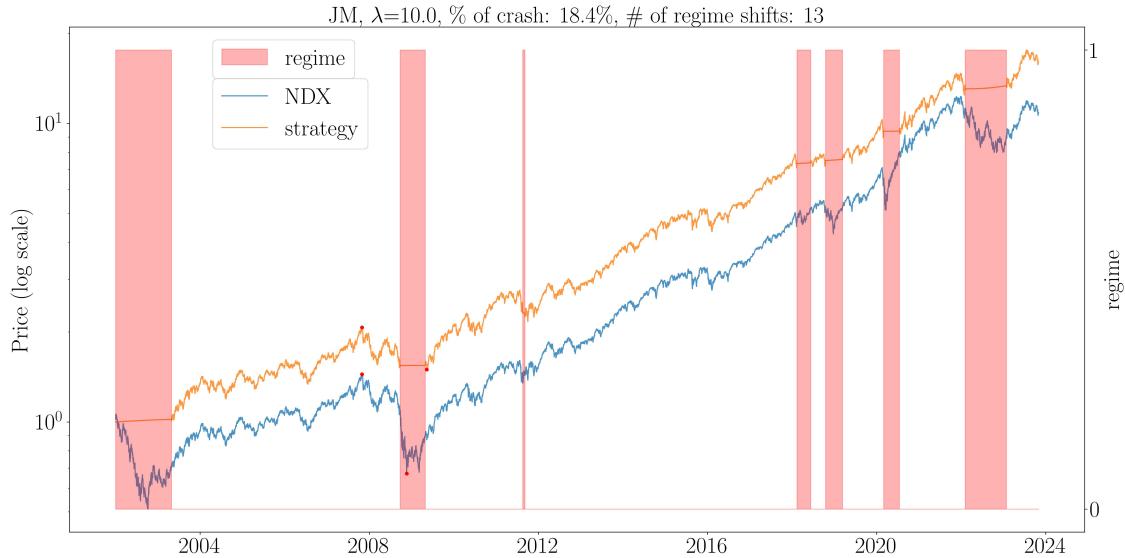


Figure 6: Performance Curve of JM-based Strategy (Orange) with Optimal  $\lambda = 10.0$  During Test Period (2002-2023), Alongside Nasdaq-100 Price (Blue) in Log Scale. Red Shaded Areas Indicate Forecasted High-Volatility Regimes with Full Risk-Free Allocation. Red Dots Mark Maximum Drawdown Peak/Trough. 10 bps Transaction Cost Applied.

## 5 Conclusion

In this article, we present an interpretable multi-step framework for the development of a regime-aware asset allocation strategy. Adhering to this framework, we develop a strategy that employs statistical jump models for the first step of regime identification, followed by an empirical evaluation of its performance through rigorous out-of-sample testing that considers transaction

costs. Our feature set for the jump model is straightforward, comprising return and volatility features derived exclusively from the return series. When it comes to selecting the jump penalty, we depart from traditional methods that rely solely on statistical criteria. Instead, we utilize a time-series cross-validation approach that directly optimizes the strategy’s performance metric over the validation period. This approach prioritizes the financial implications of the jump penalty, ensuring it benefits downstream tasks to the greatest extent. By doing so, we emphasize the practical and financial relevance of our methodology over a purely statistical perspective.

Our empirical results, spanning a 50-year period across three major US indices, suggest that the improved statistical accuracy of JMs in identifying persistent market regimes has significant financial implications for generating real profits. Specifically, throughout the validation period, the strategy’s performance exhibits an inverted U-shaped trend with decreasing jump penalty. This pattern aids in our selection of the optimal jump penalty, where we select the value corresponding to the peak performance. Subsequently, during the test period, our strategy, guided by the optimally chosen jump penalty, consistently outperforms HMM-based strategies and buy-and-hold strategies across various metrics, including annualized return, risk-adjusted return metrics like the Sharpe ratio, and maximum drawdown. Further, the visualization of regime forecasts highlights that this enhanced performance originates from the jump model’s ability to capture persistent market downturns, in contrast to HMMs, which tend to identify transient and short-lived regimes that lead to sub-optimal strategy performance.

Future directions include expanding the feature set to incorporate macroeconomic indicators and implementing feature selection in Step one. Further enhancements to Step two and three, such as exploring a multi-asset universe, developing a dedicated regime forecast algorithm, or employing advanced portfolio optimization models, hold promise for future research.

## References

- Ang, A. and Bekaert, G. (2002). Regime switches in interest rates. *Journal of Business & Economic Statistics*, 20(2):163–182.
- Ang, A. and Bekaert, G. (2004). How regimes affect asset allocation. *Financial Analysts Journal*, 60(2):86–99.

- Aydinhan, A. O., Kolm, P. N., Mulvey, J. M., and Shu, Y. (2023). Identifying patterns in financial markets: Extending the statistical jump model for regime identification. *SSRN*.
- Bae, G. I., Kim, W. C., and Mulvey, J. M. (2014). Dynamic asset allocation for varied financial markets under regime switching framework. *European Journal of Operational Research*, 234(2):450–458. 60 years following Harry Markowitz’s contribution to portfolio theory and operations research.
- Bemporad, A., Breschi, V., Piga, D., and Boyd, S. P. (2018). Fitting jump models. *Automatica*, 96:11–21.
- Bulla, J. (2011). Hidden Markov models with t components: Increased persistence and other aspects. *Quantitative Finance*, 11(3):459–475.
- Bulla, J. and Bulla, I. (2006). Stylized facts of financial time series and hidden semi-Markov models. *Computational Statistics & Data Analysis*, 51(4):2192–2209.
- Bulla, J., Mergner, S., Bulla, I., Sesboüé, A., and Chesneau, C. (2011). Markov-switching asset allocation: Do profitable strategies exist? *Journal of Asset Management*, 12:310–321.
- Chan, E., Fan, H., Sawal, S., and Viville, Q. (2023). Conditional portfolio optimization: Using machine learning to adapt capital allocations to market regimes. *SSRN*.
- Choi, K. and Hammoudeh, S. (2010). Volatility behavior of oil, industrial commodity and stock markets in a regime-switching environment. *Energy Policy*, 38(8):4388–4399.
- Cortese, F., Kolm, P., and Lindström, E. (2023a). What drives cryptocurrency returns? a sparse statistical jump model approach. *Digit Finance*.
- Cortese, F. P., Kolm, P. N., and Lindström, E. (2023b). Generalized information criteria for sparse statistical jump models. In Linde, P., editor, *Symposium i anvendt statistik, Vol 44*, Copenhagen. Copenhagen Business School.
- Fairbanksa, J. C., Griffithsb, M. D., and Wintersc, D. B. (2021). On the use of the daily fama–french risk-free rate. *Journal of Investment Management*, 19(3):39–59.

- Grootveld, H. and Hallerbach, W. (1999). Variance vs downside risk: Is there really that much difference? *European Journal of operational research*, 114(2):304–319.
- Guidolin, M. and Timmermann, A. (2007). Asset allocation under multivariate regime switching. *Journal of Economic Dynamics and Control*, 31(11):3503–3544.
- Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2):357–384.
- Hardy, M. R. (2001). A regime-switching model of long-term stock returns. *North American Actuarial Journal*, 5(2):41–53.
- Hess, M. K. (2006). Timing and diversification: A state-dependent asset allocation approach. *European Journal of Finance*, 12(3):189–204.
- James, A., Abu-Mostafa, Y. S., and Qiao, X. (2019). Machine learning for recession prediction and dynamic asset allocation. *The Journal of Financial Data Science*, 1(3):41–56.
- Kim, M. J. and Kwon, D. (2023). Dynamic asset allocation strategy: An economic regime approach. *Journal of Asset Management*, 24:136–147.
- Kolm, P. N., Ritter, G., and Simonian, J. (2021). Black–litterman and beyond: The bayesian paradigm in investment management. *The Journal of Portfolio Management*, 47(5):91–113.
- Mulvey, J. M. and Liu, H. (2016). Identifying economic regimes: Reducing downside risks for university endowments and foundations. *The Journal of Portfolio Management*, 43(1):100–108.
- Nystrup, P., Hansen, B. W., Madsen, H., and Lindström, E. (2015). Regime-based versus static asset allocation: Letting the data speak. *The Journal of Portfolio Management*, 42(1):103–109.
- Nystrup, P., Kolm, P. N., and Lindström, E. (2020a). Greedy online classification of persistent market states using realized intraday volatility features. *The Journal of Financial Data Science*, 2(3):25–39.
- Nystrup, P., Kolm, P. N., and Lindström, E. (2021). Feature selection in jump models. *Expert Systems with Applications*, 184:115558.

- Nystrup, P., Lindström, E., and Madsen, H. (2020b). Learning hidden Markov models with persistent states by penalizing jumps. *Expert Systems with Applications*, 150:113307.
- Nystrup, P., William Hansen, B., Madsen, H., and Lindström, E. (2016). Detecting change points in VIX and S&P 500: A new approach to dynamic asset allocation. *Journal of Asset Management*, 17:361–374.
- Rydén, T., Teräsvirta, T., and Åsbrink, S. (1998). Stylized facts of daily return series and the hidden Markov model. *Journal of Applied Econometrics*, 13(3):217–244.
- Schwert, G. W. (1989). Why does stock market volatility change over time? *The Journal of Finance*, 44(5):1115–1153.
- Sheikh, A. Z. and Sun, J. (2012). Regime change: Implications of macroeconomic shifts on asset class and portfolio performance. *Journal of Investing*, 21(3):36–54.
- Tu, J. (2010). Is regime switching in stock returns important in portfolio decisions? *Management Science*, 56(7):1198–1215.
- Tuck, J., Barratt, S., and Boyd, S. (2022). Portfolio construction using stratified models. In Capponi, A. and Lehalle, C.-A., editors, *Machine Learning And Data Sciences For Financial Markets*. Cambridge University Press.
- Turner, C. M., Startz, R., and Nelson, C. R. (1989). A Markov model of heteroskedasticity, risk, and learning in the stock market. *Journal of Financial Economics*, 25(1):3–22.
- Uysal, A. S. and Mulvey, J. M. (2021). A machine learning approach in regime-switching risk parity portfolios. *The Journal of Financial Data Science*, 3(2):87–108.
- van Vliet, P. and Blitz, D. (2011). Dynamic strategic asset allocation: Risk and return across the business cycle. *Journal of Asset Management*, 12:360–375.
- Zheng, K., Li, Y., and Xu, W. (2021). Regime switching model estimation: Spectral clustering hidden Markov model. *Annals of Operations Research*, 303:297–319.