A Machine Learning Approach in Regime-Switching Risk Parity Portfolios

A. Sinem Uysal and John M. Mulvey

A. Sinem Uysal
is a PhD student in the Operations Research and Financial Engineering Department at Princeton University in Princeton, NJ.
auysal@princeton.edu

John M. Mulvey
is a professor in the Operations Research and Financial Engineering Department at Princeton University in Princeton, NJ.
mulvey@princeton.edu

KEY FINDINGS

- We examine a regime prediction problem with supervised learning approaches and implement regime-switching risk parity portfolios.
- All recession periods after 1973 are captured by the random forest model, and stock market regime predictions lead to better portfolio performance.
- Regime-switching models enhance risk parity portfolios, even during a rising interest rate period. Regime-based overlay strategies provide higher risk-adjusted returns in risk parity strategies.

ABSTRACT

The authors present a machine learning approach to regime-based asset allocation. The framework consists of two primary components: (1) regime modeling and prediction and (2) identifying a regime-based strategy to enhance the performance of a risk parity portfolio. For the former, they apply supervised learning algorithms, including the random forest, based on a large macroeconomic database to estimate the probability of an upcoming recession or a stock market contraction. Out-of-sample tests show the reliability of these predictions, especially for recessions in the United States, over the period 1973 to 2020. The probability estimates are linked to a dynamic investment overlay strategy. The combined approach improves risk-adjusted returns by a substantial amount over nominal risk parity in two-asset and multi-asset test cases, even during rising interest rates in the late 1970s.

TOPICS

*Big data/machine learning, portfolio construction, performance measurement**

There is much evidence that financial markets observe abrupt changes in behavior. For example, crash periods display high volatility and increases in correlation among the return of securities, an effect called contagion. In his seminal paper, Hamilton (1989) introduced regime-switching in financial markets. Since then, many have published regime-based analysis of financial markets in economics, finance, and operations research. We observe two main research paths. The economics literature focuses on business cycle forecasting derived from various econometric methods starting from Burns and Mitchell (1946). With growing interest in big data and machine learning (ML) applications (Mullainathan and Spiess 2017; Athey 2018; Simonian and Fabozzi 2019), further studies appeared on business cycle forecasting based on ML

*All articles are now categorized by topics and subtopics. View at PM-Research.com.

methods; some recent articles on this topic are those by Yazdani (2020); Piger (2020); and James, Abu-Mostafa, and Qiao (2019). On the other hand, the finance and operations research fields predict financial market regimes rather than business cycles. A common approach is the hidden Markov model (HMM), used to define regime-switching dynamics and design investment strategies based on the inferred regime probabilities (Reus and Mulvey 2016; Nystrup, Madsen, and Lindström 2018; Fons et al. 2021).

We focus on risk-based strategies instead of the traditional mean–variance framework by Markowitz (1952). Risk budgeting overcomes the need to estimate the expected values of the asset's performance in the future. Equalizing all risk contributions is known as risk parity (RP), first introduced by Qian (2005) after the launch of the first RP fund in 1996. This framework has been studied widely, including by Chaves et al. (2011); Bruder and Roncalli (2012); Mausser and Romanko (2014); and Bai, Scheinberg, and Tutuncu (2016). Regime-switching dynamics in RP portfolios were first formally proposed by Costa and Kwon (2019).

This article's first part combines and extends the two research paths. We employ classifiers from the ML field to build predictive models for regimes for which we differentiate between two regime patterns: business cycles and stock market regime. In the second part, the inferred regime probabilities form the basis of a dynamic RP strategy. Some questions in the article are as follows: (1) Can we obtain good model performance for regime predictions with ML classifiers? (2) Which regime information is more useful in portfolio optimization? (3) How do stock market regime predictions with nonparametric classifiers perform compared with a traditional HMM in investment strategies? (4) Do we improve RP portfolio performance with regime information even in circumstances (e.g., rising interest rate environment) in which it does not perform well?

The rest of the article is organized as follow. First, we highlight previous work in the literature. Second, a detailed description of the model framework with regime prediction and RP portfolio optimization is provided. Third, we present computational results of the model, showing the benefits of regime switching for improving an RP strategy. Finally, we conclude and discuss future research directions.

LITERATURE REVIEW

The behavior of financial markets tends to change unexpectedly, and modeling these changes with regime-switching methods and other approaches has become common in finance. Ang and Timmermann (2012) pointed out several reasons for this: First, the regime concept is intuitive and usually corresponds to distinctive periods in policy and regulations in financial markets; second, regime-switching models are able to capture stylized facts about financial markets, such as fat tails, skewness, time-varying correlation, and volatility clustering. The HMM (Baum et al. 1970) has been a common approach to model regime changes in financial time-series data. Hamilton introduced regime switching to model US GNP data in his seminal work (Hamilton 1989); his regimes correspond to the National Bureau of Economic Research's (NBER's) business cycle dates.

Regime-Based Asset Allocation Strategies

Regime-based asset allocation strategies have been shown to improve portfolio performance over static asset allocation strategies (Ang and Bekaert 2004; Bulla et al. 2011). Herein, the regime-dependent parameters are inferred from the HMM for portfolio optimization. Examples of such work include that by Bae, Kim, and Mulvey (2014), who employed the HMM to identify regimes in various markets and optimize portfolios under a regime-switching framework to avoid risks generated from left-tail events. Reus and Mulvey (2016) implemented the HMM to identify regime behavior in currency markets and showed improvement in currency carry trade strategy with regime information. Nystrup, Madsen, and Lindström (2018) applied a two-state HMM to model asset returns within a control-based dynamic optimization model. Fons et al. (2021) incorporated feature selection in HMM by using a feature saliency HMM method (Adams, Beling, and Cogill 2016). They tested the regime-switching idea in factor investing, showing that the addition of a feature selection property in HMM leads to improvement in risk-adjusted portfolio performance.

Regime-based RP portfolios. Risk budgeting portfolio optimization is a popular asset allocation technique (Bruder and Roncalli 2012). Here, the risk budgets are assigned to each asset's risk contribution, and corresponding weights are found via a nonlinear optimization problem, ignoring the expected return estimations. Equalizing all risk budgets in the portfolio is known as the RP strategy. Background information on RP portfolios appears in the model framework. Two papers on regime-based RP portfolios have been published recently: Costa and Kwon (2019) examined RP portfolios under a Markov-regime-switching framework and performed experiments in equity portfolios; James, Abu-Mostafa, and Qiao (2019) considered dynamic risk budgeting between stock and bond indexes based on recession predictions from a support vector machine (SVM) model.

Business Cycle Forecasting

In the economics literature, the analysis of regimes in time-series data focuses on business cycles. In the United States, the Business Cycle Dating Committee determines the recession dates, but their methodology is not disclosed, and historically the dates are announced with a delay. For example, NBER indicated on June 2020 that, because of the COVID-19 pandemic, the US economy has been in recession since February 2020. As a consequence, the real-time prediction of economic regimes has become an important and well-studied research question. A large body of literature addresses this question, starting from Burns and Mitchell (1946) and expanded in a series of papers in the book by Stock and Watson (1993). Piger (2020) provided an extensive survey of methods of identifying business cycle turning points. Today, two ML procedures are employed by data scientists to tackle this problem: unsupervised and supervised learning.

Unsupervised learning. The first approach is similar to the traditional finance literature, in which business cycle identification is seen as an unsupervised learning problem, and some variations of Markov-switching models are applied to identify recession and expansion periods in the economy. Hamilton (1989) provided the first application of Markov-switching models in business cycle identification. Chauvet (1998) proposed a dynamic factor model with Markov-switching (DFMS) to identify phases with four monthly variables, and Chauvet and Piger (2008) expanded it by using a new real-time dataset of coincident monthly variables that include nonfarm payroll employment, the index of industrial production, real personal income excluding transfer payments, and real manufacturing and trade sales. They found that a Markov-switching model provides significant improvement over the NBER in the speed of phase identification. More recently, Camacho, Perez-Quiros, and Poncela (2018) evaluated variants of the DFMS model in identifying business cycles.

Supervised learning. The second approach treats business cycle identification as a supervised learning problem by taking NBER business cycle dates as labels and building predictive models with statistical classification methods. An early example is by Estrella and Mishkin (1998), who used probit regression to predict US recessions out of sample with various predictors: interest rate and spreads, stock prices, and monetary aggregates. They found that, beyond one quarter, the slope of the yield curve performs best as an individual predictor. Further examples include work by Estrella, Rodrigues, and Schich (2003); Kauppi and Saikkonen (2008); and Fossati (2016).

Advancements in the ML area have had a significant impact on the field of economics and challenge traditional methods (Athey 2018). The intersection of ML and econometric models and the resulting applications have become an active research topic. The availability of new large datasets and technological improvements allows researchers to explore ML methods in econometric applications. As Mullainathan and Spiess (2017) pointed out, however, finding relevant tasks and proper implementation of ML models presents an important challenge.

With the new ML methodologies, we see a change in model choices in business cycle identification. A number of authors employ ML algorithms mainly for two reasons: to improve model performance in business cycle forecasting and to incorporate high-dimensional datasets into their analyses. Qi (2001) applied neural networks to predict US recessions with 27 financial and economic indicators, whereas Ng (2012) considered a boosting ensemble of tree-based learners with a large set of macroeconomic indicators provided by James Stock and Mark Watson. Giusto and Piger (2017) proposed a nonparametric classification method known as learning vector quantization for real-time identification of business cycles and considered a dataset that contains a vintage of four monthly indicators that are suggested by NBER in US business cycle timing. Davig and Hall (2019) demonstrated naive Bayes as a recession forecasting technique; their results outperformed the Survey of Professional Forecasters[1] up to 12 months in advance. We can see the influence of developments in the ML field in finance journals as well (Simonian and Fabozzi 2019). James, Abu-Mostafa, and Qiao (2019) employed SVM for recession prediction with four coincident indicators and their lags. They showed that SVM classification errors are lower than DFMS models in the literature. Yazdani (2020) found that random forest (RF) has the best predictive performance among the various ML classifiers.

REGIME-BASED ASSET ALLOCATION FRAMEWORK

Our article extends the two research paths in regime analyses for financial markets. First, we treat regime modeling as a supervised learning problem. It differs from the traditional economics literature in two ways: (1) We incorporate not only business cycles but also stock market regimes into our models, and (2) based on regime predictions, we evaluate our model through portfolio performance. For the portfolio construction, we extend an RP approach, which is a common risk-based asset allocation strategy, based on inferred regime outputs from ML models. We demonstrate portfolio results based on regime-switching probabilities from HMM and compare them with outputs from ML models. Furthermore, we analyze possible approaches of implementing regime-switching RP portfolios and test their performance across several scenarios.

ML Approach in Market Regimes

Our regime model is set up as a supervised learning problem, wherein classification algorithms help to predict regimes. Two regime labels are considered: business cycles and stock market regime. The business cycle dates are released by NBER, which labels periods with low economic growth as recession periods and those with high economic growth as expansion periods. The stock market regimes are identified

[1] A quarterly survey for macroeconomic forecasting conducted by the Philadelphia Fed since 1990.

with the ℓ_1-trend-filtering algorithm, which was proposed by Mulvey and Liu (2016) and labels from the inputs to the supervised learning model. ℓ_1-trend filtering is a popular nonparametric unsupervised learning method that identifies trends in time-series data. For a given return time series r_t, $t = 1, \ldots, n$, assumed to consists of an underlying slowly varying trend x_t and a more rapidly varying random component z_t, which is equal to $z_t = r_t - x_t$. Our goal is to estimate the trend component x_t or, equivalently, the random component z_t. This results in an optimization problem with two competing objectives. We want the trend component, x_t, to be smooth and the random component, z_t, to be small. This problem can be formulated as an unconstrained optimization with a penalization term on the trend estimation part:

$$\min_{x \in \mathbb{R}^n} (1/2)\sum_{t=1}^{n}(r_t - x_t)^2 + \lambda \sum_{t=2}^{n}|x_{t-1} - x_t|, \tag{1}$$

where $\lambda \geq 0$ is the regularization parameter controlling the trade-off between the smoothness of x_t and the size of residuals $r_t - x_t$. The output is a piecewise linear function, and regimes can be labeled based on a prespecified threshold.

ML models. Breiman (2001b) explained two approaches in statistical modeling: data modeling and algorithmic modeling culture. The first assumes a stochastic data-generation process, and the second treats the underlying data-generation process as a black box and tries to approximate it with various functions. Here, we look at the regime modeling problem from an algorithmic approach and seek a function f that approximates the relationship between feature matrix X and regime labels y. To understand the dynamics, we explore both linear and nonlinear classifiers with ensembles for regime prediction. For further details, refer to Hastie, Tibshirani, and Friedman (2013).

Linear models. Logit and probit models are the most common linear models for classification problems. The main difference is the choice of a cumulative distribution function. Probit assumes standard normal variables, whereas logistic uses standard logistically distributed random variables for distribution specification. Early works on business cycle forecasting employ these models. We consider logistic regression along with ℓ_1 and ℓ_2 regularization terms to test linear dynamics.

Nonlinear models. Decision trees are one of the most successful nonparametric methods in ML applications in regression and classification problems. Breiman et al. (1984) proposed classification and regression trees (CART), which build decision trees by recursively partitioning the feature space and assign a label or class probability to each space for classification problems. CART takes a top-down approach by solving an optimization problem to find the best split at each node. Decision trees can capture nonlinear relationships in data and are more interpretable than other ML methods. However, there are disadvantages: Their greedy nature leads to less accurate classifiers, and a single decision tree is sensitive to training data.

Ensemble methods with tree-based classifiers help to overcome the shortcomings of single decision trees and are popular in ML owing to their high accuracy. One such method is RFs (Breiman 2001a), which work by generating an ensemble of CART trees via a bootstrap sample of training data. Another tree-based ensemble method that provides high performance in classification problems is gradient-boosted trees (Friedman 2001). Unlike RFs, gradient boosting trees are generated sequentially, and trees are grown using information from previous trees.

In our article, we observe that nonlinear models achieve better accuracies than linear models, and among nonlinear models, ensemble methods are better at capturing relationships in the data than are single decision trees. Model performance is presented in the next section with various evaluation metrics.

Evaluation metrics. The model outputs are received in two forms: probability forecasts and binary classification of regime states. Following Piger (2020), we consider multiple error metrics to evaluate a classification model's performance. A natural metric is classification accuracy (ACC). For given a probability prediction for positive state $\hat{y}_t$ and binary label of the original state y_t

$$ACC = \frac{1}{T}\sum_{t=1}^{T}[\mathbb{1}_{\{\hat{y}_t \geq c\}} y_t + (1 - \mathbb{1}_{\{\hat{y}_t \geq c\}})(1 - y_t)], \tag{2}$$

where c is the threshold to obtain regime states, which are set to natural choice of 50%. State probability forecasts are important inputs for the portfolio models, and the quadratic probability score (QPS) evaluates prediction performance in terms of probabilities. Note that QPS does not evaluate classification ability. If two models have different probability outputs, they might have same ACC score but not the same QPS, which is the mean squared error of probability forecasts.

$$QPS = \frac{1}{T}\sum_{t=1}^{T}(\hat{y}_t - y_t)^2 \tag{3}$$

In a perfect prediction setting, QPS equals 0. The lower the value of QPS, the better the model performance. Because of the imbalanced nature of the dataset, we consider other evaluation metrics to avoid inflated error metrics. The receiver operating characteristic (ROC) curve is a common choice in imbalanced classification problems. The ROC curve plots the true positive rate (recall) ($tpr = tp/tp + fn$) against the false positive rate ($fpr = fp/fp + tn$) at various threshold settings. The area under the ROC curve (AUC) generates summary statistic for the ROC metric. In a perfect prediction setting, the ROC curve will be in inverted L-shape with an AUC metric of 1. Thus, the higher the AUC value, the better the model performance.

The Matthew's correlation coefficient (MCC) is also reported as an evaluation metric and deployed widely in the field of computational biology (Matthews 1975). MCC is calculated as the following:

$$MCC = \frac{tp \times tn - fp \times fn}{\sqrt{(tp + fp)(tp + fn)(tn + fp)(tn + fn)}} \tag{4}$$

In a perfect prediction setting, MCC equals 1, and the worst value is –1. Therefore, the higher the MCC score, the better the classification performance.

Regime-Based Dynamic Asset Allocation

Next, we construct dynamic portfolios to evaluate the performance of the predictive model in an investment setting. Based on the inferred regime probabilities, asset allocations are determined in each period from the portfolio optimization. We focus on risk-based strategies instead of the traditional mean–variance framework (Markowitz 1952). Risk management has become important in portfolio analysis, and risk-based investment strategies have gained acceptance by many practitioners. Mean–variance models have practical drawbacks (Kolm, Tütüncü, and Fabozzi 2014), including sensitivity to estimated input parameters: the expected returns and the covariance matrix. Small changes in these parameters, especially in expected returns, can lead to large changes in asset allocations. Furthermore, mean–variance portfolios are concentrated in terms of the underlying risk factors.

Many straightforward approaches are proposed in the literature to manage risk within portfolio analysis. Two well-known methods are equal-weight and minimum variance portfolios. Equal weight ($1/n$) strategies are widely applied in practice and have been shown to have good out-of-sample performance (DeMiguel, Garlappi, and Uppal 2009). The downside is that it can lead to limited diversification if each asset's risk differs. The minimum variance portfolio is easy to compute but may suffer from portfolio concentration. Risk budgeting techniques seek to overcome risk concentration in portfolios.

Risk budgeting portfolio strategies. Equalizing all risk contributions in the portfolio, known as risk parity, was first introduced by Qian (2005). Starting in 2010, theoretical behaviors appeared in the academic literature. Maillard, Roncalli, and Teïletche (2010) analyzed equal risk contribution portfolios, and later Bruder and Roncalli (2012) studied properties of more a general form called risk budgeting. The main idea in the risk budgeting portfolio strategy is to decide asset allocations based on their risk contributions to the portfolio. Following Bruder and Roncalli (2012), for a portfolio with n assets, w_i represents the weight of the ith asset in the portfolio. $\mathcal{R}(w_1, \ldots, w_n)$ denotes the risk measure for portfolio $w = (w_1, \ldots, w_n)$ and verifies the following Euler decomposition if it is a homogeneous function of degree one, that is, for $f: \mathbb{R}^n \rightarrow \mathbb{R}$ and any constant c, $f(cx) = cf(x)$.

$$\mathcal{R}(w_1, \ldots, w_n) = \sum_{i=1}^{n} \underbrace{w_i . \frac{\partial \mathcal{R}(w_1, \ldots, w_n)}{\partial w_i}}_{RC_i(w_1, \ldots, w_n)} \tag{5}$$

where RC_i denotes risk contribution and the term $\partial_i \mathcal{R} = \partial \mathcal{R}(w_1, \ldots, w_n)/\partial w_i$ represents the marginal risk of asset i. For given risk budgets $\{b_1, \ldots, b_n\}$, risk budgeting allocations are obtained from the solution of the following system of nonlinear equations:

$$w^* = \{w \in [0,1]^n \mid RC_i(w_1 \ldots, w_n) = b_i, \forall i \in \{1 \ldots, n\}\} \tag{6}$$

Volatility, a common risk measure in practice, $\mathcal{R}(w) = \sqrt{w^T \Sigma w}$, is a homogeneous function of degree one, and it verifies the Euler decomposition. Under this risk measure, the marginal risk contribution of asset i is $\partial_i \mathcal{R} = (\Sigma w)_i / \sqrt{w^T \Sigma w}$, and the risk contribution of asset i will be $RC_i(w) = w_i \, \partial_i \mathcal{R} = w_i \cdot \frac{(\Sigma w)_i}{\sqrt{w^T \Sigma w}}$. Bruder and Roncalli (2012) specified the risk budgeting portfolio problem in the following form with budget and long-only constraints:

$$w^* = \left\{ w \in [0,1]^n \mid \sum_{i=1}^{n} w_i = 1, w_i.(\Sigma w)_i = b_i.(w^T \Sigma w) \right\} \tag{7}$$

where $b_i \in (0, 1]^n$ denotes the risk budget for asset i and $\Sigma_{i=1}^{n} b_i = 1$. Note that the portfolios are defined by asset weights w_i and risk budgets are in relative value. A negative risk contribution implies highly concentrated portfolios; therefore, risk budgets are limited to positive values strictly greater than zero. In the article, the authors pointed out that even with $b_i = 0$, we might get $w_i \neq 0$, which is not intuitive for investors. Therefore, the case in which risk budgets are equal to zero is omitted. The mathematical formulation in Equation 7 can be written as a least squares optimization problem:

$$
\begin{aligned}
\underset{w}{\text{minimize}} \quad & \sum_{i=1}^{n}\left(\frac{w_i(\Sigma w)_i}{w^T\Sigma w}-b_i\right)^2 \\
\text{s.t} \quad & \sum_{i}^{n} w_i = 1 \\
& 0 \le w \le 1
\end{aligned} \tag{8}
$$

and can be solved with a sequential quadratic programming algorithm. The RP portfolio presents a special case of the risk budgeting portfolio, in which risk budgets are equal, and the least squares formulation is in the following form:

$$
\begin{aligned}
\underset{w}{\text{minimize}} \quad & \sum_{i=1}^{n}\sum_{i=j}^{n}(w_i(\Sigma w)_i - w_j(\Sigma w)_j)^2 \\
\text{s.t} \quad & \sum_{i}^{n} w_i = 1 \\
& 0 \le w \le 1
\end{aligned} \tag{9}
$$

For the two-asset (n = 2) case, there exists an analytical solution that gives optimal portfolio weights.

However, for cases in which the number of assets is more than two, we need to solve optimization problems (Equations 8 and 9) to calculate the optimal weights. Note that these formulations are nonconvex, so depending on the solver's choice, we might obtain a local optimum portfolio allocation. Maillard, Roncalli, and Teïletche (2010) provided an alternative formulation for the least squares optimization approach:

$$
\begin{aligned}
\min_{y} \quad & y^T\Sigma y \\
\text{s.t} \quad & \sum_{i=1}^{n} b_i \ln(y_i) \ge c \\
& y_i \ge 0 \quad \forall i \in \{1,\ldots,n\}
\end{aligned} \tag{10}
$$

and portfolio weights are obtained by the normalization $x_i^{ERC} = y_i^* / \Sigma_{i=1}^{n} y_i^*$. They showed the equivalence of optimization problems by the KKT conditions. Note that Equation 10 is a convex optimization problem.

Challenges with RP portfolios. Risk parity portfolios have gained investors' attention, but there are critical issues. The existence of an RP portfolio is guaranteed for a long-only setup; however, Bai, Scheinberg, and Tutuncu (2016) showed that when shorting is allowed, multiple solutions exist. Furthermore, they proved that under binding asset weight constraints (Equation 9), the RP solution might not exist. Mausser and Romanko (2014) studied the computational aspects of finding RP portfolios, in which they formulated Equation 9 as a second-order cone problem. They performed tests based on various formulations and nonlinear program solver choices and found that Equation 10 gives the best results with interior point solvers. Costa and Kwon (2019) studied robust optimization formulation in a Markov-switching environment. They reported gains over the nominal RP with regime information in equity portfolios but limited gains from robust formulation.

Without leverage, from the investor point of view, the RP portfolio can be conservative. Historically, RP allocations favor fixed-income asset categories, which lead to stable but low portfolio returns, that can be enhanced with leverage and target return constraints. As a consequence, the RP performance under a high-interest-rate environment remains uncertain. Since the launch of the first RP fund in 1996, interest rates have trended lower, which might not occur in the near future. There are conflicting views on the performance of the RP portfolio under a rising interest rate environment (Kaya 2018). Chaves et al. (2011) claimed that the performance of RP is dependent on the selected asset universe. To take into account these issues, we perform portfolio analysis over disparate time periods in which the interest rate environment is different, and we conduct empirical tests with multi-asset portfolios.

Regime-based RP strategies. We consider three approaches to incorporate regime information in the experiments.

Parameter estimation. Importantly, the covariance matrix (Σ) is the only required parameter for risk-based strategies. In the nominal RP portfolio, the covariance matrix does not have any regime information, but for regime-based RP portfolios, regime prediction probabilities are used in the covariance matrix estimate ($\hat{\Sigma}$) computation. In each time period t, we calculate the crash regime predictions ($\hat{y}_t$), and historical covariance matrixes are computed for two states ($\hat{\Sigma}_{crash}, \hat{\Sigma}_{normal}$). The covariance matrix estimate then is the weighted average of these two matrixes; that is, $\hat{\Sigma} = \hat{y}_t \hat{\Sigma}_{crash} + (1 - \hat{y}_t)\hat{\Sigma}_{normal}$, where $\hat{y}_t$ is the crash regime prediction probability. The regime-based estimate of the covariance matrix is used in Equation 10 to find asset allocations.

Overlay. To enhance the returns of the RP portfolios, we consider overlay strategies. Unlike traditional leverage, our overlay strategies do not require direct capital outlay and do not affect the core portfolio structure. Mulvey, Ural, and Zhang (2007) showed that long-term investors can benefit from overlays with fix-mix strategies more than the simple leverage. We employ the risk-free rate as the borrowing rate for overlay securities and implement it based on regime information.

Dynamic risk budgeting. In this setup, risk budgets alternate between risky and risk-free assets based on regime predictions. For the periods in which crash regime prediction probability is higher than a target threshold, we allocate a larger risk budget to the less risky asset to protect the portfolio. Conversely, when crash regime prediction is low, a larger risk budget is allocated to the risky asset. We did not find this approach useful in the experiments for several reasons: It requires predefined risky and less-risky asset class notions, which may be difficult in the multi-asset case and can change over time. We observe limited gains in returns and a higher turnover rate, which might reduce or eliminate gains under transaction costs. Therefore, we excluded this strategy from our empirical results.

COMPUTATIONAL RESULTS

Sources

In the empirical tests, we collect both macroeconomic and financial market data. For regime prediction models, macroeconomic features are obtained from the dataset created by McCracken and Ng (2015), consisting of 135 macroeconomic features that represent information from various parts of the economy: (1) output and income; (2) labor market; (3) housing; (4) consumption, orders, and inventories; (5) money and credit; (6) interest and exchange rates; (7) prices; and (8) stock market.

Specific transformations suggested by the authors are applied to the macroeconomic time series to turn them into nonstationary forms, and additional lags ($l = 1$, 3, 6, 9, 12 months) of the features are considered as well. NBER business cycle dates and stock market regimes, based on the S&P 500 total return index, are labels in the prediction problem.

In the portfolio analysis, we consider monthly asset returns from major asset classes (Exhibit 1). The Kenneth R. French data library[2] provides one-month Treasury

[2] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

EXHIBIT 1

Asset Class Index Descriptions

Asset Category	Index
World Equities	MSCI World Index (MXWO)
EM Equities	MSCI Emerging Market Index (MXEF)
US Domestic Equity	S&P 500 Total Return Index
US Treasury	Barclays US Aggregate Treasury (LUATTRUU)
US Corporate Bond	Barclays US Corporate Investment Grade (LUACTRUU)
US Long Treasury	Barclays US Long Treasury (10+ years to maturity) (LUTLTRUU)
US High Yield Bond	Barclays US Corporate High Yield (LF98TRUU)
Crude Oil	S&P GSCI Commodity Total Return Index (SPGCCLAT)
Gold	LBMA Gold Price (GOLDLNPM)
Real Estate	FTSE EPRA/NAREIT Developed Total Return Index (RUGL)
Risk-Free	1-month Treasury Bill

NOTE: Bloomberg tickers are presented in parentheses.

bill returns and S&P 500 total returns, which represent the risk-free rate and the broad stock market, respectively. The overlay strategies are implemented with three securities: S&P 500 total return, Barclays US Aggregate Treasury index, and a third security based on a factor momentum tactic, introduced by Thomas and Mulvey (2020) within a retirement planning model. The reader can refer to their paper for details of the third series. We choose these securities to demonstrate the benefits of regime-switching overlays in RP portfolios.

ML Approaches in Regime-Switching Models

Backtesting framework. Our regime prediction analysis covers the sample period from 1959 to August 2020. We collect monthly macroeconomic time series as features and market regimes as labels in the classification problem. We examine two regime definitions: business cycles determined by NBER and stock market regimes identified by the ℓ_1-trend-filtering method on S&P 500 returns. The historical performance period is split into two segments: (1) 1959 to 1973 is used for hyperparameter tuning, and (2) the period from 1973 to August 2020 is used for out-of-sample testing. ML models have two types of parameters: model parameters and hyperparameters. Model parameters can be learned from data, but the hyperparameters need to be set in advance depending upon the model architecture. Important hyperparameters for the classifiers are the regularization coefficient for logistic regression, depth of the trees, number of trees that are generated in ensembles, and the maximum number of features considered in tree construction. Cross-validation is a common framework for hyperparameter tuning in the ML field. In the standard k-fold framework, training data are randomly grouped into k folds. In each iteration, the model is trained on $k-1$ folds, and the remaining fold is used for validation. After k iterations, the average of the model score is computed from validation sets. Because our dataset has a time-series property, we implement the cross-validation to keep temporal dependency and avoid look-ahead bias. We create k-folds as blocks of time periods and move training and validation sets on a rolling basis. After the model hyperparameter optimization, out-of-sample predictions are performed on a rolling window basis with a length of 150 months.

Regime prediction results. Predictive models are employed for business cycles and stock market regimes for multiple horizons ($h = 0, 1, 3, 6,$ and 12 months). Out-of-sample error metrics for linear and nonlinear classifiers appear in Exhibit 2. We observe that recessions occurred 13% of the time, whereas stock market crashes

EXHIBIT 2

Out-of-Sample (January 1973–August 2020) Model Performance Metrics of Linear and Nonlinear Classifiers

	Business Cycle				Stock Market Regime			
Model	ACC	MCC	QPS	AUC	ACC	MCC	QPS	AUC
LR	0.8853	0.5015	0.0868	0.8713	0.8244	0.2046	0.1300	0.7190
LR–ℓ_1	0.8853	0.4904	0.0826	0.8618	0.8172	0.1428	0.1390	0.6637
LR–ℓ_2	0.8871	0.5011	0.0823	0.8656	0.8190	0.1657	0.1355	0.6782
DT	0.9427	0.7564	0.0573	0.8884	0.8602	0.4893	0.1385	0.7410
RF	0.9642	0.8415	0.0337	0.9754	0.9086	0.6211	0.0724	0.9255
XGB	0.9570	0.8113	0.0332	0.9752	0.9086	0.6264	0.0735	0.9199
Horizon	**ACC**	**MCC**	**QPS**	**AUC**	**ACC**	**MCC**	**QPS**	**AUC**
$h = 0$	0.9642	0.8415	0.0337	0.9754	0.9086	0.6211	0.0724	0.9255
$h = 1$	0.9587	0.8162	0.0392	0.9777	0.9031	0.5840	0.0766	0.8996
$h = 3$	0.9440	0.7297	0.0471	0.9687	0.8917	0.4902	0.0818	0.8853
$h = 6$	0.9270	0.5951	0.0560	0.9438	0.8814	0.3899	0.0849	0.8656
$h = 12$	0.9011	0.3631	0.0716	0.8571	0.8937	0.4860	0.0861	0.8500

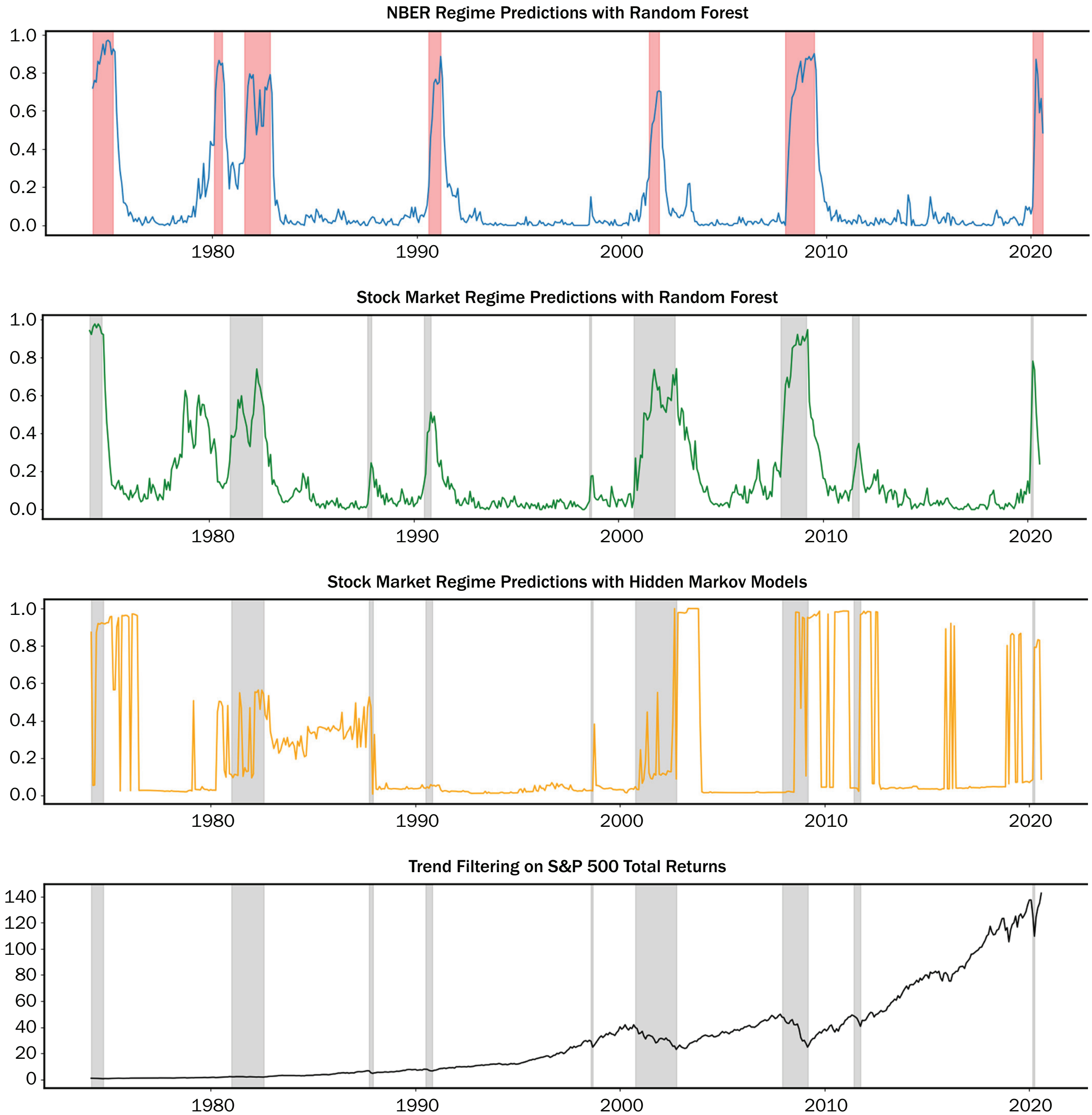
NOTE: ACC = accuracy; AUC = area under receiver operating characteristic curve; DT = decision tree; LR = logistic regression with ℓ_1 and ℓ_2 regularization; MCC = Matthew's correlation coefficient; QPS = quadratic probability score; RF = random forest; XGB = extreme gradient boosting.

occurred 17% of the time out of sample. Suppose a model always predicts a normal regime; then it has an error rate of 13% for business cycles and 17% for stock market regimes. Thus, the best classifiers' misclassification rate (1 – ACC) should be at least better than these equilibrium values. For business cycles, the RF model has the best performance among all classifiers, which aligns with the findings of Yazdani (2020). The RF has the highest accuracy rate, 96.4%, which corresponds to a 3.6% misclassification rate, lower than 13%. RF and gradient boosting trees have the best performance among all classifiers in stock market regime prediction. To stay consistent with model architecture, we choose the RF for stock market predictions. It has an accuracy rate of 90% and a misclassification rate of 10%, which is lower than the natural benchmark, 17%. In addition to ACC (accuracy score), we present other error metrics to better examine the model performance. We evaluate the performance of probability forecasts with QPS, although AUC and MCC are better metrics to measure model performance in imbalanced setting, as explained in the model framework. Overall, nonlinear models are better at capturing regimes in both setups compared to linear models across all error metrics. When we look at error metrics that are sensitive to class imbalances, ensemble models perform better than single decision trees.

Out of various classifiers, we pick the tree-based ensemble methods to forecast regimes at multiple horizons. In Exhibit 2, model errors are reported for predictions 0, 1, 3, 6, and 12 months ahead. After a one-month horizon, the model performances are deteriorating, but we still have good signals. The most recent published results on recession prediction analysis are from James, Abu-Mostafa, and Qiao (2019) and Yazdani (2020). Yazdani (2020) reported a 94% accuracy rate for the RF model over the testing period from January 2007 to the end of 2019. We observe a 96.4% accuracy rate with the RF model for recession predictions over the test period of 1973–2020. On the other hand, James, Abu-Mostafa, and Qiao (2019) used the period from 1973 to 2018 as the test sample to predict recessions with SVM and reported misclassification rates. They reported a classification error rate of 5.3% for nowcasting and 6.7% for one-month-ahead forecasting over the period of 1973–2018. We have a 3.6% classification error rate for nowcasting and 4.2%

EXHIBIT 3

Out-of-Sample Prediction Probabilities of the Three Models over the Period 1973–2020



NOTE: Red areas show National Bureau of Economic Research (NBER)–dated recession periods, and gray areas show stock market crashes identified by ℓ_1- trend-filtering algorithm with $\lambda = 0.15$.

for one-month-ahead forecasting. Previously, the stock market regime problem has been analyzed mostly in an unsupervised learning setup in which model performance evaluation by analyzing accuracy metrics is not possible.

Out-of-sample regime prediction probabilities are presented in Exhibit 3. The first plot shows recession prediction probabilities, where red areas represent NBER-dated recession dates. Note that we can capture all recession periods, even the most recent COVID-19–related crisis. The model gives recession signals starting from February 2020. The second plot shows stock market crash predictions with the RF, where gray areas show stock market crash regimes identified with the ℓ_1-trend-filtering algorithm. In the third plot, we provide out-of-sample HMM predictions based on stock market returns. Readers who are interested in the HMM and its implementation in regime detection can look at the literature review section for references. In these experiments, the signals from the HMM are less persistent and accurate than ML model predictions.

We implement three regime-switching signals in portfolio construction and compare signal quality from HMM and RF in stock market regime prediction.

Portfolio Analysis

We present computational results for portfolio analysis in three separate experiments. First, we evaluate regime-switching RP portfolios under three different regime models. Second, the RP portfolio performance is examined under two different interest rate environments. Last, we demonstrate regime-switching and nominal RP portfolios in two-asset and multi-asset cases. An overview of the multiperiod out-of-sample portfolio construction procedure is presented here. To deploy the most recent information, we implement a rolling window analysis with a length of 24 months. At the beginning of the investment period t, the regime prediction for that period is obtained from the regime-switching model, in which the 50% threshold identifies the regime state. For regime-based investment strategies, the covariance matrix is estimated with the regime information, whereas for nominal strategies only historical return data are used. We prefer the sample covariance matrix as the estimation framework for the RP portfolio because the asset universe is relatively small and RP portfolios are found to be robust to covariance misspecification (Ardia et al. 2017). Optimal RP portfolio allocations then are computed with the formulation in Equation 10. Trades are executed for the investment period t, and we repeat this procedure until the end of the investment horizon.

Portfolio performance is evaluated with various metrics. We present annualized return and volatilities and compute Sharpe ratios (historical returns of the one-month Treasury bill as the risk-free rate). Maximum drawdown measures the largest drop in the portfolio returns during the investment period. The Calmar ratio is annualized return divided by the maximum drawdown, whereas the Sortino ratio is annualized return divided by the standard deviation of downside returns (negative returns). In regime-based investment strategies, transaction costs are significant because they involve rebalancing the portfolio based on the regime information. To incorporate transaction cost effects, we consider the turnover rate metric, following DeMiguel, Garlappi, and Uppal (2009):

$$\frac{1}{T}\sum_{t=1}^{T-1}\sum_{j=1}^{N}(|\, w_{j,t}^{q} - w_{j,t+1}^{q}\,|) \tag{11}$$

for each strategy q where N denotes number of assets in the portfolio and $t = 1, \ldots, T$ represents the investment time period.

Which regime-switching model? We designate three regime-switching model specifications: The RF model provides regime predictions for two different regime labels, business cycles and stock market regimes; HMM detects stock market regimes. We evaluate signals from regime-switching models in investment strategies. We have two questions: (1) Does an RF model provide better regime-switching signals than

EXHIBIT 4

Annualized Portfolio Metrics over the Period 1975–2020

1975–2020	Ann. Ret. (%)	Ann. Vol. (%)	Max Drawdown	Sharpe Ratio	Calmar Ratio	Turnover Rate
RP-S&P 500 (RF)	8.99	5.59	0.07	0.79	1.33	0.10
RP-NBER (RF)	8.75	5.52	0.08	0.76	1.06	0.01
RP-S&P 500 (HMM)	8.71	5.56	0.07	0.74	1.17	0.12

NOTE: RP-NBER (RF) = RP portfolio with RF model with NBER regime labels; RP-S&P 500 (HMM) = RP portfolio with HMM; RP-S&P 500 (RF) = RP portfolio with RF model with S&P 500 regime labels.

HMM for the stock market regime? (2) Do stock market regime predictions improve portfolio performance more than business cycle predictions with an RF model?

To understand the regime signal's quality, we consider a two-asset case with S&P 500 total return and Barclays US Aggregate Treasury index in the regime-based RP portfolios. In the first set of experiments, the regime-switching probabilities are considered in the parameter estimation for RP portfolios. Exhibit 4 presents portfolio metrics under three different regime models. All three portfolios have similar return, volatility, and drawdown metrics, but we notice differences in the three other portfolio metrics. The RP portfolio with stock market regime predictions with the RF model has the highest Sharpe (0.79) and Calmar (1.33) ratios. The portfolio with the RF model has better performance than HMM in all three metrics: It has higher Sharpe and Calmar ratios and a lower turnover rate. The RF model with stock market regime predictions leads to a higher risk-adjusted performance and turnover rate than business cycle predictions. The lower turnover rate with the NBER model may be due to a lower regime-switching frequency in business cycles than in stock market regimes.

Exhibit 5 shows the relative performance of three regime-based RP portfolios compared with the nominal RP portfolio over time. The RF model with the stock market regime outperforms the nominal RP during the whole period, with positive relative portfolio wealth. On the other hand, the other two regime models fall behind the nominal RP after 2000. In the second set of experiments, the regime-based RP portfolios with overlay strategies are examined. Stock and bond overlays are considered in normal and crash regimes, respectively; Exhibit 6 lists the performance metrics. In both overlay strategies, the RF model with stock market regime prediction has the highest risk-adjusted performance. Stock overlays generate higher returns and volatility along with larger drawdown, but the regime model helps the portfolio to avoid a large drawdown in the 2008 crisis (the S&P 500 experienced 50% drawdown during the crash period). The RF model generates higher risk-adjusted returns and has a lower turnover rate than the HMM. Accordingly, the stock market regime prediction model leads to a better risk-adjusted return and a higher turnover rate than the business cycle prediction model. From both experiments, it is clear that the RF model with stock market regime labels provides the best out-of-sample risk-adjusted performance, and for the remainder of the article we consider it as the main baseline approach.

How does RP perform over different time periods? Before analyzing nominal and regime-switching RP portfolios, we look at the performance of RP and benchmark portfolio strategies. We focus on the two-asset case (stock and bonds) and compare the RP with the two fix-mix strategies: 60/40 and equal weight. There are conflicting views on the performance of RP portfolios in a rising interest rate environment. To address this issue, we analyze the out-of-sample performance of the portfolios over two time periods: 1975–1983 and 1983–2020. The period from 1975 to 1983 is

EXHIBIT 5

Evolution of Portfolio Wealth Relative to the Nominal RP under Three Regime Models over the Time Period 1975–2020

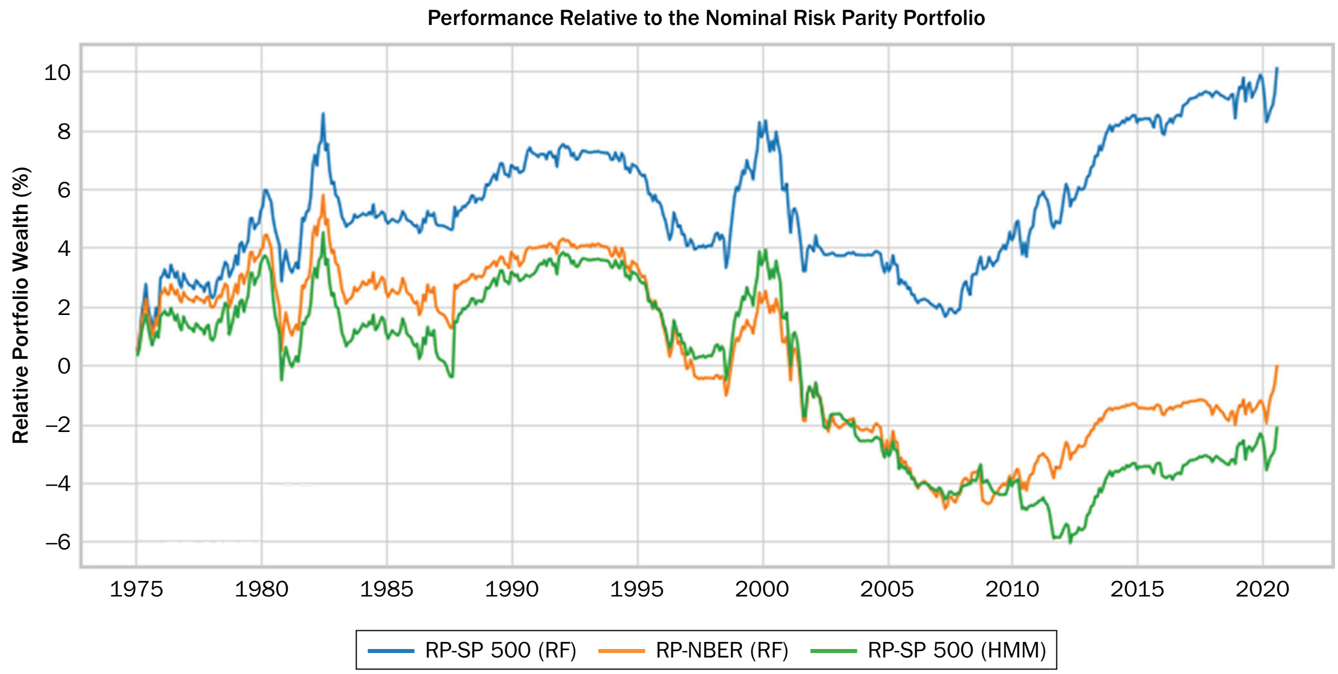


EXHIBIT 6

Annualized Portfolio Metrics for Regime-Based Overlay Strategies under Three Regime Models in Two-Asset Case

1975–2020	Ann. Ret. (%)	Ann. Vol. (%)	Max Drawdown	Sharpe Ratio	Calmar Ratio	Turnover Rate
Bond Overlay						
RP-S&P 500 (RF)	9.85	6.59	0.10	0.79	1.00	0.72
RP-S&P 500 (NBER)	9.71	7.03	0.10	0.73	0.95	0.33
RP-S&P 500 (HMM)	9.32	6.35	0.10	0.74	0.96	1.29
Stock Overlay						
RP-S&P 500 (RF)	18.81	17.48	0.36	0.79	0.52	0.72
RP-S&P 500 (NBER)	16.67	17.51	0.44	0.67	0.38	0.33
RP-S&P 500 (HMM)	14.24	17.23	0.48	0.55	0.30	1.29

characterized by increasing interest rates, whereas interest rates decline starting from 1983.

Exhibit 7 presents annualized portfolio metrics for these periods. The RP performs much better than the benchmark fix-mix portfolios (60/40, 1/*n*) over the whole period. What is interesting, however, is the RP portfolio's performance over the period 1975–1983, during which it has an annualized Sharpe ratio of 0.20, whereas both benchmark strategies have higher Sharpe ratios. During this period, bonds performed much worse than the stock market index, and we see the effect on the RP performance, which has a fixed-income-heavy asset allocation. On the other hand, the RP has a stronger performance after 1983. It has a Sharpe ratio of 0.92, whereas 60/40 and 1/*n* have 0.71 and 0.77, respectively. Although the risk-adjusted performance of the RP portfolio varies over different time periods, one characteristic that persists over the whole period is downside protection. The RP portfolio has the

EXHIBIT 7

Annualized Portfolio Metrics over Three Different Time Periods

Period	Ann. Ret. (%)	Ann. Vol. (%)	Max Drawdown	Sharpe	Calmar Ratio
1975–2020					
S&P 500 TR	13.50	15.37	0.50	0.56	0.27
US Agg Treasury	7.26	5.18	0.07	0.51	0.98
Fix-Mix 60/40	10.96	9.53	0.30	0.65	0.36
Risk-Parity	8.76	5.59	0.07	0.73	1.17
1/*n*	10.34	8.20	0.24	0.69	0.43
1975–1983					
S&P 500 TR	16.79	15.68	0.17	0.48	1.01
US Agg Treasury	9.07	6.98	0.07	0.04	1.22
Fix-Mix 60/40	13.65	10.72	0.09	0.42	1.59
Risk-Parity	10.38	7.59	0.06	0.20	1.68
1/*n*	12.87	9.64	0.08	0.40	1.66
1983–2020					
S&P 500 TR	12.81	15.30	0.50	0.59	0.25
US Agg Treasury	6.88	4.71	0.05	0.67	1.27
Fix-Mix 60/40	10.41	9.26	0.30	0.71	0.35
Risk-Parity	8.42	5.07	0.07	0.92	1.13
1/*n*	9.81	7.86	0.24	0.77	0.41

NOTE: 60/40 portfolio refers to 60% stock and 40% bond allocations.

lowest drawdown among benchmark portfolio strategies. To consider this difference in light of the interest rate environment, we examine the regime-switching RP portfolio performance in these two different historical periods.

Regime-based RP portfolios. We evaluate the regime-switching RP investment strategy's performance relative to nominal RP strategies in the two-asset and multi-asset cases. To this end, we add a factor momentum series as an overlay in normal periods and report results in both cases.

Two-asset case. Portfolio performance is tested in the two-asset case with stock and bond indexes going back to the 1970s, which tests the regime-switching RP portfolio performance in a rising interest rate environment. Exhibit 8 presents results for regime-switching and nominal RP portfolios with our overlay strategies. Regime-switching portfolios have better risk-adjusted performance than their nominal counterparts under all three scenarios. In summary, the overlay strategies enhance the RP portfolio returns, which can be attractive for investors who desire higher returns with RP portfolios. The factor momentum overlay improves nominal RP's risk-adjusted return by 30% in the whole time period. Over the period with rising interest rates (1975–1983), the nominal RP portfolio has a Sharpe ratio of 0.20, whereas the regime-based stock overlay strategy has a 0.44 Sharpe ratio, beating fix-mix benchmarks in the same period. The same findings occur in this period, in which regime-switching portfolios provide better portfolio performance than their nominal counterparts. This result is useful because uncertainty in interest rates is one of the main concerns in the RP portfolio performance, and regime-switching RP portfolios have potential to protect investments in this environment.

Relative portfolio wealth during the time periods 1975–2020 and 1975–1983 are shown in Exhibit 9. The RP with regime-switching has positive relative portfolio wealth, even in the rising interest rate environment. The 100% bond overlay has better relative performance than its regime-switching counterpart even though it has a lower Sharpe

EXHIBIT 8

Annualized Metrics for Nominal and Regime-Switching RP Strategies in Two-Asset Case

	Ann. Ret. (%)	Ann. Vol. (%)	Max Drawdown	Sharpe Ratio	Calmar Ratio	Sortino Ratio	Turnover Rate
1975–2020							
RP	8.76	5.59	0.07	0.75	1.17	3.22	0.21
RP-S&P 500 (RF)	8.99	5.59	0.07	0.79	1.33	3.37	0.10
Bond Overlay							
100% Overlay	11.69	9.93	0.19	0.70	0.62	2.21	0.21
Regime-Based	9.85	6.59	0.10	0.79	1.00	3.17	0.72
Stock Overlay							
100% Overlay	18.16	19.59	0.55	0.67	0.33	1.47	0.21
Regime-Based	18.81	17.48	0.36	0.79	0.52	1.77	0.72
Factor Momentum Overlay							
100% Overlay	11.25	7.04	0.12	0.92	0.91	3.09	0.21
Regime-Based	11.28	6.72	0.09	0.97	1.30	3.44	0.72
1975–1983							
RP	10.38	7.59	0.06	0.20	1.68	3.08	0.25
RP-S&P 500 (RF)	11.22	7.53	0.06	0.31	2.03	3.31	0.19
Bond Overlay							
100% Overlay	10.72	13.96	0.19	0.13	0.57	1.44	0.25
Regime-Based	12.69	9.71	0.09	0.38	1.34	2.91	1.45
Stock Overlay							
100% Overlay	18.55	22.10	0.29	0.41	0.63	1.44	0.25
Regime-Based	17.76	18.90	0.19	0.44	0.94	1.62	1.45
Factor Momentum Overlay							
100% Overlay	9.64	7.82	0.12	0.10	0.78	2.28	0.25
Regime-Based	11.35	8.05	0.09	0.30	1.30	2.85	1.45

ratio (Exhibit 8). On the other hand, the regime-switching bond overlay has positive relative performance in the rising interest rate environment, whereas the pure overlay strategy has negative relative portfolio wealth in the interest rate environment. Stock overlays generate positive relative wealth in both scenarios, along with the highest cumulative performance among the three options. Investors who are seeking higher returns with RP strategies may find stock overlay options profitable.

Multi-asset case. We extend the experiments to the multi-asset case, which includes major asset class indexes. Out-of-sample portfolio results are reported over the time period 1997–2020 in Exhibit 10. The regime-based RP has a higher Sharpe ratio and lower turnover rate than the nominal RP. Exhibit 11 shows the asset allocation of these two portfolios over the time period 1997–2020, wherein the nominal RP encounters more switches, which can hurt portfolio performance owing to transaction costs. The regime-switching RP strategy has positive relative performance (Exhibit 12). The overlay strategies enhance portfolio returns, especially stock overlays. The regime-switching overlay with the factor momentum series improves nominal RP's risk-adjusted return by 46%. However, a 100% bond overlay, with the US Aggregate Treasury index, has a higher risk-adjusted return than its regime-switching counterpart, which can be the result of a favorable macroeconomic environment for bonds. We want to stress that these experiments are performed only in the low interest rate environment, which might not be the case going forward. The regime-switching stock overlay has higher risk-adjusted performance than the

EXHIBIT 9

Evolution of Portfolio Wealth Relative to the Nominal RP Portfolio in Two-Asset Case

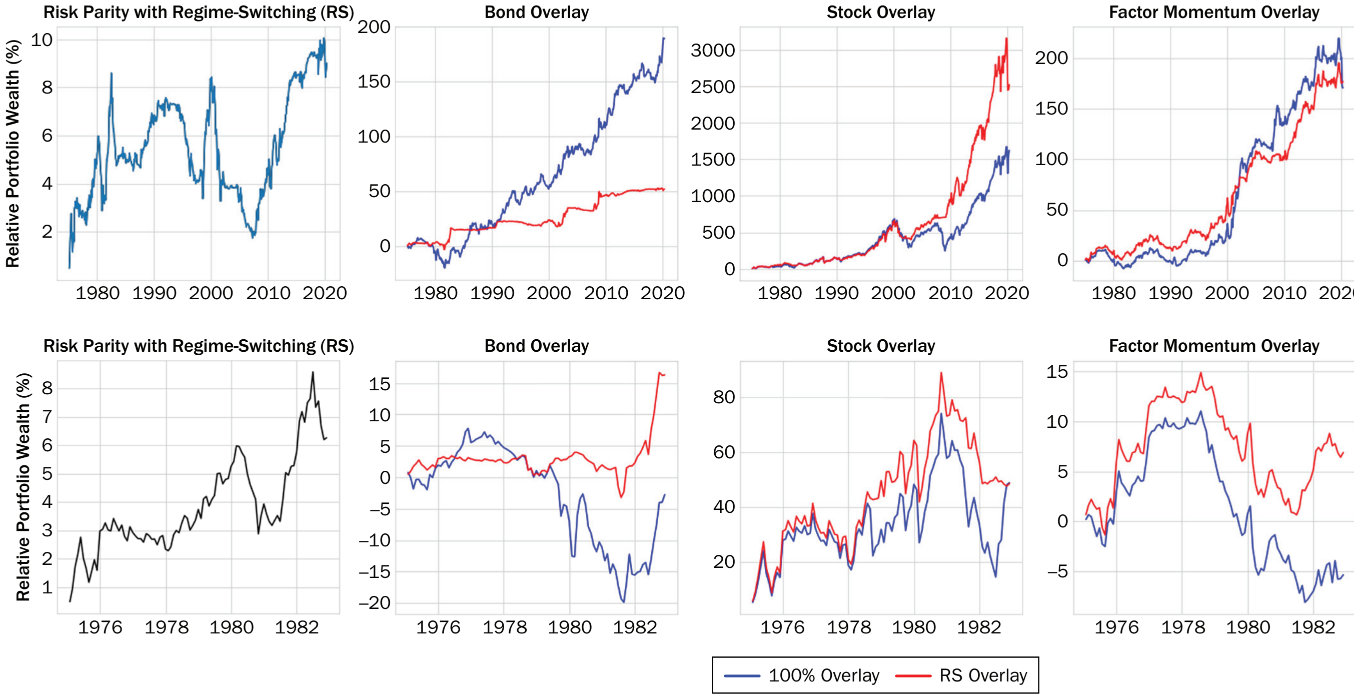


EXHIBIT 10

Annualized Metrics for Nominal and Regime-Switching RP Strategies in the Multi-Asset Case

	Ann. Ret. (%)	Ann. Vol. (%)	Max Drawdown	Sharpe Ratio	Calmar Ratio	Sortino Ratio	Turnover Rate
1997–2020							
RP	7.01	5.73	0.14	0.85	0.49	2.00	0.57
RP-S&P 500 (RF)	7.33	5.62	0.13	0.93	0.55	2.16	0.25
Bond Overlay							
100% Overlay	10.24	8.25	0.14	0.98	0.73	2.18	0.57
Regime-Based Overlay	8.29	6.49	0.14	0.95	0.61	2.26	0.92
Stock Overlay							
100% Overlay	15.62	19.81	0.58	0.67	0.27	1.17	0.57
Regime-Based Overlay	17.54	16.49	0.26	0.92	0.67	1.69	0.92
Factor Momentum Overlay							
100% Overlay	11.15	7.43	0.15	1.21	0.76	2.74	0.57
Regime-Based Overlay	10.62	6.80	0.13	1.24	0.80	2.83	0.92

100% overlay, as demonstrated by Sharpe, Calmar, and Sortino ratios. Out of the three strategies, stock overlay provides the best relative performance (Exhibit 12).

CONCLUSION

This article presents a regime-based risk parity framework for dynamic asset allocation where regime information is obtained through supervised learning models. The regime prediction model results are reasonable and better than the most recent

EXHIBIT 11

Asset Allocations for Nominal and Regime-Based RP Strategies in the Multi-Asset Universe

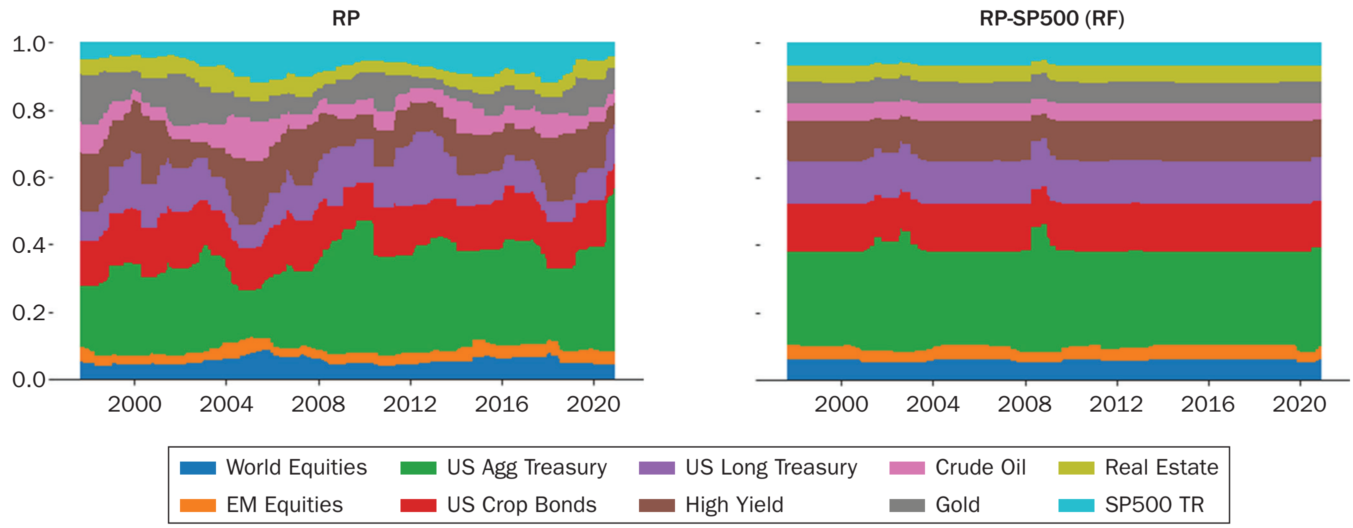
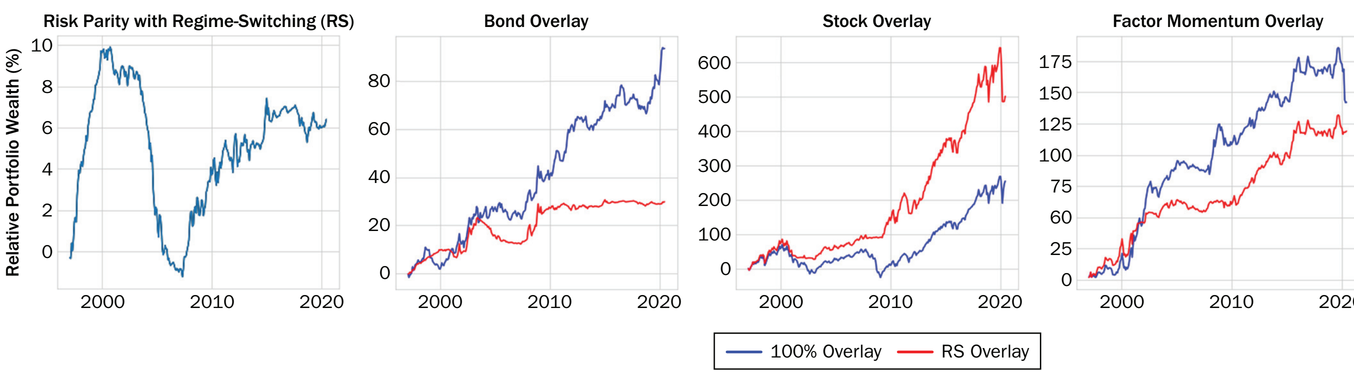


EXHIBIT 12

Evolution of Portfolio Wealth Relative to the Nominal RP Portfolio in the Multi-Asset Case



publications in business cycle forecasting. The RF model captures all recession periods out of sample; even in the most recent period, we see an increase in recession prediction probabilities starting from early February 2020. The RF model for stock market regime prediction leads to better portfolio performance than the specified HMMs. Computational results in portfolio analysis show that regime-switching RP portfolios have excellent risk-adjusted performance and positive Sharpe ratios in a rising interest rate environment. Regime-based overlay strategies enhance returns in RP portfolios, which can be beneficial for investors who are seeking higher returns within an RP framework.

REFERENCES

Adams, S., P. A. Beling, and R. Cogill. 2016. "Feature Selection for Hidden Markov Models and Hidden Semi-Markov Models." *IEEE Access* 4: 1642–1657.

Ang, A., and G. Bekaert. 2004. "How Regimes Affect Asset Allocation." *Financial Analysts Journal* 60 (2): 86–99.

Ang, A., and A. Timmermann. 2012. “Regime Changes and Financial Markets.” *Annual Review of Financial Economics* 4 (August): 313–337.

Ardia, D., G. Bolliger, K. Boudt, and J. P. Gagnon-Fleury. 2017. “The Impact of Covariance Misspecification in Risk-based Portfolios.” *Annals of Operations Research* 254: 1–16.

Athey, S. “The Impact of Machine Learning on Economics.” In *The Economics of Artificial Intelligence: An Agenda*, edited by A. Agrawal, J. Gans, and A. Goldfarb. Chicago: University of Chicago Press, 2018.

Bae, G. I., W. C. Kim, and J. M. Mulvey. 2014. “Dynamic Asset Allocation for Varied Financial Markets under Regime Switching Framework.” *European Journal of Operational Research* 234 (2): 450–458.

Bai, X., K. Scheinberg, and R. Tutuncu. 2016. “Least-Squares Approach to Risk Parity in Portfolio Selection.” *Quantitative Finance* 16: 357–376.

Baum, L. E., T. Petrie, G. Soules, and N. Weiss. 1970. “A Maximization Technique Occurring in the Statistical Analysis of Probabilistic Functions of Markov Chains.” *The Annals of Mathematical Statistics* 41 (1): 164–171.

Breiman, L. 2001a. “Random Forests.” *Machine Learning* 45 (1): 5–32.

——. 2001b. “Statistical Modeling: The Two Cultures.” *Statistical Science* 16 (3): 199–231.

Breiman, L., R. Olshen, J. Friedman, and C. Stone. *Classification and Regression Trees*. Mason, OH: Wadsworth and Brooks, 1984.

Bruder, B., and T. Roncalli. 2012. “Managing Risk Exposures Using the Risk Budgeting Approach.” *SSRN Electronic Journal* 1–33.

Bulla, J., S. Mergner, I. Bulla, A. Seboue, and C. Chesneau. 2011. “Markov-Switching Asset Allocation: Do Profitable Strategies Exist?” *The Journal of Asset Management* 12: 310–321.

Burns, A. F., and W. C. Mitchell. *Measuring Business Cycles*. New York: National Bureau of Economic Research, 1946.

Camacho, M., G. Perez-Quiros, and P. Poncela. 2018. “Markov-Switching Dynamic Factor Models in Real Time.” *International Journal of Forecasting* 34 (4): 598–611.

Chauvet, M. 1998. “An Econometric Characterization of Business Cycle Dynamics with Factor Structure and Regime Switching.” *International Economic Review* 39 (4): 969–996.

Chauvet, M., and J. Piger. 2008. “A Comparison of the Real-Time Performance of Business Cycle Dating Methods.” *Journal of Business and Economic Statistics* 26 (1): 42–49.

Chaves, D., J. Hsu, F. Li, and O. Shakernia. 2011. “Risk Parity Portfolio vs. Other Asset Allocation Heuristic Portfolios.” *The Journal of Investing* 20 (1): 108–118.

Costa, G., and R. H. Kwon. 2019. “Risk Parity Portfolio Optimization under a Markov Regime-Switching Framework.” *Quantitative Finance* 19 (3): 453–471.

Davig, T., and A. S. Hall. 2019. “Recession Forecasting Using Bayesian Classification.” *International Journal of Forecasting* 35 (3): 848–867.

DeMiguel, V., L. Garlappi, and R. Uppal. 2009. “Optimal versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy?” *The Review of Financial Studies* 22 (5): 1915–1953.

Estrella, A., and F. S. Mishkin. 1998. “Predicting U.S. Recessions: Financial Variables as Leading Indicators.” *The Review of Economics and Statistics* 80 (1): 45–61.

Estrella, A., A. P. Rodrigues, and S. Schich. 2003. “How Stable Is the Predictive Power of the Yield Curve? Evidence from Germany and the United States.” *The Review of Economics and Statistics* 85 (3): 629–644.

Fons, E., P. Dawson, J. Yau, X.-j. Zeng, and J. Keane. 2021. “A Novel Dynamic Asset Allocation System Using Feature Saliency Hidden Markov Models for Smart Beta Investing.” *Expert Systems with Applications* 163: 113720.

Fossati, S. 2016. “Dating US Business Cycles with Macro Factors.” *Studies in Nonlinear Dynamics and Econometrics* 20 (5): 529–547.

Friedman, J. H. 2001. “Greedy Function Approximation: A Gradient Boosting Machine.” *Annals of Statistics* 29 (5): 1189–1232.

Giusto, A., and J. Piger. 2017. “Identifying Business Cycle Turning Points in Real Time with Vector Quantization.” *International Journal of Forecasting* 33 (1): 174–184.

Hamilton, J. D. 1989. “A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle.” *Econometrica* 57 (2): 357–384.

Hastie, T., R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning*. New York: Springer, 2013.

James, A., Y. S. Abu-Mostafa, and X. Qiao. 2019. “Machine Learning for Recession Prediction and Dynamic Asset Allocation.” *The Journal of Financial Data Science* 3: 41–56.

Kauppi, H., and P. Saikkonen. 2008. “Predicting U.S. Recessions with Dynamic Binary Response Models.” *The Review of Economics and Statistics* 90 (4): 777–791.

Kaya, H. 2018. “How Risk Parity Can Cope with Rising Rates.” White paper, Neuberger Berman, 2018.

Kolm, P. N., R. Tütüncü, and F. J. Fabozzi. 2014. “60 Years of Portfolio Optimization: Practical Challenges and Current Trends.” *European Journal of Operational Research* 234: 356–371.

Maillard, S., T. Roncalli, and J. Teïletche. 2010. “The Properties of Equally Weighted Risk Contribution Portfolios.” *The Journal of Portfolio Management* 36 (4): 60–70.

Markowitz, H. 1952. “Portfolio Selection.” *The Journal of Finance* 7 (1): 77–91.

Matthews, B. W. 1975. “Comparison of the Predicted and Observed Secondary Structure of T4 Phage Lysozyme.” *BBA—Protein Structure* 405 (2): 442–451.

Mausser, H., and O. Romanko. 2014. “Computing Equal Risk Contribution Portfolios.” *IBM Journal of Research and Development* 58 (4): 5:1–5:12.

McCracken, M. W., and S. Ng. “FRED-MD: A Monthly Database for Macroeconomic Research.” Working paper, 2015.

Mullainathan, S., and J. Spiess. 2017. “Machine Learning: An Applied Econometric Approach.” *Journal of Economic Perspectives* 31 (2): 87–106.

Mulvey, J. M., and H. Liu. 2016. “Identifying Economic Regimes: Reducing Downside Risks for University Endowments and Foundations.” *The Journal of Portfolio Management* 43 (1): 100–108.

Mulvey, J. M., C. Ural, and Z. Zhang. 2007. “Improving Performance for Long-Term Investors: Wide Diversification, Leverage, and Overlay Strategies.” *Quantitative Finance* 7: 175–187.

Ng, S. 2012. “Boosting Recessions.” *Canadian Journal of Economics* 91 (5): 1689–1699.

Nystrup, P., H. Madsen, and E. Lindström. 2018. “Dynamic Portfolio Optimization across Hidden Market Regimes.” *Quantitative Finance* 18 (1): 83–95.

Piger, J. 2020. “Turning Points and Classification.” *Advanced Studies in Theoretical and Applied Econometrics* 52: 585–624.

Qi, M. 2001. “Predicting US Recessions with Leading Indicators via Neural Network Models.” *International Journal of Forecasting* 17 (3): 383–401.

Qian, E. "Risk Parity Portfolios." Panagora Asset Management, September 2005.

Reus, L., and J. M. Mulvey. 2016. "Dynamic Allocations for Currency Futures under Switching Regimes Signals." *European Journal of Operational Research* 253 (1): 85–93.

Simonian, J., and F. J. Fabozzi. 2019. "Triumph of the Empiricists: The Birth of Financial Data Science." *The Journal of Financial Data Science* 1 (1): 10–13.

Stock, J. H., and M. W. Watson. *Business Cycles, Indicators, and Forecasting*. Chicago: University of Chicago Press, 1993.

Thomas, D., and J. M. Mulvey. "Factor Momentum in Retirement Planning." Working paper, Princeton University, ORFE Department, 2020.

Yazdani, A. 2020. "Machine Learning Prediction of Recessions: An Imbalanced Classification Approach." *The Journal of Financial Data Science* 2 (4): 21–32.