DTU Library

Dynamic portfolio optimization across hidden market regimes

Nystrup, Peter; Madsen, Henrik; Lindström, Erik

Dynamic Portfolio Optimization
Across Hidden Market Regimes

Peter Nystrup*†‡, Henrik Madsen‡ and Erik Lindström§

†Sampension, Tuborg Havnevej 14, 2900 Hellerup, Denmark
‡Department of Applied Mathematics and Computer Science, Technical University of Denmark,
Asmussens Allé, Building 303B, 2800 Kgs. Lyngby, Denmark
§Centre for Mathematical Sciences, Lund University, Box 118, 221 00 Lund, Sweden

Regime-based asset allocation has been shown to add value over rebalancing to static weights and, in particular, reduce potential drawdowns by reacting to changes in market conditions. The predominant approach in previous studies has been to specify in advance a static decision rule for changing the allocation based on the state of financial markets or the economy. In this article, model predictive control (MPC) is used to dynamically optimize a portfolio based on forecasts of the mean and variance of financial returns from a hidden Markov model with time-varying parameters. There are computational advantages to using MPC when estimates of future returns are updated every time a new observation becomes available, since the optimal control actions are reconsidered anyway. MPC outperforms a static decision rule for changing the allocation and realizes both a higher return and a significantly lower risk than a buy-and-hold investment in various major stock market indices. This is after accounting for transaction costs, with a one-day delay in the implementation of allocation changes, and with zero-interest cash as the only alternative to the stock indices. Imposing a trading penalty that reduces the number of trades is found to increase the robustness of the approach.

1.   Introduction

The objective of portfolio optimization is to find an optimal tradeoff between risk and return over a fixed planning horizon. Traditionally, investors decide on a strategic asset allocation (SAA) based on a single-period optimization, inspired by the mean–variance framework of Markowitz (1952). The purpose is to develop a static, "all-weather" portfolio that optimizes efficiency across a range of economic scenarios. Even if the SAA is reconsidered on an annual basis, it is unlikely to change significantly, as long as the purpose is "all-weather" efficiency.

In the presence of time-varying investment opportunities, portfolio weights should be adjusted as new information arrives to take advantage of favorable economic regimes and withstand adverse regimes (Sheikh and Sun 2012). The abrupt regime changes that financial markets tend to undergo present a big challenge to traditional SAA. Although some changes may be transitory, the new behavior often persists for several periods after a change (Ang and Timmermann 2012).

---

*Corresponding author. Email: pnys@dtu.dk

Regime-based asset allocation (RBAA) has indeed been shown to add value over rebalancing to static weights and, in particular, reduce potential drawdowns by reacting to changes in market conditions (see Ang and Bekaert 2004, Guidolin and Timmermann 2007, Bulla *et al.* 2011, Kritzman *et al.* 2012, Nystrup *et al.* 2015a, 2017). The predominant approach is to specify in advance a static decision rule for changing the allocation based on the state of financial markets or the economy.

The parameters of the decision rule can be optimized in sample, but it does not guarantee that the decision rule is optimal for the problem at hand. A disadvantage is, therefore, that a large number of different specifications might have to be tried, in order to find a decision rule with good performance. Testing many different specifications increases the risk of inferior performance out of sample. Further, it can be argued that a static decision rule is hardly optimal when the underlying model used for regime inference is time-varying, as in Bulla *et al.* (2011) and Nystrup *et al.* (2015a, 2017).

An alternative approach is to dynamically optimize the portfolio based on the inferred regime probabilities and parameters taking into account transaction costs, risk aversion, and possibly other constraints. Herzog *et al.* (2007) and Boyd *et al.* (2014) proposed to use model predictive control (MPC) to solve this constrained, stochastic control problem. In MPC, a statistical model of the process is used to predict its future evolution and choose the best control action.

The great strength of MPC is the capability to solve control problems under constraints in a computationally feasible manner. Even so, it is commonly assumed that asset prices can be described by a linear factor model with constant variance and that there are no transaction costs in order to derive analytical expressions for when the allocation should be changed (see, e.g., Herzog *et al.* 2007, Costa and Araujo 2008, Calafiore 2008, 2009). This limits the practical impact of the results. Transaction costs are important when comparing the performance of static and dynamic strategies, because frequent rebalancing can offset the potential excess return of a dynamic strategy. Moreover, transaction costs stabilize the optimization problem (Brodie *et al.* 2009, Ho *et al.* 2015).

In this article, asset returns are modeled by a two-state hidden Markov model (HMM) with time-varying parameters, similar to the model considered in Nystrup *et al.* (2015a, 2016b, 2017). From a statistical perspective, the HMM is a more realistic description of asset price dynamics than a linear factor model with constant variance. It is well suited to capture the stylized behavior of financial series, including volatility clustering, leptokurtosis, and time-varying correlations (see, e.g., Rydén *et al.* 1998, Ang and Timmermann 2012). From an economic perspective, the HMM can describe the abrupt changes in market conditions and investment opportunities that arise due to changes in risk aversion and structural changes in the state of the economy.

Instead of a static decision rule for changing the portfolio based on the inferred regime, MPC is used to dynamically optimize the portfolio based on forecasted means and variances. Using an HMM, the forecasts are mean-reverting and only change when the regime probabilities change. Thus, the allocation is still determined indirectly by the inferred regime. MPC, however, is applicable to forecasts from any type of model. The impact of transaction costs and risk aversion is analyzed in a live-sample setting using available market data. MPC is compared with previous approaches to RBAA under realistic assumptions about transaction costs and implementation.

The article is structured as follows: Section 2 introduces the HMM, its estimation, and use for forecasting. Section 3 is concerned with dynamic portfolio optimization and MPC. The empirical results are presented in section 4. Finally, section 5 concludes.

2.   The Hidden Markov Model

The HMM is a popular choice for inferring the hidden state of financial markets. It can match the tendency of financial markets to change their behavior abruptly and the phenomenon that the new behavior often persists for several periods after a change (Ang and Timmermann 2012). In addition, it is well suited to capture the stylized behavior of many financial series including volatility clustering and leptokurtosis, as shown by Rydén *et al.* (1998).

In an HMM, the probability distribution that generates an observation depends on the state of an unobserved Markov chain. A sequence of discrete random variables $\{S_t : t \in \mathrm{N}\}$ is said to be a first-order Markov chain if, for all $t \in \mathrm{N}$, it satisfies the Markov property:

$$\Pr\left(S_{t+1}|\, S_t, \ldots, S_1\right) = \Pr\left(S_{t+1}|\, S_t\right). \tag{1}$$

The conditional probabilities $\Pr\left(S_{t+1} = j|\, S_t = i\right) = \gamma_{ij}$ are called transition probabilities. A Markov chain with transition probability matrix $\boldsymbol{\Gamma} = \{\gamma_{ij}\}$ has stationary distribution $\boldsymbol{\pi}$, if $\boldsymbol{\pi}^{\mathrm{T}}\boldsymbol{\Gamma} = \boldsymbol{\pi}^{\mathrm{T}}$ and $\mathbf{1}^{\mathrm{T}}\boldsymbol{\pi} = 1$, where $\mathbf{1}$ is a column vector with all entries one. The Markov chain is said to be stationary if $\boldsymbol{\delta} = \boldsymbol{\pi}$, where $\boldsymbol{\delta}$ is the initial distribution, i.e., $\delta_i = \Pr\left(S_1 = i\right)$.

As an example, consider the two-state model with Gaussian conditional distributions:

$$Y_t \sim N\left(\mu_{S_t}, \sigma^2_{S_t}\right),$$

where

$$\mu_{S_t} = \begin{cases} \mu_1, & \text{if } S_t = 1, \\ \mu_2, & \text{if } S_t = 2, \end{cases} \quad \sigma^2_{S_t} = \begin{cases} \sigma^2_1, & \text{if } S_t = 1, \\ \sigma^2_2, & \text{if } S_t = 2, \end{cases} \quad \text{and} \quad \boldsymbol{\Gamma} = \begin{bmatrix} 1 - \gamma_{12} & \gamma_{12} \\ \gamma_{21} & 1 - \gamma_{21} \end{bmatrix}.$$

When the current state S_t is known, the distribution of Y_t depends only on S_t, and not on previous states or observations.

The sojourn times are implicitly assumed to be geometrically distributed:

$$\Pr\left('\text{staying } t \text{ time steps in state } i'\right) = \gamma_{ii}^{t-1}\left(1 - \gamma_{ii}\right). \tag{2}$$

The geometric distribution is memoryless, implying that the time until the next transition out of the current state is independent of the time spent in the state.

In order to improve its fit to the distributional and temporal properties of daily returns, the Gaussian HMM has been extended by considering other sojourn-time distributions than the memoryless geometric distribution (Bulla and Bulla 2006), other conditional distributions than the Gaussian distribution (Bulla 2011), and a continuous-time formulation as an alternative to the dominating discrete-time models (Nystrup *et al.* 2015b). As an alternative to increasing the model complexity, Nystrup *et al.* (2016b) obtained good results using an adaptive estimation approach that allowed for time variation in the parameters of a two-state Gaussian HMM. This approach was adopted in Nystrup *et al.* (2015a, 2017) and will be adopted in this article as well.

2.1. *Adaptive Parameter Estimation*

The parameters of an HMM are typically estimated using the maximum-likelihood method. The two most popular approaches to maximizing the likelihood are direct numerical maximization and the Baum–Welch algorithm, a special case of the expectation–maximization (EM) algorithm (Baum *et al.* 1970, Dempster *et al.* 1977).

Every observation is assumed to be of equal importance, no matter how long the sample period is. This approach works well when the sample period is short and the underlying process does not change over time. The time-varying behavior of the parameters documented in previous studies (Rydén *et al.* 1998, Bulla 2011, Nystrup *et al.* 2016b), however, calls for an adaptive approach that assigns more weight to the most recent observations, while keeping in mind past patterns at a reduced confidence.

As pointed out by Cappé *et al.* (2005), it is possible to evaluate derivatives of the likelihood function with respect to the parameters for virtually any model that the EM algorithm can be applied to. As a consequence, instead of resorting to a specific algorithm such as the EM algorithm,

the likelihood can be maximized using gradient-based methods. Lystig and Hughes (2002) described an algorithm for exact computation of the score vector and the observed information matrix in HMMs that can be performed in a single pass through the data. Their algorithm was derived from the forward–backward algorithm.

The reason for exploring gradient-based methods is the flexibility to make the estimator recursive and adaptive.[1] The estimation of the parameters through a maximization of the conditional log-likelihood function can be done online using the estimator

$$\hat{\theta}_t = \arg\max_\theta \sum_{n=1}^{t} w_n \log \Pr\left(Y_n \,|\, Y_{n-1}, \dots, Y_1, \theta\right) = \arg\max_\theta \tilde{l}_t\left(\theta\right) \tag{3}$$

with $w_n = 1$.[2] The online estimator can be made adaptive by introducing a different weighting. A popular choice is to use exponential weights $w_n = f^{t-n}$, where $0 < f < 1$ is the forgetting factor (Parkum *et al.* 1992, Kulhavỳ and Zarrop 1993). The speed of adaption is then determined by the effective memory length

$$N_{\text{eff}} = \frac{1}{1-f}. \tag{4}$$

Maximizing the second-order Taylor expansion of $\tilde{l}_t\left(\theta\right)$ around $\hat{\theta}_{t-1}$ with respect to θ and defining the solution as the estimator $\hat{\theta}_t$ leads to

$$\hat{\theta}_t = \hat{\theta}_{t-1} - \left[\nabla_{\theta\theta}\tilde{l}_t\left(\hat{\theta}_{t-1}\right)\right]^{-1} \nabla_\theta \tilde{l}_t\left(\hat{\theta}_{t-1}\right). \tag{5}$$

This is equivalent to a specific case of the generalized autoregressive score (GAS) model of Creal *et al.* (2013). Using the estimator (5) it is possible to reach quadratic convergence, whereas the GAS model, in general, converges only linearly (see Cappé *et al.* 2005).

Scaling by the Hessian in (5) is equivalent to scaling by the variance of the score function, because the expectation of the score is zero. The variance of the score function is known as the Fisher information

$$I_t\left(\theta\right) = \mathrm{E}\left[-\nabla_{\theta\theta}l_t\right] = \mathrm{E}\left[\nabla_\theta l_t \nabla_\theta l_t^{\mathrm{T}}\right]. \tag{6}$$

Approximating the Hessian by the Fisher information leads to the recursive, adaptive estimator

$$\hat{\theta}_t \approx \hat{\theta}_{t-1} + A\left[I_t\left(\hat{\theta}_{t-1}\right)\right]^{-1} \nabla_\theta \tilde{l}_t\left(\hat{\theta}_{t-1}\right). \tag{7}$$

The tuning constant A can be adjusted to increase or decrease the speed of convergence without changing the effective memory length, although it is common to choose $A \approx 1/N_{\text{eff}}$. The inverse of the Fisher information can be updated recursively using (6) and the matrix inversion lemma. It is necessary to apply a transformation to all constrained parameters for the estimator (7) to converge and it is advisable to start the estimation at $t > 1$ to avoid large initial steps.

The time variation of the parameters is observation driven based on the score of the likelihood function. Although the parameters are stochastic, they are perfectly predictable given the past observations. This is contrary to parameter-driven models, in which the parameters are stochastic processes with their own source of error. No prior knowledge is assumed about the parameters, and

---

[1]See Khreich *et al.* (2012) for a survey of techniques for incremental learning of HMM parameters.
[2]An online estimator processes its input observation-by-observation in a sequential fashion, without having the entire input sequence available from the start.

no attempt is made to identify the drivers of the variations (see, e.g., Brennan *et al.* 1997).

The use of the score function for updating the parameters is intuitive, as it defines the steepest ascent direction for improving the model's local fit in terms of the likelihood at time t given the current parameter values (Creal *et al.* 2013). For HMMs, the score function must consider previous observations and cannot reasonably be approximated by the score function of the most recent observation, as it is often done for other models (Khreich *et al.* 2012). This leads to a significant increase in computational complexity. In order to compute the weighted score function, the algorithm of Lystig and Hughes (2002) has to be run for each iteration and the contribution of each observation has to be weighted.

2.2. *Forecasting*

The first step toward calculating the forecast distributions is to estimate the current state probabilities given the past observations and parameters. The vector of state probabilities is

$$\boldsymbol{\alpha}_{T|T}^{\mathrm{T}} = \frac{\boldsymbol{\delta}^{\mathrm{T}}\mathbf{P}_1\left(y_1\right)\prod_{t=2}^{T}\boldsymbol{\Gamma}_t\mathbf{P}_t\left(y_t\right)}{\boldsymbol{\delta}^{\mathrm{T}}\mathbf{P}_1\left(y_1\right)\left(\prod_{t=2}^{T}\boldsymbol{\Gamma}_t\mathbf{P}_t\left(y_t\right)\right)\mathbf{1}}, \tag{8}$$

where the i'th entry is $\left(\alpha_{T|T}\right)_i = \Pr\left(\left. S_T = i\right| Y_T, \ldots, Y_1\right)$ and $\mathbf{P}_t\left(y_t\right)$ is a diagonal matrix with the state-dependent conditional densities $p_i\left(y_t\right) = \Pr\left(\left. Y_t = y_t\right| S_t = i, \theta_t\right)$ as entries.

Once the current state probabilities are estimated, the state probabilities k steps ahead can be calculated by multiplying $\boldsymbol{\alpha}_{T|T}$ with the transition probability matrix k times:

$$\boldsymbol{\alpha}_{T+k|T}^{\mathrm{T}} = \boldsymbol{\alpha}_{T|T}^{\mathrm{T}}\boldsymbol{\Gamma}_T^k. \tag{9}$$

The parameters are assumed to stay constant in the absence of a model describing their evolution.

The density forecast is the average of the state-dependent conditional densities weighted by the forecasted state probabilities. When the conditional distributions are distinct normal distributions, the forecast distribution will be a non-normal mixture (Frühwirth-Schnatter 2006).[1] Using Monte Carlo simulation, Boyd *et al.* (2014) showed that the results of dynamic portfolio optimization are not particularly sensitive to higher order moments. Consequently, for the present application, only the first and second moment of the forecast distribution are considered.

The first two moments of a mixture distribution are

$$\mu = \sum_{i=1}^{m}\mu_i\alpha_i \tag{10}$$

$$\sigma^2 = \sum_{i=1}^{m}\left(\mu_i^2 + \sigma_i^2\right)\alpha_i - \mu^2 \tag{11}$$

with α_i denoting the weights—that is, the forecasted state probabilities.

Before calculating the moments of the mixture distribution, the conditional means and variances of the returns are calculated based on the moments of the log-returns. Within each state, the returns r_t are assumed to be iid with log-normal distribution

$$\log\left(1 + r_t\right) \sim N\left(\mu, \sigma^2\right),$$

where μ and σ^2 are the mean and variance of the log-returns. Thus, the mean and variance of the

---

[1]The non-normality can easily be captured by generating scenarios using the current state probabilities as initial probabilities.

returns are given by

$$\mathrm{E}\left[r_t\right] = \exp\left(\mu + \sigma^2/2\right) - 1 \tag{12}$$

$$\mathrm{Var}\left[r_t\right] = \left(\exp\left(\sigma^2\right) - 1\right)\exp\left(2\mu + \sigma^2\right). \tag{13}$$

The forecasted mean and variance will be mean-reverting, as the forecast horizon extends and the state probabilities converge to the stationary distribution of the Markov chain. The rate of convergence is determined by the size of the second largest eigenvalue of the transition probability matrix, which for a two-state Markov chain is $|\lambda_2| = |\gamma_{11} + \gamma_{22} - 1|$. The more persistent the states are—or equivalently, the larger the size of λ_2—the lower the rate of convergence.

3.   Dynamic Portfolio Optimization

The approach outlined in the previous section can be applied online to forecast the mean and variance of returns at discrete horizons. The forecasted means and variances are inputs to a multi-period portfolio optimization. Every day a decision has to be made whether or not to change the current portfolio allocation, knowing that the decision will be reconsidered the next day with new input. In the risk-neutral case, the possible gain or saving from changing allocation has to exceed the costs involved.

The planning horizon should at least be long enough to reach the stationary distribution of the underlying Markov chain, whereafter the forecast does not change. In principle, the forecast horizon should be infinitely long, but in reality no one has an infinite horizon.

There is a limit to how far ahead in time it is meaningful to make predictions. For sufficiently long horizons, it is not possible to make better predictions than the long-term mean and variance, which is also the reason that the forecasted mean and variance converge to their stationary values, when the forecast horizon extends. Thus, looking only a limited number of steps into the future is not just an approximation necessary to make the optimization problem computationally feasible, it also seems perfectly reasonable.

The formulation of the dynamic portfolio optimization problem as a stochastic control problem is inspired by Boyd *et al.* (2014), however, the objective function and the way transaction costs are handled are significantly different. Boyd *et al.* (2014) assumed that infinite amounts of cash could be entered into the portfolio on any given day. Although it would be possible to constrain the amount of cash that can be entered into the portfolio per day, it is for most purposes more realistic to assume that only a finite amount of cash is available initially. At any later point in time, the amount of cash available depends on the portfolio's development and transaction costs incurred.

3.1.   *Stochastic Control Formulation*

Let $h_t \in \mathrm{R}^n$ denote the portfolio holdings at time t, where $(h_t)_i$ is the dollar value of asset i at the end of day t, with $(h_t)_i < 0$ meaning a short position in asset i. Assets can be bought and sold at the end of each day. Let $u_t \in \mathrm{R}^n$ denote the dollar values of the trades, with $(u_t)_i > 0$ meaning that asset i is bought at the end of day t.

The post-trade portfolio is defined as

$$h_t^+ = h_t + u_t, \quad t = 0, \ldots, T-1, \tag{14}$$

which is also the portfolio at the beginning of day $t+1$. The total value of the portfolio before trading is $V_t = \mathbf{1}^\mathrm{T} h_t$ and the total value of the post-trade portfolio is $V_t^+ = \mathbf{1}^\mathrm{T} h_t^+ \leq V_t$. Working with holdings h_t, rather than weights h_t/V_t, simplifies the notation.

The portfolio is assumed to be self-financing with transaction costs proportional to the total trade

volume, that is,

$$\mathbf{1}^{\mathrm{T}} u_t + \kappa^{\mathrm{T}} |u_t| = 0, \quad t = 0, \ldots, T - 1, \tag{15}$$

where κ is a vector of commission rates and the absolute value is elementwise. The constraint states that $-\mathbf{1}^{\mathrm{T}} u_t$, which is the total gross proceeds from sales minus the total gross proceeds from purchases, equals $\kappa^{\mathrm{T}} |u_t|$, the total transaction cost for purchases and sales.

For optimization purposes, the constraint (15) is replaced by the convex relaxation

$$\mathbf{1}^{\mathrm{T}} u_t + \kappa^{\mathrm{T}} |u_t| \le 0, \quad t = 0, \ldots, T - 1, \tag{16}$$

which allows the possibility of discarding money.

The post-trade portfolio is held until the end of the next day. The (pre-trade) portfolio at the end of the next day is given by

$$h_{t+1} = (\mathbf{1} + r_{t+1}) \circ h_t^+, \quad t = 0, \ldots, T - 1, \tag{17}$$

where $r_{t+1} \in \mathrm{R}^n$ is the vector of asset returns from day t to day $t + 1$ and $\circ$ denotes Hadamard (elementwise) multiplication of vectors. As illustrated in figure 1, the dynamics are linear, but unknown at time t.

Figure 1. Timeline of portfolio dynamics.

The returns r_t are random variables with mean and covariance

$$\mathrm{E}\,[r_t] = \bar{r}_t, \quad \mathrm{E}\left[(r_t - \bar{r}_t)(r_t - \bar{r}_t)^{\mathrm{T}}\right] = \Sigma_t, \quad t = 1, \ldots, T.$$

The trades are determined in each period by a trading policy $\phi_t : \mathrm{R}^n \to \mathrm{R}^n$:

$$u_t = \phi_t\,(h_t), \quad t = 0, \ldots, T - 1.$$

Let $\mathcal{C}_t \subseteq \mathrm{R}^n$ denote the post-trade portfolio constraint set. Since $\mathcal{C}_t$ is nonempty, it follows that for any value of h_t, there exists a u_t for which

$$h_t^+ = h_t + u_t \in \mathcal{C}_t. \tag{18}$$

Explicit constraints are imposed only on the post-trade portfolio h_t^+, because this can be controlled by buying and selling (i.e., through u_t), whereas the pre-trade portfolio h_t is determined by the random return r_t in the previous period and, therefore, not directly controllable.

The portfolio may be subject to constraints on the post-trade holdings, such as minimum and maximum allowed holdings for each asset:

$$h_t^{\min} \le h_t^+ \le h_t^{\max}, \tag{19}$$

where the inequalities are elementwise and $h_t^{\min}$ and $h_t^{\max}$ are given vectors of minimum and maximum asset holdings in dollars. For a long-only portfolio with no short positions allowed $h^{\min} = 0$. Position limits can also be expressed relative to the total portfolio value, for example,

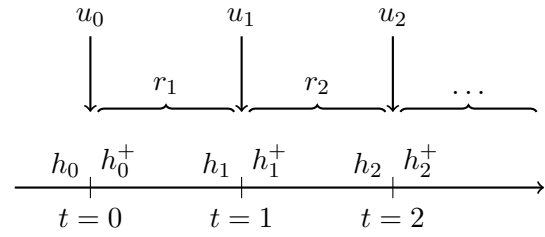$$y_t^+ \le V_t^+ H_t^{\max}, \tag{20}$$

with $H_t^{\max} \in \mathrm{R}^n$.

The overall objective is to maximize

$$J = \mathrm{E}\left[V_T - \sum_{t=0}^{T-1} \psi_t\left(h_t, u_t\right)\right], \tag{21}$$

where the expectation is over the sequence of returns $r_1, \ldots, r_T$ conditional on all past observations, $V_T = \mathbf{1}^{\mathrm{T}} h_T$ is the terminal value of the portfolio, and $\psi_t : \mathrm{R}^n \times \mathrm{R}^n \to \mathrm{R}$ is a cost, with units of dollars, for period t. This is a stochastic control problem with linear dynamics and convex objective function, ensuring the existence of a unique solution (Boyd and Vandenberghe 2004). The data for the problem is the distribution of r_t, the stage costs ψ_t, and the initial portfolio h_0.

3.1.1.   Risk-Averse Control.
The traditional risk-adjustment charge is proportional to the variance of the next period portfolio value given the current post-trade portfolio, which corresponds to

$$\psi_t\left(h_t, u_t\right) = \gamma \frac{\mathrm{Var}\left[V_{t+1}\middle| h_t^+\right]}{V_t^+} = \gamma \frac{\left(h_t^+\right)^{\mathrm{T}} \Sigma_{t+1} h_t^+}{V_t^+}, \tag{22}$$

where $\gamma \geq 0$ is a unitless risk-aversion parameter. In control theory, the inclusion of a risk penalty is referred to as risk-averse control (Whittle 1981). To ensure that the tradeoff between terminal value and variance does not depend on the portfolio value, the variance is scaled by the post-trade value V_t^+.

If the returns are independent, then the sum of the variances is the variance of the terminal value. In that case, the objective function (21) with the risk-adjustment charge (22) is equivalent to the mean–variance criterion of Markowitz (1952).[1] It is a special case of expected utility maximization with a quadratic utility function. While the utility approach was theoretically justified by von Neumann and Morgenstern (1953), in practice few, if any, investors know their utility functions; nor do the functions which financial engineers and economists find analytically convenient necessarily represent a particular investor's attitude toward risk and return (Dai *et al.* 2010a). The mean–variance criterion remains the most commonly used in portfolio optimization (Meucci 2005).

As an alternative to including a risk penalty in the objective function, Boyd *et al.* (2014) proposed to constrain the portfolio standard deviation to a fraction of the portfolio value. In a single-period setting, the two formulations are equivalent. A risk limit might be preferable, because it is easier to quantify than a risk-aversion parameter. In a multi-period setting, however, a constraint on the portfolio standard deviation leads to excessive trading and inferior performance. The resulting constant-risk portfolio does not consider the attractiveness of the risk-return tradeoff. In order to maximize the return, the portfolio will be right at the risk limit most of the time, which leads to forced trading, whenever the volatility forecast increases unexpectedly.

3.1.2.   Trading Aversion.
Boyd *et al.* (2014) gave examples of other convex constraints and cost terms that arise in practical investment problems and can easily be included. One option is to include a penalty for trading

$$\psi_t\left(u_t\right) = \rho^{\mathrm{T}}\left|u_t\right|, \tag{23}$$

where ρ is a vector of trading-aversion parameters. This could reflect a conservatism toward trading, for example, due to the uncertainty related to the parameter estimates and forecasts. Inflating the

---

[1]Everything is scaled by V_t.

transaction cost κ in (15)–(16) would have the same effect. In order to distinguish it from the actual transaction cost, the trading penalty (23) is instead included in the objective function (21), similarly to the variance penalty (22).

It is well known that estimation errors can cause mean–variance optimized portfolios to perform poorly (see Michaud 1989, DeMiguel *et al.* 2009). Brodie *et al.* (2009) showed that the mean–variance optimization problem can be reformulated as a constrained, least-squares regression problem. Imposing a trading penalty (23) that is proportional to the trade volume is a convex relaxation of constraining the number of trades. Similar to the least absolute shrinkage and selection operator (LASSO) in regression analysis (Tibshirani 1996), this ℓ_1-penalty regularizes the optimization problem and reduces the risk due to estimation errors (Ho *et al.* 2015).

3.2. *Model Predictive Control*

MPC, also referred to as receding horizon control or rolling horizon planning, is widely used in some industries; primarily for systems with slow dynamics such as energy systems, chemical plants, and supply chains (see, e.g., Bemporad 2006), but it is also used, for example, for steering autonomous vehicles (see, e.g., Falcone *et al.* 2007). MPC typically works very well in practice, even for short horizons.

MPC is based on the simple idea that in order to determine u_t, all future (unknown) returns are replaced by their forecasted mean values $\hat{r}_\tau$, $\tau = t+1, \ldots, T$. This turns the stochastic control problem into a deterministic optimization problem

$$
\begin{aligned}
\text{maximize} \quad & V_T - \sum_{\tau=t}^{T-1} \psi_\tau \left(h_\tau, u_\tau \right) \\
\text{subject to} \quad & h_{\tau+1} = \left(\mathbf{1} + \hat{r}_{\tau+1} \right) \circ \left(h_\tau + u_\tau \right), \quad \tau = t, \ldots, T-1
\end{aligned}
\tag{24}
$$

with variables $h_{t+1}, \ldots, h_T$ and $u_t, \ldots, u_{T-1}$. Note that h_t is not a variable, but the (known) current portfolio holdings.

Solving this convex optimization problem yields an optimal sequence of trades $u_t^\star, \ldots, u_{T-1}^\star$. This sequence is a plan for future trades over the remaining trading horizon under the highly unrealistic assumption that future returns will be equal to their forecasted values. An alternative is to forecast the unconditional distribution and generate a number of scenarios, but this is computationally much more challenging.

The MPC policy takes $\phi^{\text{MPC}}(h_t) = u_t^\star$, that is, only the first trade in the planned sequence of trades is executed. At the next step, the process is repeated, starting from the new portfolio h_{t+1}. In the case of a mean–variance objective function, Herzog *et al.* (2007) showed that future asset allocation decisions do not depend on the trajectory of the portfolio, but solely on the current tradeoff between satisfying the constraints and maximizing the objective. As emphasized by Boyd *et al.* (2014), there are computational advantages to using MPC in cases when estimates of future return statistics are updated online. In this case, the expected returns $\bar{r}_t$ are simply replaced with the most recent return estimates.

MPC for stochastic systems is a suboptimal control strategy, however, it uses new information advantageously and is better than pure open-loop control (Herzog *et al.* 2007). The open-loop policy would be to execute the entire sequence of trades $u_t^\star, \ldots, u_{T-1}^\star$ based on the initial portfolio without recourse. Using Monte Carlo simulation, Boyd *et al.* (2014) showed that, in any practical sense, the MPC policy is optimal.

Truncated MPC. The MPC policy described in (24) plans a sequence of trades for the full time interval $t, \ldots, T$. A common variation is to look a limited number of steps, K, into the future. At each time t the optimization problem is

$$
\begin{aligned}
\text{maximize} \quad & V_{t+K}^{\text{term}} \left(h_{t+K} \right) - \sum_{\tau=t}^{t+K-1} \psi_\tau \left(h_\tau, u_\tau \right) \\
\text{subject to} \quad & h_{\tau+1} = \left(\mathbf{1} + \hat{r}_{\tau+1} \right) \circ \left(h_\tau + u_\tau \right), \quad \tau = t, \ldots, t+K-1
\end{aligned}
\tag{25}
$$

Algorithm 1 MPC approach to dynamic portfolio optimization

(i) Update HMM parameters based on the most recent returns
(ii) Forecast the mean and variance K steps into the future
(iii) Compute an optimal sequence of trades $u_t^\star, \ldots, u_{t+K-1}^\star$ based on the current portfolio
(iv) Execute the first trade, $u_t^\star$, in the sequence and return to step i

---

with variables $h_{t+1}, \ldots, h_{t+K}$ and $u_t, \ldots, u_{t+K-1}$. Here K is the number of steps of look-ahead and V_{t+K}^{term} is the terminal value.

If the terminal value is appropriately chosen, then the truncated MPC policy is exactly the same as the full look-ahead policy. If K is large enough to reach the stationary distribution of the underlying Markov chain, then $V_{t+K}^{\text{term}}(h_{t+K})$ can be replaced by $V_{t+K} = \mathbf{1}^{\mathrm{T}} h_{t+K}$, since the risk–return tradeoff does not change after this point. If transaction costs are very high, then the choice of K can affect the result, even when K is large enough to reach the stationary distribution of the underlying Markov chain. The choice, however, should still reflect how far ahead in time it is meaningful to make predictions.

Algorithm 1 summarizes the four steps in the MPC approach to solving the dynamic portfolio optimization problem. At time t, a new measurement is obtained, which is used to update the parameters of the HMM and forecast the mean and variance K steps into the future. The next step is to compute an optimal sequence of trades $u_t^\star, \ldots, u_{t+K-1}^\star$ based on the current portfolio and the forecasts. Only the first trade in the sequence is executed before a new measurement is obtained and the procedure is repeated. Computing the optimal sequence of trades for $K = 100$ by solving the optimization problem (25) takes less than 18 milliseconds using CVXPY (Diamond and Boyd 2016) with the open-source solver ECOS (Domahidi *et al.* 2013).

4.   Empirical Results

4.1.   *Data*

The asset universe considered consists of various major stock market indices and cash. Cash positions are assumed to be risk-free and yield zero interest, hence, the only source of performance is the stock indices. The stock indices are considered one at a time. By only considering one risky asset and one risk-free asset, correlations can be disregarded. It is natural to focus on a stock index, since portfolio risk is typically dominated by stock market risk (see, e.g., Goyal *et al.* 2015). Previous studies on RBAA have also focused on stocks and cash, sometimes in combination with bonds (Ang and Bekaert 2004, Guidolin and Timmermann 2007, Bulla *et al.* 2011, Kritzman *et al.* 2012, Nystrup *et al.* 2015a).

In the first subsections, the data analyzed is 4,943 daily log-returns of the MSCI World Total Return Index covering the period from 1997 through 2015.[1] Then in section 4.6, the analysis is repeated for S&P 500, TOPIX, DAX, FTSE, and MSCI EM. Figure 2 shows the MSCI World index and its daily log-returns over the 19-year data period. The volatility forms clusters, as large price movements tend to be followed by large price movements and vice versa, as noted by Mandelbrot (1963).[2] RBAA aims to exploit this persistence of the volatility, since risk-adjusted returns, on average, are substantially lower during turbulent periods, irrespective of the source of turbulence, as shown by Kritzman and Li (2010).

Similar to previous studies, the regime detection will focus on the log-returns of the stock indices. The observed regimes in financial markets are related to the phases of the business cycle (Campbell

---

[1]The log-returns are calculated using $r_t = \log(P_t) - \log(P_{t-1})$, where P_t is the closing price of the index on day t and log is the natural logarithm.
[2]A quantitative manifestation of this fact is that while returns themselves are uncorrelated, absolute and squared returns display a positive, significant, and slowly decaying autocorrelation function.
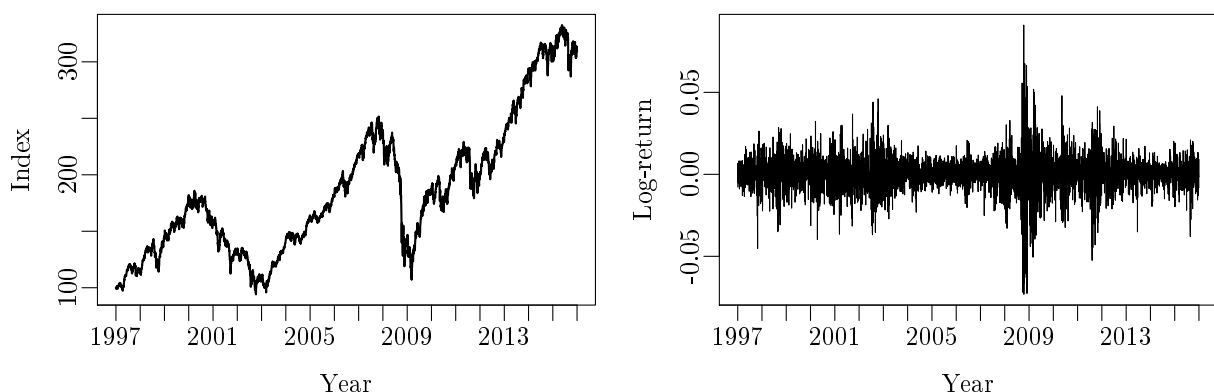
Figure 2. MSCI World Total Return Index and its daily log-returns.

1998, Cochrane 2005). As argued in Nystrup *et al.* (2015a), however, the link is complex and difficult to exploit for investment purposes due to the large lag in the availability of data related to the business cycle. Besides, stock markets generally lead the economy (Siegel 1991). The focus is, therefore, on readily available market data instead of attempting to establish the link to the business cycle.

4.2. HMM Parameters

Figure 3 shows the result of applying the adaptive estimator (7) on the daily log-returns of the MSCI World index to estimate the parameters of a two-state Gaussian HMM. An effective memory length of $N_{\text{eff}} = 260$ days was used with the tuning constant $A = 1/N_{\text{eff}}$. The first 260 observations were used for initialization.

The choice of memory length affects the parameter estimates and can be viewed as a tradeoff between bias and variance. A shorter window yields a faster adaption to changes, but larger variance of the estimates, as fewer observations are included in the estimation. In Nystrup *et al.* (2016b), an effective memory length of about one year was found to give the best forecasts. In agreement with this, forecasts based on a memory length of 260 days are found to give good results when used as inputs for MPC of an investment portfolio.

Similar to the finding in previous studies (Rydén *et al.* 1998, Bulla 2011, Nystrup *et al.* 2016b), the HMM parameters are seen to fluctuate a lot over the 19-year data period. State one is the most persistent, has the lowest variance, and—most of the time—has a positive mean. State two has a much higher variance and a negative mean value most of the time. The period in the early 2000s, when the mean return in both states was negative, exposes a weakness to having pre-defined rules for changing the portfolio based on the regime; regardless of the regime, an allocation to stocks would lose money in those years.

The probabilities of staying in the current state, γ_{11} and γ_{22}, appear to be less than 0.99 the majority of the time. A probability of 0.99 would imply an expected sojourn time of $1/(1-0.99) = 100$ days. The number of steps of look-ahead, K, is, therefore, chosen to be 100.[1] This is found to be a sufficiently large number in that a further increase does not affect the results.

4.3. Optimal Thresholds for Changing Allocation

The optimal thresholds for changing allocation depend on the specific parameter values. Using MPC, the current parameter values are taken into account when deciding whether to change the

---

[1]The number of steps of look-ahead, K, could also be chosen based on the maximum absolute value of the second largest eigenvalue of the transition probability matrix, $|\lambda_2| = |1 - \gamma_{11} - \gamma_{22}|$, which is just below 0.99.
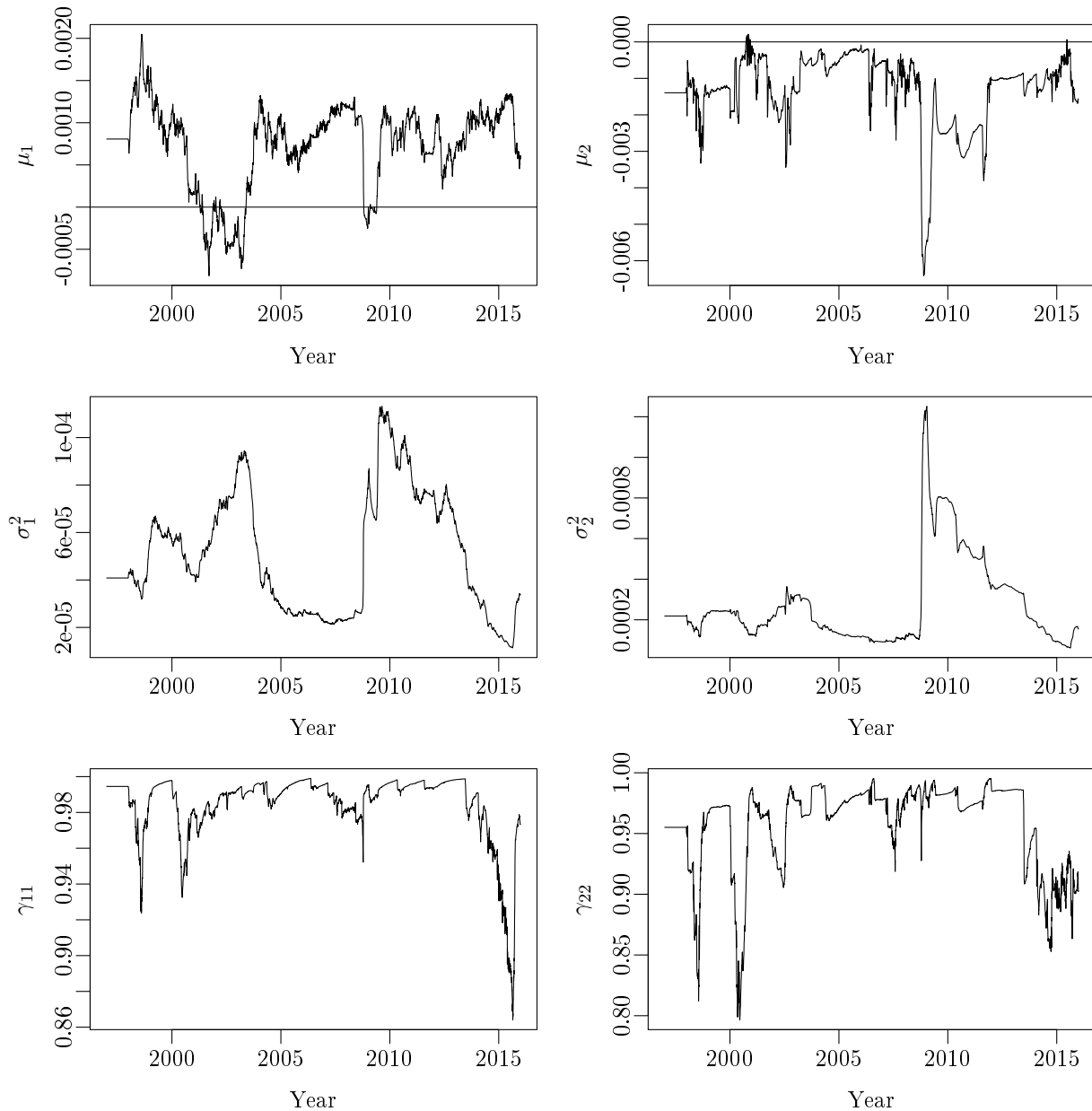
Figure 3. Parameters of a two-state Gaussian HMM estimated adaptively using an effective memory length of $N_{\mathrm{eff}} = 260$ days.

portfolio, instead of having a static decision rule for changing the allocation based on the inferred regime. As an example, figure 4 shows the optimal thresholds for changing allocation for different levels of risk aversion and transaction costs for a long-only portfolio based on the parameter values on January 2, 2012. The thresholds are not necessarily optimal ex post.

Based on the parameter values on January 2, 2012, a risk-neutral investor with $\gamma = 0$ that can trade at zero cost should be fully invested in stocks when the probability of currently being in state one (the state with low variance and positive mean) exceeds 0.55. Whenever the probability falls below 0.55, the risk-neutral investor should sell all stocks. The risk-neutral investor is always fully allocated to either stocks or cash.

A risk-averse investor with $\gamma = 2$ should be fully invested in stocks when the probability of currently being in state one exceeds 0.86 and—similar to a risk-neutral investor—fully allocated to cash when the probability is below 0.55. In between, the portfolio is a mix of stocks and cash.

A risk-neutral investor that can trade stocks at a cost of $\kappa = 0.02$ per transaction should buy
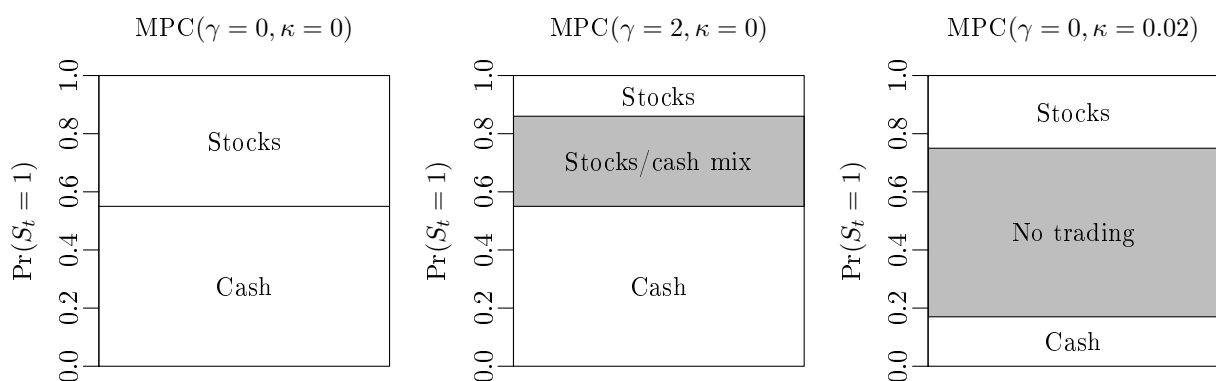
Figure 4. Optimal thresholds for changing allocation for different levels of risk aversion, γ, and transaction costs, κ, based on the parameter values on January 2, 2012.

stocks when the probability of currently being in state one exceeds 0.75 and sell all stocks when the probability falls below 0.17. In between, the risk-neutral investor should keep the initial portfolio. The thresholds are not symmetric, as $0.75 \neq 1 - 0.17$.

Institutional investors can trade stocks at a much lower cost than 2%, but this high value clearly illustrates that the presence of transaction costs leads to a no-trade zone. Based on the parameter values at the end of 2015, for example, a risk-neutral investor that can trade at a cost of 2% should never sell stocks, even if the probability of currently being in the good state is zero, because of the low persistence of the bad state and the not excessively negative mean value. The higher the transaction costs, the larger the no-trade zone. Dai *et al.* (2010b) derived similar results in a continuous-time framework.

4.4. *Comparison of MPC Results*

Table 1 summarizes the performance of MPC with and without risk (22) and trading (23) aversion for $K = 100$ steps of look-ahead. Short positions were not allowed. Transaction costs of $\kappa = 0.001$ (10 basis points per transaction) have been deducted in all three cases. This is a realistic cost for an institutional investor.

The first approach, MPC($\gamma = 0, \rho = 0$), is pure return maximization with no risk or trading penalty (except for the 10 basis point transaction cost). This approach yields the highest annualized return (AR) despite having an annual turnover (AT) of 4.61. An AT of 4.61 means that the entire portfolio is shifted from stocks to cash or vice versa, on average, 4.61 times per year.

Introducing a risk penalty by setting $\gamma = 2$ increases the AT from 4.61 to 5.30 and leads to a lower standard deviation (SD), lower maximum drawdown (MDD)[1], and lower AR. The Calmar ratio (CR)[2] is unaffected. To ensure that the performance comparison is not distorted by autocorrelation in the daily returns, the reported SDs have been adjusted for autocorrelation using the procedure outlined by Kinlaw *et al.* (2015). As γ increases and more emphasis is put on the variance forecast, the Sharpe ratio (SR)[3] deteriorates. This could indicate that the mean changes faster than the variance. The variance is crucial when distinguishing between market regimes (Nystrup *et al.* 2016a), however, basing the allocation decision on both the mean and variance forecasts adds another source of estimation error. This should be further explored in a future study encompassing more assets.

The third approach, MPC($\gamma = 0, \rho = 0.02$), is return maximization with a 2% trading penalty on top of the 10 basis point transaction cost. The trading penalty (23) with $\rho = 0.02$ is included in the objective function when determining the optimal sequence of trades in step iii of algorithm 1,

---

[1] The maximum drawdown is the largest relative decline from a historical peak in the index value.
[2] The Calmar ratio is the annualized return divided by the maximum drawdown.
[3] The Sharpe ratio is the annualized return divided by the standard deviation adjusted for autocorrelation.

Table 1.   Performance of MPC with and without risk and trading aversion.

	MPC($\gamma = 0, \rho = 0$)	MPC($\gamma = 2, \rho = 0$)	MPC($\gamma = 0, \rho = 0.02$)
Annualized return	0.076	0.062	0.070
Standard deviation	0.12	0.10	0.11
Sharpe ratio	0.65	0.61	0.67
Maximum drawdown	0.26	0.21	0.23
Calmar ratio	0.29	0.29	0.31
Annual turnover	4.61	5.30	1.17

Notes: Transaction costs of $\kappa = 0.001$ per transaction have been deducted. γ is the risk-aversion parameter and ρ is the trading-aversion parameter. Short positions were not allowed.
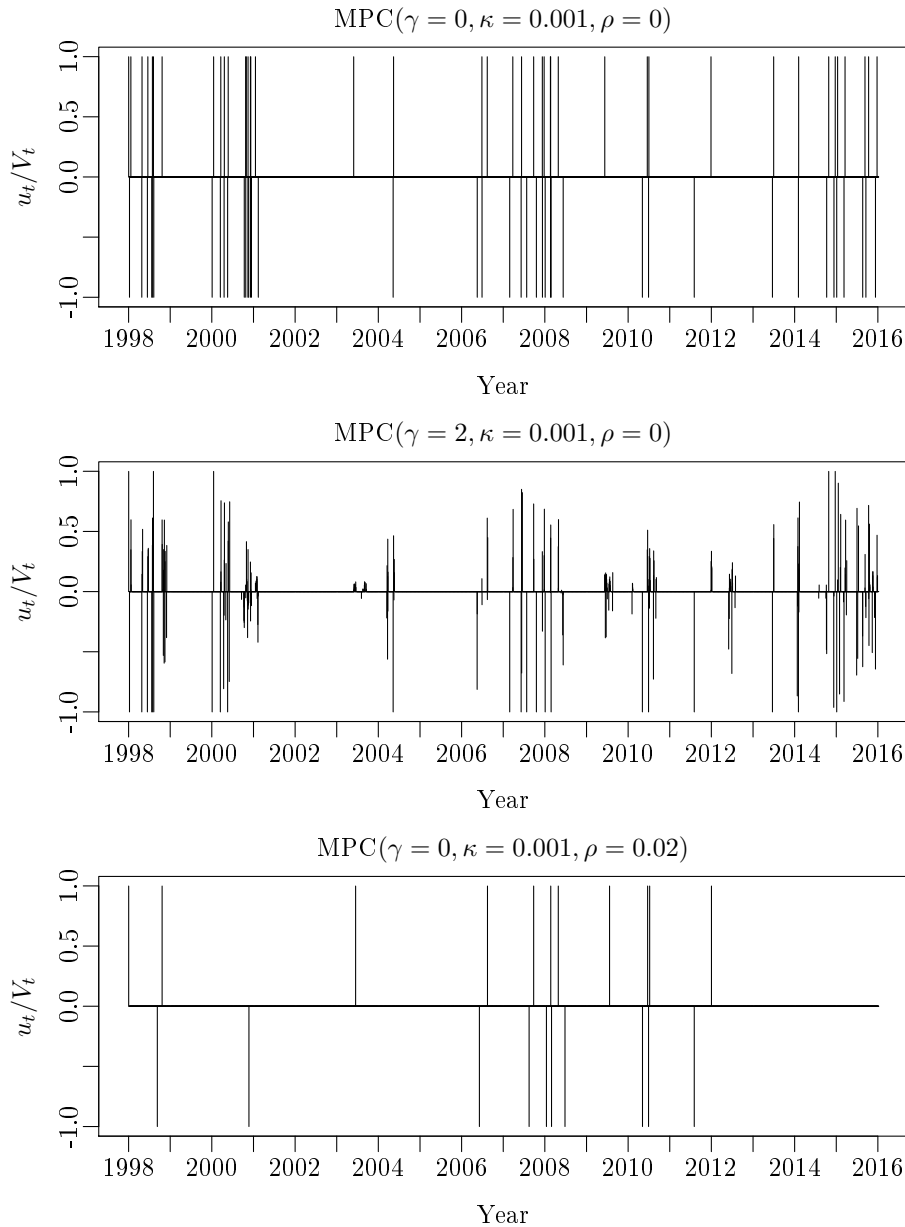

Figure 5. Trades in the MSCI World index relative to the portfolio value at the time based on MPC with and without risk and trading aversion.
Notes: Transaction costs of $\kappa = 0.001$ per transaction have been deducted. γ is the risk-aversion parameter and ρ is the trading-aversion parameter. Short positions were not allowed.

Table 2. Performance of MPC with and without a trading penalty compared to the MSCI World index, rule-based RBAA, and SAA, when allocation changes are subject to a one-day delay.

	MPC($\gamma = 0, \rho = 0$)	MPC($\gamma = 0, \rho = 0.02$)	MSCI World	RBAA	SAA
Annualized return	0.065	0.069	0.056	0.046	0.041
Standard deviation	0.12	0.11	0.18	0.12	0.13
Sharpe ratio	0.56	0.63	0.30	0.38	0.32
Maximum drawdown	0.26	0.25	0.57	0.34	0.44
Calmar ratio	0.25	0.28	0.10	0.13	0.09
Annual turnover	4.16	1.17	0.00	1.94	0.09

Notes: The SAA portfolio is rebalanced monthly to a fixed allocation of 69% stocks and 31% cash, which equals the average allocation of MPC($\gamma = 0, \rho = 0$). Transaction costs of $\kappa = 0.001$ per transaction have been deducted. γ is the risk-aversion parameter and ρ is the trading-aversion parameter. Short positions were not allowed.

but the actual transaction cost applied is still $\kappa = 0.001$. This subjective trading penalty leads to a significantly lower AT and a slightly lower AR compared to the unpenalized case, while the SR and CR are roughly unchanged.

Figure 5 shows the transactions in the MSCI World index relative to the portfolio value at the time for the three approaches. Introducing a risk penalty by setting $\gamma = 2$ leads to more frequent trading and a higher AT compared to pure return maximization. A trading penalty, on the other hand, leads to significantly fewer trades. The trading penalty appears to be effective at reducing the number of trades that are reversed within a short timespan and may, therefore, be preferred in some applications.

4.5.   *Comparison with Rule-Based Approach when Allocation Changes are Delayed*

The results presented in the previous subsection are based on the assumption that it is possible to trade at the closing price after it is known and the parameters and forecasts have been updated. It is often more realistic to assume that allocation changes cannot be implemented until the end of the next day, as illustrated in figure 6. To ensure that the long-only constraint is still satisfied at all times, trading decisions have to be implemented as fractions of the asset holding.[1]

Figure 6.   Timeline of portfolio dynamics when trades are delayed by one day.

In table 2, the performance of the risk-neutral MPC approach with and without a trading penalty is reported when allocation changes are subject to a one-day delay. Transaction costs of 10 basis points have been deducted from the reported results. The AR of MPC with no risk or trading penalty is about one percentage point lower when allocation changes are delayed, in spite of the AT being lowered from 4.61 to 4.16.

Imposing a trading penalty actually increases the AR when allocation changes are subject to a one-day delay. The AR, SR, and CR of the risk-neutral MPC approach with trading aversion are almost unchanged compared to when there is no delay. This suggest that a trading penalty increases the robustness of the MPC approach, similarly to what it does in a single-period setting (Brodie *et al.* 2009, Ho *et al.* 2015). The delay has no impact on the AT of the penalized MPC approach that is still substantially lower than for the unpenalized approach.

In table 2, the performance of the risk-neutral MPC approaches is compared with the MSCI World index, a rule-based RBAA approach, and an SAA portfolio that is rebalanced monthly to a
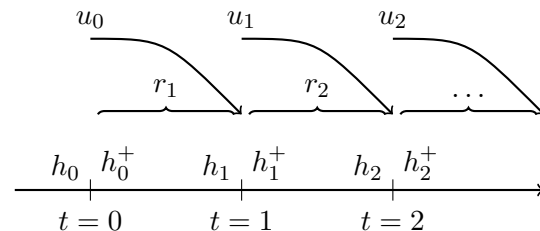
---

[1]If a decision is made on day t to sell \$80 worth of stocks out of a total holding of \$100, then 80% of the stocks are sold on day $t + 1$, regardless of their value.
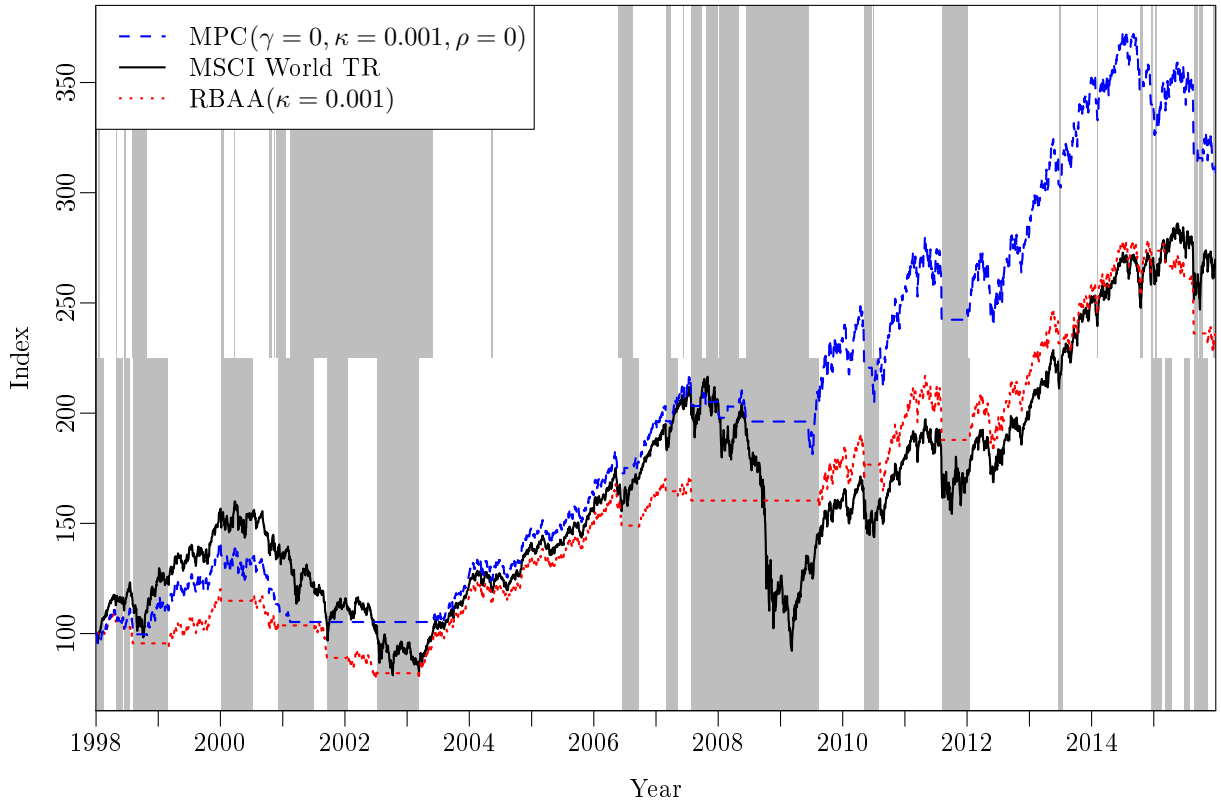
Figure 7. Performance of the return-maximizing MPC approach compared with rule-based RBAA and the MSCI World index. In the shaded periods, the MPC and the RBAA portfolios (top and bottom half, respectively) were fully allocated to cash.
Notes: Transaction costs of $\kappa = 0.001$ per transaction have been deducted. γ is the risk-aversion parameter and ρ is the trading-aversion parameter. Short positions were not allowed.

fixed allocation of 69% stocks and 31% cash, which equals the average allocation of the unpenalized MPC approach.

The rule-based RBAA approach is the same as in Nystrup *et al.* (2017). Like the risk-neutral MPC approaches, it is either fully allocated to stocks or cash. The allocation is changed when the probability that a regime change has occurred exceeds a threshold of 0.9998. The underlying HMM is estimated using (7) with an effective memory length of two years, since this was found to give better results. A similar memory length was used in Nystrup *et al.* (2015a, 2017).

The risk-neutral MPC approaches have realized the highest AR and have outperformed the MSCI World index that has a significantly higher SD and MDD. This is under the assumption that cash positions yield zero interest. Further, with approximately the same SD and a lower MDD than the RBAA and SAA portfolios, the MPC approaches have realized a substantially higher SR and CR.

RBAA outperforms both SAA and the index in terms of SR and CR. Its AT of 1.94 is higher than that of the penalized MPC approach, but less than half of that of the unpenalized MPC approach. The performance of RBAA relative to the index is not as convincing as when there are other alternatives to invest in than zero-interest cash (see Nystrup *et al.* 2015a, 2017).

In figure 7, the performance of MPC with no trading or risk penalty is compared to the rule-based RBAA strategy and the MSCI World index when allocation changes are subject to a one-day delay. In the shaded periods, the MPC (top half) and the RBAA portfolios (bottom half) were fully allocated to cash. The allocations are different from what would be expected if the regimes were based on a business cycle indicator.

The MPC portfolio performs better than the RBAA portfolio during the build-up and burst of the dot-com bubble around the year 2000. It is fully allocated to stocks most of the time leading

Table 3.   Performance of MPC with no risk or trading penalty when applied to various major stock indices with no delay in allocation changes.

	MSCI World	S&P500	TOPIX	DAX	FTSE	MSCI EM
Annualized return	0.076 (0.056)	0.110 (0.108)	0.044 (0.025)	0.066 (0.073)	0.051 (0.078)	0.119 (0.072)
Standard deviation	0.12 (0.18)	0.14 (0.16)	0.17 (0.24)	0.16 (0.22)	0.14 (0.15)	0.18 (0.27)
Sharpe ratio	0.65 (0.30)	0.79 (0.68)	0.26 (0.10)	0.41 (0.33)	0.36 (0.54)	0.66 (0.27)
Maximum drawdown	0.26 (0.57)	0.36 (0.55)	0.48 (0.72)	0.39 (0.73)	0.36 (0.48)	0.38 (0.65)
Calmar ratio	0.29 (0.10)	0.30 (0.20)	0.09 (0.03)	0.17 (0.10)	0.14 (0.16)	0.31 (0.11)
Annual turnover	4.61 (0.00)	3.29 (0.00)	4.69 (0.00)	5.89 (0.00)	4.59 (0.00)	8.18 (0.00)

Notes: The numbers in parentheses are the summary statistics for a buy-and-hold investment in the respective indices. Transaction costs of $\kappa = 0.001$ per transaction have been deducted. Short positions were not allowed.

up to the peak and fully allocated to cash throughout the downturn. The MPC portfolio stays fully allocated to cash throughout the downturn, because the mean value in both states is negative in this period, cf. figure 3. The RBAA portfolio times the subsequent rebound better and does well by staying fully allocated to cash throughout the crash in 2008. The MPC portfolio times the rebound in 2009 better and gradually extends its lead over the following years.

The MPC portfolio is slightly behind the MSCI World index at the peak in year 2000, but then moves ahead of the index during the downturn. The lead is maintained in the following years leading up to the crash in 2008, during which the lead is significantly extended. Part of the lead is lost during the market rebound in the first half of 2009, before the MPC allocation is shifted to stocks. At the end of the sample, the performance gap is substantial. The outperformance relative to the index comes from the two major downturns, but this is hardly surprising, since there is no other source of return than the index itself. In risk-adjusted terms, the outperformance is conclusive.

4.6.   *Application to Other Indices*

Table 3 summarizes the results from applying MPC with no risk or trading penalty (i.e., MPC($\gamma = 0, \rho = 0$)) to various major stock market indices. For the MSCI World index, the numbers are the same that were reported in table 1. Recall that the testing period for the MSCI World index spans 1998 through 2015. For S&P 500, TOPIX, DAX, and FTSE, the data period includes 1984 through 2015. The first two years are used for initialization, leaving 30 years for testing. For the MSCI EM index, daily data is only available from 1988 and onwards, thus the testing period is four years shorter. All indices are net total return.

The MPC approach realizes a higher SR and CR than a buy-and-hold investment (as summarized in parentheses) in five out of six indices, with FTSE being the only exception. The best performance relative to the underlying index is obtained for MSCI World, TOPIX, and MSCI EM. For all indices except DAX and FTSE, the AR of the MPC approach is higher than that of the underlying index despite high ATs—which could easily be reduced by introducing a trading penalty. In all six cases, the SD and MDD are lower than those of the underlying index. The results in table 3 show that the MPC approach in combination with the adaptively estimated HMM has worked well for multiple major stock indices across different time periods. By introducing a trading penalty and calibrating its level to each index individually, it would be possible to further improve the results.

5.   Conclusion

This article has shown the strength of using MPC for dynamic portfolio optimization in combination with an online method for forecasting the mean and variance of financial returns. There are computational advantages to using MPC in cases when estimates of future return statistics are updated every time a new observation becomes available, since the optimal control actions are reconsidered anyway.

Based on forecasts from an adaptively-estimated HMM, the MPC approach realized a higher

return and a significantly lower risk than a buy-and-hold investment in various major stock indices. This was after accounting for transaction costs. Imposing an additional trading penalty increased the robustness, by reducing the number of trades, and improved the performance, when allocation changes were subject to a delay.

MPC also outperformed RBAA based on a static decision rule for changing the portfolio. The performance of rule-based RBAA has been stronger in previous studies, where there were more investment opportunities than stocks and zero-interest cash. Thus, there is potential for using MPC for optimal control of multi-asset portfolios.

To keep things simple and illustrate the strength of the approach, the focus of this article was on stocks and cash, but it naturally extends to a multi-asset portfolio. Another possibility for future work would be to specify a model for the parameter changes, possibly including relevant explanatory variables, in an attempt to improve the forecasts and take the stochasticity into account in the portfolio optimization.

Acknowledgment

Funding

References

Ang, A. and Bekaert, G., How regimes affect asset allocation. *Financial Analysts Journal*, 2004, **60**, 86–99.

Ang, A. and Timmermann, A., Regime changes and financial markets. *Annual Review of Financial Economics*, 2012, **4**, 313–337.

Baum, L.E., Petrie, T., Soules, G. and Weiss, N., A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *Annals of Mathematical Statistics*, 1970, **41**, 164–171.

Bemporad, A., Model predictive control design: New trends and tools. In *Proceedings of the 45th IEEE Conference on Decision and Control*, pp. 6678–6683, 2006.

Boyd, S., Mueller, M.T., O'Donoghue, B. and Wang, Y., Performance bounds and suboptimal policies for multi-period investment. *Foundations and Trends in Optimization*, 2014, **1**, 1–72.

Boyd, S. and Vandenberghe, L., *Convex Optimization*, 2004, Cambridge University Press, New York.

Brennan, M.J., Schwartz, E.S. and Lagnado, R., Strategic asset allocation. *Journal of Economic Dynamics and Control*, 1997, **21**, 1377–1403.

Brodie, J., Daubechies, I., Mol, C.D., Giannone, D. and Loris, I., Sparse and stable Markowitz portfolios. *Proceedings of the National Academy of Sciences of the United States of America*, 2009, **106**, 12267–12272.

Bulla, J., Hidden Markov models with t components. Increased persistence and other aspects. *Quantitative Finance*, 2011, **11**, 459–475.

Bulla, J. and Bulla, I., Stylized facts of financial time series and hidden semi-Markov models. *Computational Statistics & Data Analysis*, 2006, **51**, 2192–2209.

Bulla, J., Mergner, S., Bulla, I., Sesboüé, A. and Chesneau, C., Markov-switching asset allocation: Do profitable strategies exist?. *Journal of Asset Management*, 2011, **12**, 310–321.

Calafiore, G.C., Multi-period portfolio optimization with linear control policies. *Automatica*, 2008, **44**, 2463–2473.

Calafiore, G.C., An affine control method for optimal dynamic asset allocation with transaction costs. *SIAM Journal on Control and Optimization*, 2009, **48**, 2254–2274.

Campbell, J.Y., Asset prices, consumption, and the business cycle. Working Paper 6485, National Bureau of Economic Research, 1998.

Cappé, O., Moulines, E. and Rydén, T., *Inference in Hidden Markov Models*, 2005, Springer: New York.

Cochrane, J., Financial markets and the real economy. Working Paper 11193, National Bureau of Economic Research, 2005.

Costa, O.L. and Araujo, M.V., A generalized multi-period mean–variance portfolio optimization with Markov switching parameters. *Automatica*, 2008, **44**, 2487–2497.

Creal, D., Koopman, S.J. and Lucas, A., Generalized autoregressive score models with applications. *Journal of Applied Econometrics*, 2013, **28**, 777–795.

Dai, M., Xu, Z.Q. and Zhou, X.Y., Continuous-time Markowitz's model with transaction costs. *SIAM Journal on Financial Mathematics*, 2010a, **1**, 96–125.

Dai, M., Zhang, Q. and Zhu, Q.J., Trend following trading under a regime switching model. *SIAM Journal on Financial Mathematics*, 2010b, **1**, 780–810.

DeMiguel, V., Garlappi, L. and Uppal, R., Optimal versus naive diversification: How inefficient is the $1/N$ portfolio strategy?. *Review of Financial Studies*, 2009, **22**, 1915–1953.

Dempster, A.P., Laird, N.M. and Rubin, D.B., Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 1977, **39**, 1–38.

Diamond, S. and Boyd, S., CVXPY: A Python-embedded modeling language for convex optimization. *Journal of Machine Learning Research*, 2016, **17**, 1–5.

Domahidi, A., Chu, E. and Boyd, S., ECOS: An SOCP solver for embedded systems. In *Proceedings of the European Control Conference*, pp. 3071–3076, 2013.

Falcone, P., Borrelli, F., Asgari, J., Tseng, H.E. and Hrovat, D., Predictive active steering control for autonomous vehicle systems. *IEEE Transactions on control systems technology*, 2007, **15**, 566–580.

Frühwirth-Schnatter, S., *Finite Mixture and Markov Switching Models*, 2006, Springer: New York.

Goyal, A., Ilmanen, A. and Kabiller, D., Bad habits and good practices. *Journal of Portfolio Management*, 2015, **41**, 97–107.

Guidolin, M. and Timmermann, A., Asset allocation under multivariate regime switching. *Journal of Economic Dynamics and Control*, 2007, **31**, 3503–3544.

Herzog, F., Dondi, G. and Geering, H.P., Stochastic model predictive control and portfolio optimization. *International Journal of Theoretical and Applied Finance*, 2007, **10**, 203–233.

Ho, M., Sun, Z. and Xin, J., Weighted elastic net penalized mean–variance portfolio design and computation. *SIAM Journal on Financial Mathematics*, 2015, **6**, 1220–1244.

Khreich, W., Granger, E., Miri, A. and Sabourin, R., A survey of techniques for incremental learning of HMM parameters. *Information Sciences*, 2012, **197**, 105–130.

Kinlaw, W., Kritzman, M. and Turkington, D., The divergence of high- and low-frequency estimation: Implications for performance measurement. *Journal of Portfolio Management*, 2015, **41**, 14–21.

Kritzman, M. and Li, Y., Skulls, financial turbulence, and risk management. *Financial Analysts Journal*, 2010, **66**, 30–41.

Kritzman, M., Page, S. and Turkington, D., Regime shifts: Implications for dynamic strategies. *Financial Analysts Journal*, 2012, **68**, 22–39.

Kulhavý, R. and Zarrop, M.B., On a general concept of forgetting. *International Journal of Control*, 1993, **58**, 905–924.

Lystig, T.C. and Hughes, J.P., Exact computation of the observed information matrix for hidden Markov models. *Journal of Computational and Graphical Statistics*, 2002, **11**, 678–689.

Mandelbrot, B., The variation of certain speculative prices. *Journal of Business*, 1963, **36**, 394–419.

Markowitz, H.M., Portfolio selection. *Journal of Finance*, 1952, **7**, 77–91.

Meucci, A., *Risk and Asset Allocation*, 2005, Springer: Berlin.

Michaud, R.O., The Markowitz optimization Enigma: Is 'optimized' optimal?. *Financial Analysts Journal*, 1989, **45**, 31–42.

Nystrup, P., Hansen, B.W., Larsen, H.O., Madsen, H. and Lindström, E., Dynamic allocation or diversification: A regime-based approach to multiple assets. *Journal of Portfolio Management*, 2017, **44**, forthcoming.

Nystrup, P., Hansen, B.W., Madsen, H. and Lindström, E., Regime-based versus static asset allocation: Letting the data speak. *Journal of Portfolio Management*, 2015a, **42**, 103–109.

Nystrup, P., Hansen, B.W., Madsen, H. and Lindström, E., Detecting change points in VIX and S&P 500: A new approach to dynamic asset allocation. *Journal of Asset Management*, 2016a, **17**, 361–374.

Nystrup, P., Madsen, H. and Lindström, E., Stylised facts of financial time series and hidden Markov models in continuous time. *Quantitative Finance*, 2015b, **15**, 1531–1541.

Nystrup, P., Madsen, H. and Lindström, E., Long memory of financial time series and hidden Markov models with time-varying parameters. *Journal of Forecasting*, 2016b.

Parkum, J.E., Poulsen, N.K. and Holst, J., Recursive forgetting algorithms. *International Journal of Control*, 1992, **55**, 109–128.

Rydén, T., Teräsvirta, T. and Åsbrink, S., Stylized facts of daily return series and the hidden Markov model. *Journal of Applied Econometrics*, 1998, **13**, 217–244.

Sheikh, A.Z. and Sun, J., Regime change: Implications of macroeconomic shifts on asset class and portfolio performance. *Journal of Investing*, 2012, **21**, 36–54.

Siegel, J.J., Does it pay stock investors to forecast the business cycle?. *Journal of Portfolio Management*, 1991, **18**, 27–34.

Tibshirani, R., Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 1996, **58**, 267–288.

von Neumann, J. and Morgenstern, O., *Theory of Games and Economic Behavior*, 3rd , 1953, Princeton University Press, Princeton.

Whittle, P., Risk-sensitive linear/quadratic/Gaussian control. *Advances in Applied Probability*, 1981, **13**, 764–777.