J Math Model Algor (2014) 13:59–85
DOI 10.1007/s10852-012-9214-4

A Higher-Order Hidden Markov Chain-Modulated Model for Asset Allocation

Xiaojing Xi · Rogemar Mamon · Matt Davison

Received: 21 January 2012 / Accepted: 30 October 2012 / Published online: 24 November 2012

Abstract This paper presents an analysis of asset allocation strategies when the asset returns are governed by a discrete-time higher-order hidden Markov model (HOHMM), also called the weak hidden Markov model. We assume the drifts and volatilities of the asset returns switch over time according to the state of the HOHMM, in which the probability of the current state depends on the information from previous time-steps. The "switching" and "mixed" strategies are studied. We use a multivariate filtering technique in conjunction with the EM algorithm to obtain estimates of model parameter at a given time. This, in turn, aids investors in determining the optimal investment strategy for the next time step. Numerical implementation is applied to data on Russell 3000 value and growth indices. We benchmark the respective performances of portfolio using three classical investment measures.

Keywords Markov chain · Regime-switching · Filtering · Investment strategy · Portfolio performance

1 Introduction

It is well documented in the asset allocation literature that the inclusion of market regime-switching dynamics has considerable impact on the optimal portfolio strategy

X. Xi · R. Mamon · M. Davison
Department of Applied Mathematics, University of Western Ontario, London, ON, Canada

R. Mamon (✉) · M. Davison
Department of Statistical and Actuarial Sciences, University of Western Ontario,
1151 Richmond Street, London, ON, Canada N6A 5B7
e-mail: rmamon@stats.uwo.ca

M. Davison
Richard Ivey School of Business, University of Western Ontario, London, ON, Canada

of individual investors. In financial portfolio management, the portfolio risk cannot be entirely eliminated although it can be controlled with an optimal asset allocation strategy combining different types and amount of investment. Each investment has its own unique risk and return characteristics. A well-designed investment aims at the maximisation of expected portfolio return whilst controlling the level of risk. A fundamental example of the single-period mean-variance asset allocation problem is given by the Nobel prize-winning work of Markowitz [24], in which the variance is employed as a measure of risk and the efficient allocation of wealth amongst different investment classes is provided. Since practical asset allocation problems involve inter-temporal decisions, Samelson [30] and Merton [26] considered asset allocation problem in a multi-period model and in a continuous-time model, respectively. In particular, Merton utilised stochastic optimal control theory to derive a closed-form solution for an asset allocation strategy under certain assumptions. A key assumption of these early works in the literature is that the dynamics of asset returns are linear processes with constant coefficients. However, the state of the economy and the financial market varies randomly over time. Investors are concerned with regime-switching uncertainty affecting the portfolio return. Such uncertainty affects the future payoffs and therefore could alter the optimal asset allocation. Ang and Bekaert [3] introduced a regime-switching model with time-varying correlations and volatilities for asset allocation. They reported evidence of shifting regimes in the US, UK and German equity markets. In a subsequent study, Ang and Bekaert [4] expanded the lists of markets and assets in the investigation of optimal asset allocation under a regime-switching framework. Their out-of-sample test shows that the regime-switching strategy dominates a non-regime dependent strategy. Bauer et al. [5] observed the tendency of changing correlation and volatility amongst assets, and considered a regime-switching technique for portfolio optimisation. An out-of-sample backtesting was applied on a six-asset portfolio consisting of equities, bonds, commodities and real estate. The results demonstrated a significant information gain from using a regime-switching strategy. Guidolin and Timmermann [18] considered asset allocation decisions under a regime-switching model for asset returns with four separate regimes. It was found that the optimal allocations vary considerably across these states and change over time as investors revise their estimate of the state probabilities.

Various works that support model assumptions in which parameters change over time in accordance with the evolution of an unobserved Markov chain have been proposed. Elliott and van der Hoek [15] put forward a model for the rates of asset returns driven by a Markov chain in discrete time. Their work features filtering and prediction techniques in the model identification and outlines how their method could be applied to the asset allocation problem using mean-variance type utility criterion. Graflund and Nilsson [17] investigated dynamic portfolio selection within a Markovian switching framework. Their results highlight the economic importance of regimes and suggest that ignoring the regime will require significant compensation. In the study of Ammann and Verhofen [2], Markov Chain Monte Carlo methods were applied to estimate a multivariate regime-switching model. Two clearly separable regimes characterised by different mean returns, volatilities and correlations were found. The results of their out-of-sample backtest suggests that the buy-and-hold strategy based on regime-switching model can be profitable. Bulla et al. [7] focused on daily stock market return series at five major regional markets over the last four decades. They presented an out-of sample performance analysis with transaction costs taken into account and concluded that the strategy is improved by considering Markovian switching model. Erlwein et al. [16] developed and compared investment strategies in allocating funds to either growth or value stocks, whose price dynamics are driven by a hidden Markov model (HMM). Their investigation shows that the HMM-based strategies are more stable and outperform the pure growth strategy in terms of higher Sharpe ratios and lower variance of the performance. Elliott et al. [14] considered a mean-variance portfolio selection problem where the appreciation rate of the risky asset is modulated by a continuous-time Markov chain. They employed the gauge transformation technique to obtain robust filters and developed the filter-based Expectation Maximisation (EM) algorithm in calculating the estimates of the unknown parameters. An explicit solution to the mean-variance portfolio problem is derived using the filtering results.

The original form of HMM is now in widespread use because of its ability to capture the switching of market or economic states. However, there is growing evidence that many financial series have longer memories than may be captured by original form HMM models. Some even suggest that financial time series are of the long memory form pioneered by Hurst [20]. Lobato and Savin [21] and Ray and Tsay [28] found evidence of longer-range dependence in the volatility of S&P 500 returns. McCarthy et al. [25] found evidence of long memory in percentage changes on Treasury debt security yields. However, Couillard and Davison [10] noted that caution must be taken in applying statistical tests for longer-range Hurst-like dependence. Some effort has been made to develop long memory financial models. For instance, Dajcman [11] examined a time varying long memory parameter for eight European stock market returns by using an auto-regressive fractionally integrated moving average model. Others suggest that much of this behaviour can be explained using shorter memory models, an approach that this paper also takes. Maheu [23] suggested that GARCH models can in some circumstances account for the long memory property found in financial market volatility.

There is a vast literature devoted to modelling longer-range dependence using single-state stochastic models. Unfortunately, Rydén et al. [29] suggested that an ordinary HMM cannot describe the stylized fact of very slow decay in return autocorrelations. So some extension of the HMM literature is needed. One approach is to use a hidden semi-Markov model. For example, Yu et al. [34] developed a recursive formula to estimate the Hurst parameter corresponding to a second-order HMM. They showed that this approach could capture longer-range dependence in web server workload when the distribution of at least one state is heavy tailed. In a more financial application, Bulla and Bulla [6] explored goodness of fit of two hidden semi-Markov models to 18 pan-European sector return indices, obtaining promising results. In this the current work, we take a related approach which, although motivated by [6, 34], is slightly different. We choose to extend the single-state HMM models to involve multiple time lags, much as a simple AR(1) model can be extended to AR(q) by using multiple time lags. We call this approach the higher-order hidden Markov model (HOHMM).

In this paper, we apply an HOHMM, in particular a second-order HMM, to an asset allocation problem. Our financial motivation for this comes from the fact that practitioners often like to divide a market not only into high and low volatility states, but also into “trending” and “choppy” market states. A one-state model is unlikely to pick this up, since on average returns are serially uncorrelated because some of the time they have a positive serial correlation and at other times a negative serial correlation. The hope is that a two-state model might better capture this behaviour. Also, as mentioned by Solberg [32], the real significance of HOHMM is to establish that the Markov chain assumption is not really as restrictive as it first appears. One is not limited to a dependence on just one prior time epoch but can make the dependency extend to any finite number of prior epochs, thereby capturing more information from the past. This, in turns, widens the literature on models aiming to reflect longer-range dependence in financial models.

In the HOHMM model the transition matrix between one state and the next state is itself dependent on the information in the prior states. An nth-order Markov chain is dependent on the prior n state. The higher the order, the more extended the dependency, and therefore more information from the past can be captured. Xi and Mamon [33] proposed an HOHMM for discrete-time continuous-range observations and provided a detailed implementation of the model to a financial dataset. Hess [19] considered conditional CAPM strategies based on regime forecasts from an autoregressive Markov regime-switching behaviour with lag two. The improvement of the portfolio performance by using the proposed strategy is examined through in-sample and out-of-sample analyses. An application of higher-order Markovian switching model for risk measurement is presented by Siu et al. [31]. Other applications of HOHMM, such as in option pricing, can be found in Ching et al. [9].

In this paper, we investigate optimal investment strategies for asset allocation under a weak Markov-switching framework. In particular, we assume the log returns of risky assets are modulated by a second-order multivariate Markov chain, whose current behaviour depends on its behaviour at the previous two time steps. The states of the higher-order Markov chain are interpreted as states of the economy. Compared to the previous research conducted in [33], we extend the single-variate HOHMM to a multivariate case by modifying the Radon-Nikodým derivative. This extension allows us to investigate the application of HOHMM involving multivariate financial series, such as those occurring in asset allocation. We use the same asset allocation strategies from [16]. Compared to their research, we relax the Markov assumption by increasing the first-order HMM to second-order HMM. The filtering technique for HOHMM is implemented on updated market data, which includes the period of the subprime crisis. The numerical results show how an HOHMM captures information during a crisis period and the resulting impact on the strategy. The HOHMM has the advantage that it can capture the longer-ranging dependence of the states of the market, and therefore it is more appropriate when memories are evident in financial series. From the investors' view, tactical investment decisions require the evaluation of the expected future payoff on risky assets. More economic insights can be gained if relevant historical information can be incorporated into the unobservable market state; this will be beneficial to investors from both the economic and statistical perspectives. Although a higher-order Markov chain, more specifically a Markov chain of order higher than two, leads to more information incorporated in the HMM, the number of model parameters involved increases exponentially. Ching et al. [8] apply a higher-order multivariate HMM to a sequence of multivariate categorical data and show that an nth-order, s-variate, N-state Markov chain model requires ns^2N^2 parameters. To facilitate the dynamic estimation of this huge number of parameters, we use a transformation that converts an HOHMM into a regular HMM thereby enabling the estimation algorithm to perform smoothly. The transformation, which is essentially a mapping of states, is employed to eventually recover the required number of parameters. Our asset allocation strategies rely on the estimates of parameters and forecasted return through the mathematical techniques of HOHMMs.

This paper is organised as follows. In Section 2, we present the multidimensional HOHMM filtering and estimation techniques. HOHMM filtering procedure is applied to the Russell 3000 value and growth indices data, whose log-returns are assumed to follow a normal distribution with regime-switching dependent on two previous time epoch. The EM algorithm is then applied to obtain the online recursive estimates of the model parameters. In Section 3, we utilise the optimal estimates to forecast the two indices and conclude that a two-state HOHMM is sufficient to capture the characteristics of our data based on four error metrics. We investigate an investment strategy switching between the Russell 3000 value and growth indices. The switching decision determined by the one-step ahead forecasts return of each index. In Section 5, a mixed investment strategy is considered. The optimal weights of investment between the two indices are obtained by solving a mean-variance problem under the regime-switching setting. The estimation of the optimal weights incorporates the parameter estimates as well as the states of a higher-order Markov chain. Portfolio performance is investigated in Section 6, where we use three classical measures for benchmarking. Furthermore, a bootstrap analysis is used to compare the stability of portfolios with various level of transaction costs. Section 7 concludes the paper.

2 Filtering and Parameter Estimation

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space under which $\mathbf{x}_k$ is a Markov chain with finite-state space in discrete time $(k = 0, 1, 2 \ldots)$. To simplify the discussion and present a complete characterization of the parameter estimation, we only consider a second-order Markov chain. That is,

$$\begin{aligned} P\,(\mathbf{x}_{k+1} = x_{k+1}|\mathbf{x}_0 = x_0, \ldots, \mathbf{x}_{k-1} = x_{k-1}, \mathbf{x}_k = x_k) \\ = P(\mathbf{x}_{k+1} = x_{k+1}|\mathbf{x}_{k-1} = x_{k-1}, \mathbf{x}_k = x_k). \end{aligned}$$

Without loss of generality, the N-state higher-order Markov chain takes value from the canonical basis $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_N\} \subset \mathbf{R}^N$, where $\mathbf{e}_i$ is the vector with unity in the ith element and zero elsewhere. We interpret $\langle \mathbf{x}_k, \mathbf{e}_i \rangle$ as the event that the economy is in the state i at time k. Here $\langle \cdot, \cdot \rangle$ denotes the inner product in $\mathbf{R}^N$. The element a_{lmv}, $l, m, v \in 1, \ldots, N$, of the transition probability matrix $\mathbf{A}$ refers to the probability that the process enters state l given that the current states is the mth state and the previous state was in v.

Instead of studying the higher-order Markov chain directly, we introduce a mapping α, to embed the second-order Markov chain into the first-order Markov chain, and then apply the regular filtering method. This idea is analogous to the embedding of higher-order ODEs into a system of first-order ODEs and solving the system by regular methods, see Abell and Braslton [1]. The mapping α is defined by

$$\alpha(\mathbf{e}_r, \mathbf{e}_s) = \mathbf{e}_{rs}, \text{ for } 1 \le r, s \le N,$$

where $\mathbf{e}_{rs}$ is an $\mathbf{R}^{N^2}$−unit vector with unity in its $((r-1)N+s)th$ position. Note that

$$\langle \alpha(\mathbf{x}_k, \mathbf{x}_{k-1}), \mathbf{e}_{rs}\rangle = \langle \mathbf{x}_k, \mathbf{e}_r\rangle\langle \mathbf{x}_{k-1}, \mathbf{e}_s\rangle$$

represents the identification of the new first-order Markov chain with the canonical basis. We also define the new $N^2 \times N^2$ transition probability matrix $\mathbf{\Pi}$ of the new Markov chain by

$$\pi_{ij} = \begin{cases} a_{lmv} \text{ if } i = (l-1)N+m, \; j = (m-1)N+v \\ 0 \quad \text{ otherwise.} \end{cases}$$

Note that at time k, each non-zero element π_{ij} represents the probability

$$\pi_{ij} = a_{lmv} = P(\mathbf{x}_k = \mathbf{e}_l | \mathbf{x}_{k-1} = \mathbf{e}_m, \; \mathbf{x}_{k-2} = \mathbf{e}_v),$$

and each zero represents an impossible transition. It is known [31] that the new Markov chain $\alpha(\mathbf{x}_k, \mathbf{x}_{k-1})$ has a semi-martingale representation

$$\alpha(\mathbf{x}_k, \mathbf{x}_{k-1}) = \mathbf{\Pi}\alpha(\mathbf{x}_{k-1}, \mathbf{x}_{k-2}) + \mathbf{v}_k, \tag{1}$$

where $\{\mathbf{v}_k\}_{k\geq 1}$ is a sequence of $\mathbf{R}^{N^2}$ martingale increments. In a comprehensive monograph on HMM, MacDonald and Zucchini [22] devoted a section introducing the HOHMM and give a detailed example of transforming a second-order two-state Markov chain into a regular two-state Markov chain. An efficient recursive algorithm for computing the likelihood from consecutive observations under a second-order HMM is given. However, no application of the second-order HOHMM MLE is given in this book. Du Preez et al. [27] developed a computing algorithm to reduce any HOHMM to a corresponding first-order HMM. The algorithm is applied to language recognition. In contrast to their research objectives, we focus on financial time series applications. In particular, we obtain a reduced first-order HMM using a transformation and estimate parameters of asset price logreturns.

Let $\mathbf{y}_k = (y_k^1, y_k^2, \ldots, y_k^d)$ be a d-dimensional process. Each component y_k^g, $1 \leq g \leq d$, is the sequence of log returns of an asset price with the dynamics

$$y_{k+1}^g = f^g(\mathbf{x}_k) + \sigma^g(\mathbf{x}_k) z_{k+1}^g.$$

Here $\{z_k\}$ is a sequence of $N(0,1)$ independent, identically distributed (IID) random variables and independent of $\mathbf{x}$. The function f^g and σ^g are determined by the vectors $\mathbf{f}^g = (f_1^g, f_2^g, \ldots, f_N^g)^\top$ and $\boldsymbol{\sigma}^g = (\sigma_1^g, \sigma_2^g, \ldots, \sigma_N^g)^\top$ in $\mathbf{R}^N$, and $f^g(\mathbf{x}_k) = \langle \mathbf{f}^g, \mathbf{x}_k\rangle$ and $\sigma^g(\mathbf{x}_k) = \langle \boldsymbol{\sigma}^g, \mathbf{x}_k\rangle$ represent the mean and variance of y_k^g, respectively and $\top$ denotes the transpose of a matrix. Note that all components of the vector observation process have the same underlying higher-order Markov chain.

It must be noted that we do not observe the underlying higher-order Markov chain from the financial market directly. Under the real world measure P, the state $\mathbf{x}_k$ is contained in the noisy observations $\mathbf{y}_k$, $k \geq 1$. We aim to "filter" the noise out of the observations. By the Kolmogorov Extension Theorem, there exists a reference probability measure $\bar{P}$, under which the observation $\mathbf{y}_k$ are $N(0,1)$ IID random variables and therefore $\bar{P}$ is deemed to be an easier measure to work with. The filters are derived under the reference measure. We perform a measure change to construct the real-world measure P from the ideal-world measure $\bar{P}$ by invoking a discrete-time version of Girsanov's theorem. Let $\phi(z)$ denote the probability density function of a standard normal random variable z. For each component g, define

$$\lambda_l^g = \frac{\phi(\sigma^g(\mathbf{x}_{l-1})^{-1}(y_l^g - f^g(\mathbf{x}_{l-1})))}{\sigma^g(\mathbf{x}_{l-1})\phi(y_l^g)}, \tag{2}$$

and the Radon-Nikodým derivative of P with respect to $\bar{P}$, $\frac{dP}{d\bar{P}}|_{\mathscr{Y}_k} = \Lambda_k$, is given by

$$\Lambda_k = \prod_{g=1}^{d}\prod_{l=1}^{k}\lambda_l^g,\ k \geq 1,\ \Lambda_0 = 1. \tag{3}$$

To obtain the estimates of $\alpha(\mathbf{x}_k, \mathbf{x}_{k-1})$ under the real world measure, we first perform all calculations under the reference probability measure $\bar{P}$. Then Bayes' theorem is employed to relate the conditional expectation under two different measures. Note that we can also consider another reference probability measure $\tilde{P}$ under which the y_k's are $N(0, \sigma^2)$, $\sigma \neq 1$. In such a case, we define

$$\tilde{\lambda}_l^g = \frac{\phi(\sigma^g(\mathbf{x}_{l-1})^{-1}(y_l^g - f^g(\mathbf{x}_{l-1})))}{\phi(\sigma^g(\mathbf{x}_{l-1})^{-1}y_l^g)}.$$

Based on our numerical experiment, since $\tilde{\lambda}$ is much larger than λ, the speed of convergence with $\tilde{\lambda}$ is ten steps slower than using λ.

Let us derive the conditional expectation of $\alpha(\mathbf{x}_k, \mathbf{x}_{k-1})$ given $\mathscr{Y}_k$ under P. Write

$$p_k^{ij} = P(\mathbf{x}_k = \mathbf{e}_i, \mathbf{x}_{k-1} = \mathbf{e}_j|\mathscr{Y}_k) = E[\langle\alpha(\mathbf{x}_k, \mathbf{x}_{k-1}), \mathbf{e}_{ij}\rangle|\mathscr{Y}_k], \tag{4}$$

with $\mathbf{p}_k = (p_k^{11}, \ldots, p_k^{ij}, \ldots, p_k^{NN}) \in \mathbf{R}^{N^2}$. Using Bayes' theorem, we have

$$\mathbf{p}_k = E[\alpha(\mathbf{x}_k, \mathbf{x}_{k-1})|\mathscr{Y}_k] = \frac{\bar{E}[\Lambda_k\alpha(\mathbf{x}_k, \mathbf{x}_{k-1})|\mathscr{Y}_k]}{\bar{E}[\Lambda_k|\mathscr{Y}_k]}. \tag{5}$$

Defining $\mathbf{q}_k = \bar{E}[\Lambda_k\alpha(\mathbf{x}_k, \mathbf{x}_{k-1})|\mathscr{Y}_k]$ and $\mathbf{1} = (1, \ldots, 1)^\top \in \mathbf{R}^{N^2}$, we see that

$$\sum_{i,j}^{N}\langle\alpha(\mathbf{x}_k, \mathbf{x}_{k-1}), \mathbf{e}_{ij}\rangle = \langle\alpha(\mathbf{x}_k, \mathbf{x}_{k-1}), \mathbf{1}\rangle = 1,$$

so that

$$\langle\mathbf{q}_k, \mathbf{1}\rangle = \bar{E}[\Lambda_k\langle\alpha(\mathbf{x}_k, \mathbf{x}_{k-1}), \mathbf{1}\rangle|\mathscr{Y}_k] = \bar{E}[\Lambda_k|\mathscr{Y}_k]. \tag{6}$$

With Eqs. 5 and 6, we get the explicit form of the conditional distribution

$$\mathbf{p}_k = \frac{\mathbf{q}_k}{\langle\mathbf{q}_k, \mathbf{1}\rangle}.$$

Now, we need a recursive filter for the process $\mathbf{q}_k$ in order to estimate the state process $\alpha(\mathbf{x}_k, \mathbf{x}_{k-1})$. Define the diagonal matrix $\mathbf{B}_k$ by

$$\mathbf{B}_k = \begin{pmatrix} b_k^1 & & & & & \\ & \ddots & & & & \\ & & b_k^N & & & \\ & & & \ddots & & \\ & & & & b_k^1 & \\ & & & & & \ddots \\ & & & & & & b_k^N \end{pmatrix}, \tag{7}$$

where

$$b_k^i = \prod_{g=1}^{d} \frac{\phi((y_k^g - f_i^g)/\sigma_i^g)}{\sigma_i^g \phi(y_k^g)}. \tag{8}$$

Notation For any $\mathcal{Y}_k$-adapted process X_k, write $\hat{X}_k = E[X_k|\mathcal{Y}_k]$ and $\gamma(X)_k = \bar{E}[\Lambda_k X_k|\mathcal{Y}_k]$. Again invoking Bayes' theorem, we have

$$\hat{X}_k = \frac{\gamma(X)_k}{\bar{E}[\Lambda_k|\mathcal{Y}_k]}. \tag{9}$$

To estimate the parameters of the model, recursive filters shall be derived for the following processes:

1. J^{rst},the number of jumps from $(\mathbf{e}_s, \mathbf{e}_t)$ to state $\mathbf{e}_r$ up to time k.
2. O_k^{rs}, the occupation time of the higher-order Markov chain spent in state $(\mathbf{e}_r, \mathbf{e}_s)$ up to time k,
3. O_k^r, the occupation time spent by the higher-order Markov chain in state $\mathbf{e}_r$ up to time k,
4. $T_k^r(h)$, the level sum for the state $\mathbf{e}_r$, where h is a function with the form $h(y) = y$ or $h(y) = y^2$.

To obtain on-line estimates for the quantities of the above four related process, we shall take advantage of the semi-martingale representation in Eqs. 1 and 9. We can then obtain recursive equations for the vector quantities $J_k^{rst}\alpha(\mathbf{x}_k, \mathbf{x}_{k-1})$, $O_k^{rs}\alpha(\mathbf{x}_k, \mathbf{x}_{k-1})$, $O_k^r\alpha(\mathbf{x}_k, \mathbf{x}_{k-1})$ and $T_k^r(h)\alpha(\mathbf{x}_k, \mathbf{x}_{k-1})$. The recursive relation of these vector processes and $\mathbf{q}_k$ under a multi-dimensional observation set-up are given in the following proposition.

Proposition 1 *Let $\mathbf{V}_r$, $1 \leq r \leq N$ be an $N^2 \times N^2$ matrix such that the $((i-1)N + r)$th column of $\mathbf{V}_r$ is $\mathbf{e}_{ir}$ for $i = 1 \ldots N$ and zero elsewhere. If $\mathbf{B}$ is the diagonal matrix defined in Eq. 7, then*

$$\mathbf{q}_{k+1} = \mathbf{B}_{k+1}\mathbf{\Pi}\mathbf{q}_k \tag{10}$$

and

$$\gamma(J^{rst}\alpha(\mathbf{x}_{k+1},\mathbf{x}_k))_{k+1} = \mathbf{B}_{k+1}\mathbf{\Pi}\gamma(J^{rst}\alpha(\mathbf{x}_k,\mathbf{x}_{k-1}))_k + b^r_{k+1}\langle\mathbf{\Pi}\mathbf{e}_{st},\mathbf{e}_{rs}\rangle\langle\mathbf{q}_k,\mathbf{e}_{st}\rangle\mathbf{e}_{rs}, \quad (11)$$

$$\gamma(O^{rs}\alpha(\mathbf{x}_{k+1},\mathbf{x}_k))_{k+1} = \mathbf{B}_{k+1}\mathbf{\Pi}\gamma(O^{rs}\alpha(\mathbf{x}_k,\mathbf{x}_{k-1}))_k + b^r_{k+1}\langle\mathbf{q}_k,\mathbf{e}_{rs}\rangle\mathbf{\Pi}\mathbf{e}_{rs}, \quad (12)$$

$$\gamma(O^{r}\alpha(\mathbf{x}_{k+1},\mathbf{x}_k))_{k+1} = \mathbf{B}_{k+1}\mathbf{\Pi}\gamma(O^{r}\alpha(\mathbf{x}_k,\mathbf{x}_{k-1}))_k + b^r_{k+1}\mathbf{V}_r\mathbf{\Pi}\mathbf{q}_k, \quad (13)$$

$$\gamma(T^{r}(h)\alpha(\mathbf{x}_{k+1},\mathbf{x}_k))_{k+1} = \mathbf{B}_{k+1}\mathbf{\Pi}\gamma(T^{r}(h)\alpha(\mathbf{x}_k,\mathbf{x}_{k-1}))_k + h(y^g_{k+1})b^r_{k+1}\mathbf{V}_r\mathbf{\Pi}\mathbf{q}_k. \quad (14)$$

Proof See [33] for an analogous proof of each filter under the single observation setting. □

Similar to Eq. 6, by summing the components, Eqs. 11–14 give expressions for $\gamma(J^{rst})_k, \gamma(O^{rs})_k, \gamma(O^r)_k$ and $\gamma(T^r(g))_k$.

Now we make use of the expectation maximization (EM) algorithm to estimate the optimal parameters. The calculation is similar to the technique as in single observation set-up. The estimates are expressed in terms of the recursions in Eqs. 11–14, which are provided in the following proposition.

Proposition 2 *Suppose the observation is d-dimensional and the set of parameters* $\{\hat{a}_{rst},\ \hat{f}^g_r,\ \hat{\sigma}^g_r\}$ *determines the dynamics of* $y^g_k,\ k \geq 1,\ 1 \leq g \leq d$, *then the EM estimates for these parameters are given by*

$$\hat{a}_{rst} = \frac{\widehat{J}^{rst}_k}{\widehat{O}^{st}_k} = \frac{\gamma(J^{rst})_k}{\gamma(O^{st})_k},\ \forall \text{ pairs } (r,s),\ r \neq s, \quad (15)$$

$$\hat{f}^g_r = \frac{\widehat{T}^r_k}{\widehat{O}^r_k} = \frac{\gamma(T^r(y^g))_k}{\gamma(O^r)_k}, \quad (16)$$

$$\hat{\sigma}^g_r = \sqrt{\frac{\widehat{T}^r((y^g)^2)_k - 2\hat{f}^g_r\widehat{T}^r(y^g)_k + (\hat{f}^g_r)^2\widehat{O}^r_k}{\widehat{O}^r_k}}. \quad (17)$$

Proof See [33] for an analogous proof of each estimate under the single observation setting. □

Given the observation up to time k, new parameters $\hat{a}_{rst}(k),\ \hat{f}^g_r(k),\ \hat{\sigma}^g_r(k),\ 1 \leq r,s,t \leq N$ are given by Eqs. 15–17. The recursive filters for the unobserved Markov chain and the related process in Proposition 1 can be re-evaluated using the new estimates. Consequently, it allows the algorithm to update the parameters automatically.

3 Forecasting Indices

Suppose an investor wants to choose a portfolio with two investments to diversify his/her risk. In order to have such diversification the two assets should act differently during different periods in the economic cycle. For example, growth and value stocks tend to perform well at different times of the economic cycle, so switching between the classes at appropriate times may add value. We apply the iterative procedure derived in the previous section to two weekly datasets of stock indices: Russell 3000 growth and Russell 3000 value indices. The data were recorded from June 1995 to December 2010; thus there are 783 data points in each dataset. Both indices are constructed based on the Russell 3000 index, in which the underlying companies are all incorporated in the U.S. and representing approximately 98 % of the investable U.S. equity market. Companies within the Russell 3000 that exhibit higher price-to-book and forecasted earnings are used to form the Russell 3000 growth index. This subindex therefore measures the performance of the broad growth segment of the US equity market. The Russell 3000 value index includes Russell 3000 companies with lower price-to-book value and lower forecasted growth values. Therefore the Russell 3000 value index measures the performance of the value stocks in the US equity market.

The regime-switching models are developed to capture particular behaviour of the evolution of an asset price. We segregate the observation data into four intervals to investigate the index values and returns. Tables 1, 2 and 3 provide descriptive statistics of the Russell 3000 Index return together with the growth- and value-subindex returns for the entire period as well as the subperiods. The descriptive statistics demonstrate the possible segregation of the actual data into different states according to the levels of mean and volatility. We find the subperiods characterised by different levels of mean and volatility. For example, we can see that the log return y_k has a higher volatility when the mean is negative, and vice versa. If the data has only one state, the model will collapse to one regime. As a result, the estimated parameters of each state will be close to each other.

We consider the two indices as a two-dimensional observation process. The dynamics of the log returns are given by

$$y_{k+1}^{RV} = \log \frac{\text{RValue}(k+1)}{\text{RValue}(k)} = f^{RV}(\mathbf{x}_k) + \sigma^{RV}(\mathbf{x}_k) w_{k+1}^{RV}$$
$$y_{k+1}^{RG} = \log \frac{\text{RGrowth}(k+1)}{\text{RGrowth}(k)} = f^{RG}(\mathbf{x}_k) + \sigma^{RG}(\mathbf{x}_k) w_{k+1}^{RG}$$

Table 1 Summary statistics of Russell 3000 growth returns

	Entire data	06/95–07/98	07/98–09/03	09/03–08/08	09/05/08–12/31/10
Max	0.1659	0.0669	0.1659	0.0429	0.1090
Min	−0.1806	−0.0462	−0.1683	−0.0522	−0.1806
Median	0.0026	0.0052	0.0007	0.0012	0.0034
Mean	0.0010	0.0048	−0.0008	0.0009	0.0005
Std	0.0299	0.0196	0.0379	0.0184	0.0399
Skewness	−0.4478	−0.0281	−0.0721	−0.3730	−0.8401
Kurtosis	5.0859	0.1581	2.6943	0.2936	3.7011

Table 2 Summary statistics of Russell 3000 value return

	Entire data	06/95–07/98	07/98–09/03	09/03–08/08	09/05/08–12/31/10
Max	0.1381	0.0554	0.0719	0.0615	0.1381
Min	−0.2167	−0.0551	−0.1162	−0.0609	−0.2167
Median	0.0025	0.0052	−0.0007	0.0028	0.0031
Mean	0.0011	0.0042	0.0001	0.0011	−0.0005
Std	0.0266	0.0168	0.0258	0.0187	0.0465
Skewness	−0.8031	−0.2649	−0.2585	−0.4757	−0.7317
Kurtosis	8.1906	0.5851	1.6751	0.7746	4.0000

where

$$\mathbf{f}^{RV} = (f_1^{RV}, \ldots, f_N^{RV}) \in \mathbf{R}^N,\ \mathbf{f}^{RG} = (f_1^{RG}, \ldots, f_N^{RG}) \in \mathbf{R}^N,$$
$$\boldsymbol{\sigma}^{RV} = (\sigma_1^{RV}, \ldots, \sigma_N^{RV}) \in \mathbf{R}^N,\ \boldsymbol{\sigma}^{RG} = (\sigma_1^{RG}, \ldots, \sigma_N^{RG}) \in \mathbf{R}^N,$$

are governed by the same HOHMM $\mathbf{x}$. Here, w_k^{RV} and w_k^{RG} are $N(0, 1)$ IID random variables independent of each other. The data are processed in batches of ten observation points. At the end of each pass through the data, $\mathbf{f}$, $\boldsymbol{\sigma}$, $\mathbf{A}$ and $\boldsymbol{\Pi}$ are updated with new estimates using the formulas given in the previous section. These new estimates are in turn used as initial parameters for the next pass. This means the parameters are updated roughly every two and a half months. We process the data in batches in order to lower computational expenses. Furthermore, the use of data batches is consistent with the idea of suboptimal schemes; see page 15 of Elliott et al. [13]. In our case, such choice of ten data points provides the best fitting performance of the forecasts to the actual data. Investors can choose any length of a batch to update their information according to their needs. In our numerical experiment, we find that updating parameters every two and half month is sufficient to capture market information. Whilst utilising batches with fewer data points improves forecasting errors slightly, it does not lead to a better portfolio performance. Figure 1 displays the plot of the evolution of $\mathbf{f}^{RV}$, $\mathbf{f}^{RG}$, $\boldsymbol{\sigma}^{RV}$, $\boldsymbol{\sigma}^{RG}$ and the transition matrix $\mathbf{A}$ under the two-state HOHMM setting.

The optimal investment strategy is developed based on the forecasts of index returns. To assess the predictive performance of the model, we calculate the one-step ahead forecasts for both indices through the following equations

$$E[\text{RValue}_{k+1}|\mathscr{Y}_k] = \text{RValue}_k \sum_{i,j=1}^{N} \langle \mathbf{p}_k, \mathbf{e}_{ij} \rangle \exp(f_i^{RV} + (\sigma_i^{RV})^2/2) \tag{18}$$

Table 3 Summary statistics of Russell 3000 return

	Entire data	06/95–07/98	07/98–09/03	09/03–08/08	09/05/08–12/31/10
Max	0.1204	0.0611	0.0995	0.1204	0.1204
Min	−0.1986	−0.0507	−0.1267	−0.1986	−0.1986
Median	0.0024	0.0048	−0.0008	0.0038	0.0042
Mean	0.0012	0.0045	−0.0002	−0.0015	0.0000
Std	0.0272	0.0177	0.0302	0.0465	0.0426
Skewness	−0.6741	−0.1051	−0.2204	−0.6765	−0.8040
Kurtosis	5.8949	0.2693	2.0180	2.9595	3.9150

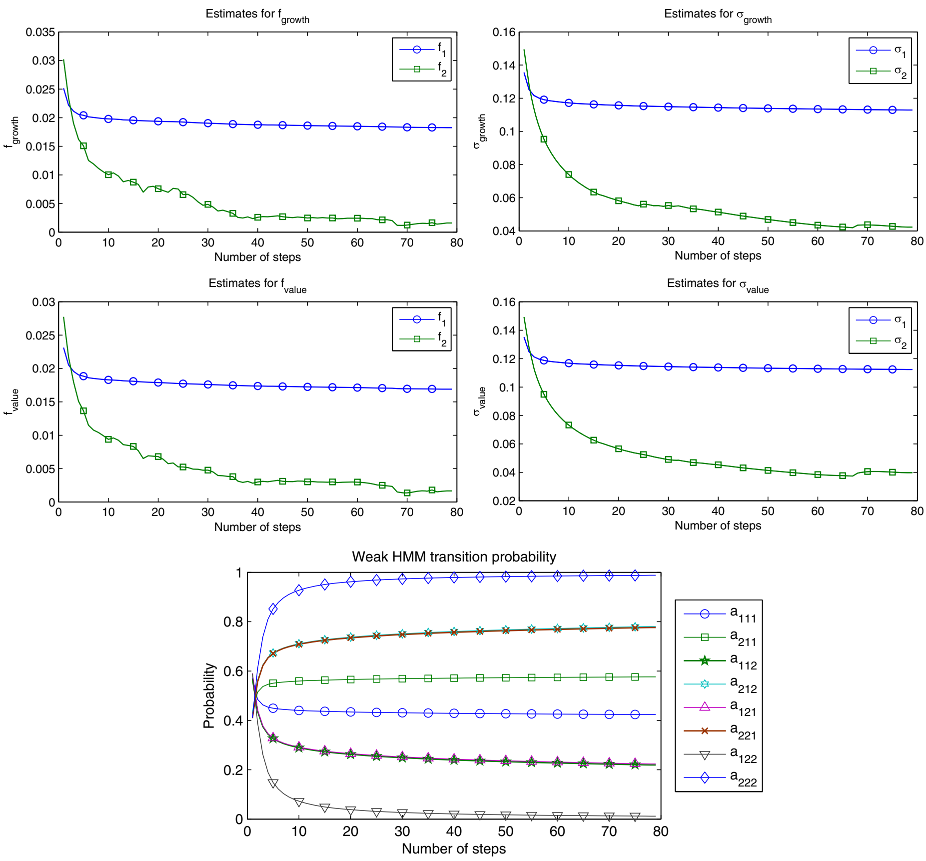

Fig. 1 Evolution of parameter estimates under the 2-state setting

$$E[\text{RGrowth}_{k+1}|\mathscr{Y}_k] = \text{RGrowth}_k \sum_{i,j=1}^{N} \langle \mathbf{p}_k, \mathbf{e}_{ij} \rangle \exp(f_i^{RG} + (\sigma_i^{RG})^2/2). \tag{19}$$

In this paper, we do not model the correlation amongst assets explicitly. However, the two indices are governed by the same hidden higher-order Markov chain, and

Table 4 Error measures: root mean square error (RMSE), absolute mean error (AME), relative absolute error (RAE) and absolute percentage error (APE) for one-step ahead forecasts under 1-, 2- and 3-state HOHMM set-ups

	1-state HOHMM	2-state HOHMM	3-state HOHMM
RAE value	0.0966	0.0976	0.1044
APE value	0.0188	0.0192	0.0216
MAE value	38.6423	39.0510	41.7675
RMSE value	54.4720	54.6768	56.7974
RAE growth	0.1159	0.1170	0.1241
APE growth	0.0215	0.0218	0.0241
MAE growth	43.0295	43.4276	46.0558
RMSE growth	64.4588	64.7098	66.5664

Table 5 Error measures restricted to financial crisis period: root mean square error (RMSE), absolute mean error (AME), relative absolute error (RAE) and absolute percentage error (APE) for one-step ahead forecasts under 1-, 2- and 3-state HOHMM set-ups

	1-state HOHMM	2-state HOHMM	3-state HOHMM
RAE value	1.0365	0.9836	0.9828
APE value	0.0191	0.0127	0.0142
MAE value	13.5314	15.8644	16.4198
RMSE value	17.8370	16.4198	20.0430
RAE growth	0.9530	0.4213	0.6690
APE growth	0.0088	0.0043	0.0075
MAE growth	11.9003	5.7440	10.0917
RMSE growth	13.5887	8.3321	12.0772

thus, they are correlated implicitly. Actual filters with correct correlation structure between Brownian motions driving the logreturns of growth and value indices will presumably be better. So, one may view that this study is as a lower bound for the validity of a larger study.

Our goal in this exercise is not to obtain a better forecast of either growth or value returns over a short time horizon: the relative size of the random to the deterministic terms in the stock price model makes this impossible. Instead, our goal is to determine whether HOHMM can improve our management of a given portfolio. However, for completeness, we do present an assessment of the goodness of fit of the one-step ahead forecasts. To make this comparison we use four criteria: root mean square error (RMSE), absolute mean error (AME), relative absolute error (RAE) and absolute percentage error (APE) for $N = 1$, $N = 2$, and $N = 3$. The results of this error analysis are given in Table 4.

The results show that the two-state model tends to outperform the three-state model in all forecasting metrics. Although the one-state model has a slight improvement, Table 4 shows that the performance of the 2-state model is statistically nearly indistinguishable with that of the one-state model. However, Table 5 depicts a similar error analysis restricted to the time window surrounding the financial crisis of 2007–2008. Table 5 clearly shows the advantages of a 2-state model over the one-state model especially in the modelling of the growth index. These results on forecast performance are interesting.

But the true test of a model like this is in its application to a trading strategy. We want to decide if these filter results enable us to better manage a portfolio. That is the topic of this paper's next section.

4 A Switching Investment Strategy

There are various asset allocation strategies that one can devise. Here we focus on a dynamic asset allocation which assumes active changes to an investment based on short-term market forecasts for returns. The switching strategy utilise the forecasted risk-adjusted returns of the indices as signals to switch investments between the Russell 3000 growth and Russell 3000 value index. The forecasted risk-adjusted return is calculated by dividing the forecasted returns by the realised volatility of each index covering the previous 20 data points.

We implement the switching strategy on a 15-year dataset recorded from June 1995 to December 2010. The observation data is divided into 15 intervals and each interval covers roughly one-year dataset. We suppose a hypothetical starting investment of \$100 and then apply the forecasting method on the period considered. At the beginning of each interval, the signals from the forecasted risk-adjusted return on both indices are compared. The full amount is invested in the index with the higher forecasted risk-adjusted return. We also assume that the transaction cost is a fixed percentage of total investment. When asset allocation changes, this transaction cost is subtracted from total investment.

The overall performance of the switching strategy is compared with that of the pure investment strategy on the basis of the log-return of the terminal wealth. Let X^{RG} and X^{RV} denote the differences of log-returns from the switching and pure strategies, which are defined by

$$X_i^{RG} = \log \frac{\mathrm{SW}_i}{100} - \log \frac{RG_i}{100} \tag{20}$$

$$X_i^{RV} = \log \frac{\mathrm{SW}_i}{100} - \log \frac{RV_i}{100}, \tag{21}$$

for $i = 1, 2, \ldots, 15$. Here, SW_i denotes the terminal wealth of the portfolio with switching strategy at the end of the ith interval. RG_i and RV_i denote the terminal wealth of the investment, holding 100 % of Russell 3000 growth index and Russell 3000 value index, respectively, at the end of the ith interval. The portfolio performance under varying transaction costs from five basis points (1 bp=0.01 %) to 70 bps is presented in Table 6. In addition, we present the performance of both switching and pure indices strategies using the usual HMM forecasts. Our study shows that the HOHMM-based switching strategy has higher values in Mean(X^{RV}) and Mean(X^{RG}) than those from HMM-based switching strategy yielding negative values. HOHMM-based strategy shows higher std(X^{RV}) and lower std(X^{RG}) than those based on HMM strategy. As we can observe, HOHMM-based Mean(X^{RV}) and Mean(X^{RG}) are positive and slightly decrease as transaction cost increases. This means that on average the log return from the switching strategy is higher than that from the pure index investments. Compared with the pure growth and value strategies, the HOHMM switching strategy has either the highest or the second highest terminal value in 15 intervals. Figure 2 displays number of the intervals in which switching strategy has the highest and the second highest terminal values for transaction cost varying from 1 bp to 80 bps. We observe, however, high values of std(X^{RV}) and std(X^{RG}), which indicate a high risk of employing the switching strategy. We next introduce a mixed strategy to address the diversification of risks.

Table 6 Performance comparison for HOHMM- and HMM-based switching strategies with varying transaction costs

Transaction	5 bps		20 bps		50 bps		70 bps	
cost	HOHMM	HMM	HOHMM	HMM	HOHMM	HMM	HOHMM	HMM
Mean (X^{RG}) %	3.9106	−0.6895	3.5163	−0.9281	2.7261	−1.4062	2.1979	−1.7258
Std (X^{RG}) %	2.4102	6.7138	2.4701	6.5882	2.7086	6.3598	2.9403	6.2259
Mean (X^{RV}) %	11.2393	−9.5292	10.8450	−9.7677	10.0547	−10.2458	9.5266	−10.5654
Std (X^{RV}) %	18.4701	15.8552	18.2859	16.0045	17.9291	16.3117	17.7003	16.5226

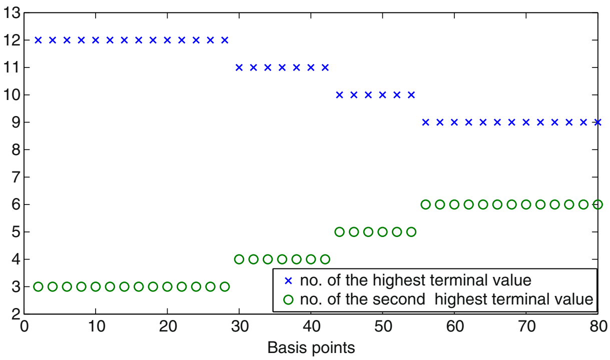

Fig. 2 Numbers of the intervals switching strategy has the highest and the second highest terminal values for varying transaction cost

5 A Mixed Investment Strategy

Selecting an investment strategy is similar to the asset allocation decision problem in that one tries to maximise strategy return whilst controlling portfolio risk. The risk is evaluated in terms of the variance of the portfolio's return. The mean-variance problem entails maximising the expected portfolio's return and minimising the variance of the portfolio's return. In this section, we investigate a mixed asset allocation strategy on two assets whose dynamics are modulated by HOHMM. With this strategy investors determine the optimal weight of each asset to allocate based on the estimated parameters and the state of the higher-order Markov chain. The development here follows the applications of results in [15, 16]. Suppose an investor is guided by an optimization equation MV, which is a linear combination of the expected portfolio's return and variance of the portfolio's return. Let $\mathbf{w} = (w_g, w_v)$ denote the weight of Russell 3000 growth and value indices, respectively. To solve this mean-variance problem, we wish to estimate the optimal $\mathbf{w}$ which maximises the function

$$MV(\mathbf{w}) = vE[w_g y^{RG}_{k+1} + w_v y^{RV}_{k+1} | \mathscr{Y}_k] - \mathrm{Var}[w_g y^{RG}_{k+1} + w_v y^{RV}_{k+1} | \mathscr{Y}_k],$$

where y^{RG} and y^{RV} are the returns of Russell 3000 growth and value indices, and v is a nonnegative risk aversion factor. The optimal weights are given in the following proposition.

Proposition 3 *Let $v > 0$ be the risk aversion factor. Suppose that neither short selling nor borrowing is allowed. The optimal weight w_g is given by*

$$w_g = \begin{cases} \frac{v(\langle \mathbf{f}^{RG}, \hat{\mathbf{x}}_k\rangle - \langle \mathbf{f}^{RV}, \hat{\mathbf{x}}_k\rangle) + 2\langle \boldsymbol{\sigma}^{RV}, \hat{\mathbf{x}}_k\rangle^2}{2(\langle \boldsymbol{\sigma}^{RG}, \hat{\mathbf{x}}_k\rangle^2 + \langle \boldsymbol{\sigma}^{RV}, \hat{\mathbf{x}}_k\rangle^2)} & \text{when } -2\langle \boldsymbol{\sigma}^{RV}, \hat{\mathbf{x}}_k\rangle^2 < v(\langle \mathbf{f}^{RG}, \hat{\mathbf{x}}_k\rangle - \langle \mathbf{f}^{RV}, \hat{\mathbf{x}}_k\rangle) \\ & < 2\langle \boldsymbol{\sigma}^{RG}, \hat{\mathbf{x}}_k\rangle^2 \\ 1 & \text{when } v(\langle \mathbf{f}^{RG}, \hat{\mathbf{x}}_k\rangle - \langle \mathbf{f}^{RV}, \hat{\mathbf{x}}_k\rangle) > 2\langle \boldsymbol{\sigma}^{RG}, \hat{\mathbf{x}}_k\rangle^2 \\ 0 & \text{when } v(\langle \mathbf{f}^{RG}, \hat{\mathbf{x}}_k\rangle - \langle \mathbf{f}^{RV}, \hat{\mathbf{x}}_k\rangle) < -2\langle \boldsymbol{\sigma}^{RV}, \hat{\mathbf{x}}_k\rangle^2 \end{cases},$$

and the optimal weight w_v is given by $w_v = 1 - w_g$.

Proof See [16] for proof. □

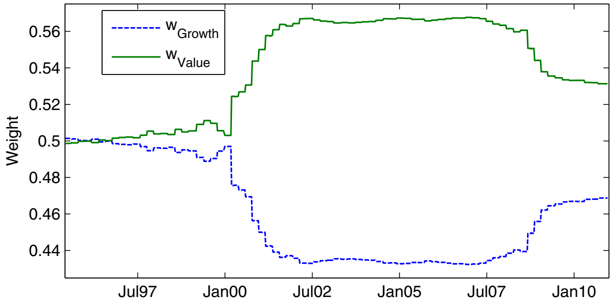

Fig. 3 Optimal weights for Russell 3000 value and growth indices in the HOHMM-based mixed strategy with $v = 0.08$

Note that

$$\langle \hat{\mathbf{x}}_k, \mathbf{e}_r \rangle = \sum_{i=1}^{N} E[\langle \alpha(\mathbf{x}_k, \mathbf{x}_{k-1}), \mathbf{e}_{ri} \rangle | \mathscr{Y}_k], \text{ for } r = 1 \dots N,$$

where $\mathbf{e}_{ri} = \alpha(\mathbf{e}_r, \mathbf{e}_i)$ indicating that indeed, the optimal weight depends on the state of the embedded MC $\alpha(\mathbf{x}_k, \mathbf{x}_{k-1})$.

Note that the weights belong to the interval [0, 1] since neither short selling nor borrowing is allowed. Similar to the previous section, we divide the observation data into 15 intervals. For each interval, the optimal weights are calculated for each time k utilising the optimal parameters and the estimated states of the weak Markov chain. Investors can allocate investment using different parameter updating frequency depending on their goal. To achieve consistency in comparison with the switching and pure index strategies, the weights employed for each index is updated at the beginning of each interval; transaction cost will also be considered. To gauge the strategy performance, we shall focus on the terminal value of the portfolio.

Figure 3 exhibits a plot of optimal weights for Russell 3000 growth and value indices. The risk aversion factor v is a scaling constant which is chosen by the investor. Here, we allow this factor to vary from $v = 0$ (totally avoiding risk) to $v = 5$ (seeking some risk). The evolution of optimal weights for Russell 3000 growth index with different values of v is shown in Fig. 4. When v is small, the investor is relatively conservative. The switching of market's regime has less impact on his/her asset

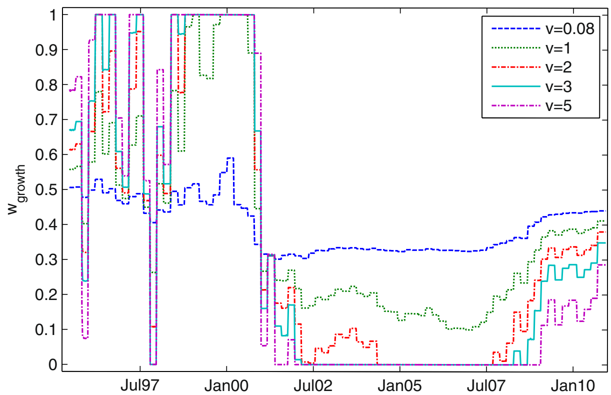

Fig. 4 Evolution of optimal weights for Russell 3000 growth index in the HOHMM-based mixed strategy with varying v's

Table 7 Performance comparison between HOHMM- and HMM-based mixed strategies with varying transaction costs

Transaction cost	5 bps		20 bps		50 bps		70 bps	
	HOHMM	HMM	HOHMM	HMM	HOHMM	HMM	HOHMM	HMM
Mean (X^{RG}) %	4.5606	2.9287	3.5750	1.8686	1.5993	−0.2564	0.2789	−1.6767
Std (X^{RG}) %	10.664	11.8866	10.3087	11.5399	9.7089	10.9647	9.4039	10.6797
Mean (X^{RV}) %	−2.7680	−5.9108	−3.7536	−6.9710	−5.7292	−9.0960	−7.0497	−10.5163
Std (X^{RV}) %	8.5098	7.6077	8.9849	8.1914	10.0169	9.4298	10.7537	10.2950

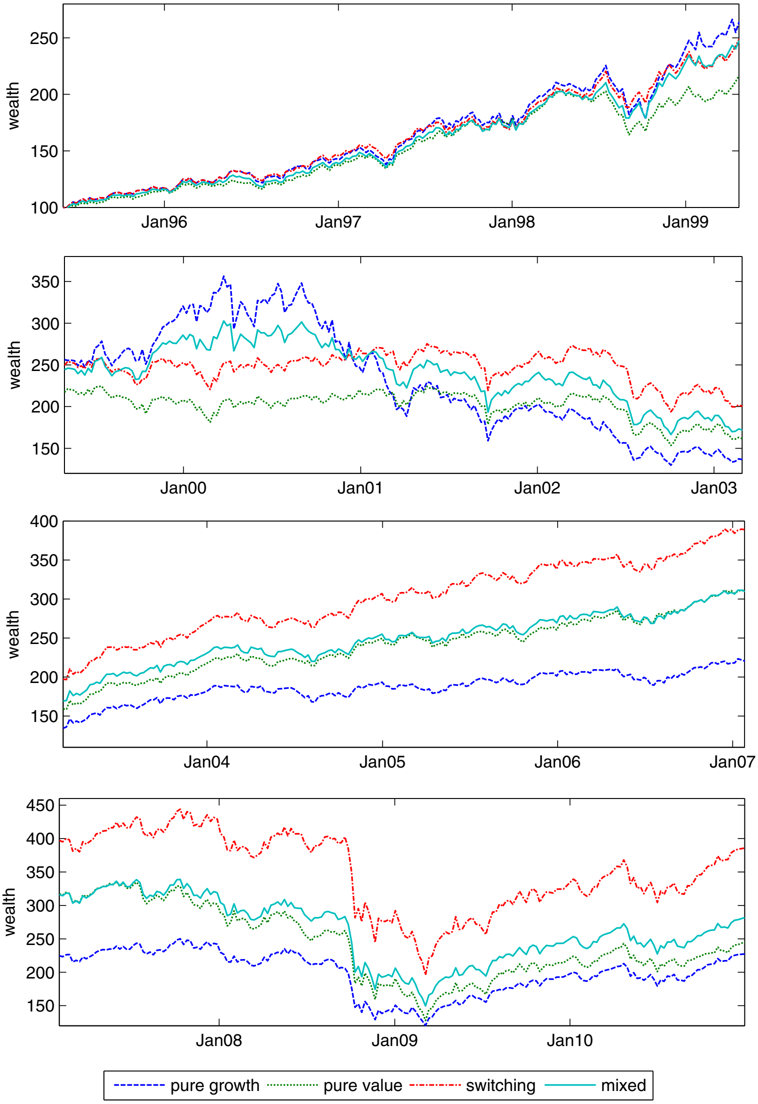

Fig. 5 Switching, mixed, pure growth and pure value strategies comparison between 1995 and 2010

allocation as can be viewed from the stable variation of weights for the Russell 3000 growth index. The investor with higher v appears to aggressively react to market regime switching. The Russell 3000 growth index has higher risk than the value index before September 2001 and it has lower risk after that. Consequently, the weight to allocate in growth index is higher than 0.5 before this time and it drops below 0.5 when the index has less uncertainty.

Table 7 shows the overall performance of the mixed strategy, which is compared with the pure growth and pure value strategies with $v = 0.08$, obtained though the analogue formulae of Eqs. 20 and 21. The standard deviations of the differences of returns, std(X^{RG}) and std(X^{RV}), are lower than that of using switching strategy as we expected. Mean(X^{RG}) and Mean(X^{RV}) decrease as transaction cost increases. Compared to the mixed strategy based on the forecasts under the usual HMM framework, the HOHMM-based mixed strategy produces higher values in both mean and standard deviation. The HOHMM setting certainly carries more opportunities to explore the trade off between expected return and risk, which means higher risk may lead to higher return.

Figure 5 presents the evolution of investment under the switching, mixed and pure index strategies. Each subplot covers three years of data. We can see that based on the value of the investments, the mixed strategy does not always outperform other strategies. It is not straightforward to establish from the plots which strategy is the best. We shall then evaluate the portfolio performance through some classical measures in investment.

6 Evaluating the Portfolio Performance

In this section, we carry out performance comparisons amongst portfolio allocation strategies developed in the previous sections using historical data and simulated data. We begin by discussing and evaluating the simple performance of the portfolio. We evaluate the portfolio performance through a benchmark. In this case, the Russell 3000 index is a natural benchmark since both the Russell 3000 growth and value indices are its subindices. The comparison of four portfolios with the benchmark is made using three classical measures on returns; these are the Sharpe ratio, Jensen's alpha and the appraisal ratio.

Table 8 demonstrates the performance of both the switching and the mixing strategies as compared to the performance of the simple "all in the value" or "all in

Table 8 Unconditional and conditional performance of both switching strategy and mixed strategy

	Mean return (%)	Std error of mean (%)	Volatility (%)	Std error of volatility (%)	Weeks
Switching strategy	8.80	4.90	19.04	1.08	784
Switching invest in value index	10.40	5.90	14.49	0.65	375
Switching invest in growth index	7.80	7.80	24.01	2.31	409
Pure value index	5.70	5.00	17.96	1.01	784
Pure growth index	5.20	5.70	17.02	1.08	784
Mixed strategy	6.80	5.00	19.38	0.98	784
Weight of value index > 0.5	5.70	5.20	19.86	1.02	735
Weight of growth index > 0.5	25.50	9.80	9.66	0.65	49

the growth" subindices. These strategies either switch between the Russell value and growth indices (whose properties are summarised in the middle two rows) or create a mixed portfolio of the two units, blended between them, as described in this paper. All quantities are given in annualised units. The unconditional average performance of the switching strategy is given in row 1, the unconditional average performance of the mixing strategy is given in row 6, and the unconditional average performance of the constituent subindices are given in rows 4 and 5. All of the unconditional averages are computed for 784 weeks. Conditional average performances are also presented, in rows 2 and 3 for the switching strategy and in the final two rows (rows 7 and 8) for the mixing strategy. The number of weeks during which the various conditions presented applied are given in the rightmost column of the table.

From Table 8, ignoring the standard error measurements for the time being, both the new strategies appear to yield a higher return than their simple counterpart. The high standard errors of these measurements highlight the ever present difficulty of estimating strategy returns. However, both of these strategies obtain these higher returns at the cost of increased strategy risk from the 17 % range to the 19 % range. The conditional performances summarised here explain this. In the switching strategy, higher returns during the times at which the strategy recommends the growth index come at a significantly increased cost in risk. In contrast, the mixed strategy is able to obtain huge returns during the small fraction of the time that it allocates more than half the wealth to the growth portfolio at low risk; this only occurs because the riskiest days of the market were all experienced when the strategy opted for a majority in the value index.

The mixed strategy results displayed in Table 7 appear counterintuitive, because it appears that the usual risk-reduction effect of diversification are not in evidence. To understand this, one has to recall that the growth and value index returns are very strongly correlated to one another reducing the benefits of diversification in this case.

We now go on to evaluate a number of risk-adjusted portfolio measures. The first measure is the Sharpe ratio, denoted by SR, and

$$\mathrm{SR} = \frac{E[R_{\mathrm{portfolio}} - R_{\mathrm{riskfree}}]}{\sqrt{\mathrm{Var}(R_{\mathrm{portfolio}} - R_{\mathrm{riskfree}})}},$$

where R_{riskfree} is the risk-free interest rate. This is used to characterise how well the return of an asset compensates the investors for the risk taken. The higher the Sharpe ratio the higher is the return with the same level of risk. In Table 9, we tabulate the Sharpe ratio of five investment strategies using the dataset divided into 15 intervals. Note that the switching strategy has the same Sharpe ratio with that of one of either pure value or pure growth strategy. At the beginning of each interval, the switching strategy allocates to one of the subindices with a number of shares depending on the value of the switching investment and the chosen index at previous time step. Hence, the switching portfolio and the chosen subindex have the same return in one interval. Such similarity is also true in other measures since all the calculations are based on the returns. The differences amongst the strategies are small. Out of the 15 intervals, the switching strategy shows a better performance than the benchmark in 11 intervals and the mixed strategy outperforms the benchmark in six intervals. Both the HOHMM switching and HOHMM mixed strategies have higher risk-adjusted mean than the benchmark. In particular, the switching strategy shows the highest risk-adjusted mean in all of the strategies.

Table 9 Sharpe ratio for five investment strategies using 15 intervals

Period	Switching strategy	Mixed strategy	Pure Russell 3000 value	Pure Russell 3000 growth	Pure Russell 3000 index
1	0.2164	0.2016	0.1645	0.2164	0.2020
2	0.2254	0.2178	0.2254	0.2020	0.2177
3	−0.0083	−0.0298	−0.0523	−0.0083	−0.0292
4	0.0999	0.1485	0.0999	0.1699	0.1491
5	0.0141	0.0095	0.0141	0.0058	0.0184
6	−0.0751	−0.1347	−0.0751	−0.1652	−0.1360
7	−0.1488	−0.1930	−0.1488	−0.2394	−0.1984
8	0.1252	0.1304	0.1252	0.1331	0.1311
9	0.0206	−0.0393	0.0206	−0.1104	−0.0475
10	0.0701	0.0630	0.0555	0.0701	0.0633
11	0.0976	0.0701	0.0976	0.0346	0.0653
12	−0.0839	−0.1471	−0.1904	−0.0839	−0.1405
13	−0.1293	−0.1294	−0.1293	−0.1264	−0.1292
14	0.0259	0.0243	0.0259	0.0219	0.0241
15	0.6620	0.5511	0.4190	0.6620	0.5581
Mean	0.0741	0.0495	0.0435	0.0521	0.0499
	(4.96×10^{-4})	(4.74×10^{-4})	(3.92×10^{-4})	(5.42×10^{-4})	(4.7×10^{-4})
Std	0.1976	0.1890	0.1577	0.2168	0.1906
	(5.99×10^{-4})	(4.46×10^{-4})	(3.10×10^{-4})	(5.82×10^{-4})	(4.50×10^{-4})
Mean/Std	0.3751	0.2622	0.2755	0.2405	0.2618
	(2.24×10^{-3})	(2.55×10^{-3})	(2.66×10^{-3})	(2.53×10^{-3})	(2.50×10^{-3})

Numbers inside the parentheses are standard errors

We calculate Jensen's alpha, which is often used to measure the abnormal return of a portfolio over the expected return. This is denoted by α_J and it is the constant in the regression model,

$$\alpha_J = R_{\text{portfolio}} - [R_{\text{riskfree}} - \beta_{\text{portfolio}}(R_{\text{benchmark}} - R_{\text{riskfree}})].$$

A positive alpha indicates the portfolio has a higher marginal return. Table 10 shows the Jensen's alpha for four allocation strategies. Although the differences amongst the values of α are very small, there are 11 and five positive α's out of 15 for the switching and mixed strategies, respectively. It indicates the marginal returns in these periods are higher than that of the benchmark.

Finally, we consider the Treynor and Black's appraisal ratio (AR), also known as the information ratio. It is defined as the ratio between relative return and the relative risk and is given by

$$\text{AR} = \frac{E[R_{\text{portfolio}} - R_{\text{benchmark}}]}{\sqrt{\text{Var}(R_{\text{portfolio}} - R_{\text{benchmark}})}}.$$

The formula is very similar to the Sharpe ratio. Whereas the Sharpe ratio measures return relative to a riskless asset, the AR looks at returns relative to a risky benchmark. The higher the AR, the higher is the active return of the portfolio given the same risk level. Table 11 reports the AR of four investment strategies. The switching strategy outperforms the mixed strategy in 11 intervals. In particular, we have the highest mean and lowest standard deviation under this measure. We observe

Table 10 Jensen's alpha for four investment strategies using 15 intervals

Period	Switching strategy ($\times 10^{-4}$)	Mixed strategy ($\times 10^{-4}$)	Pure Russell 3000 value ($\times 10^{-4}$)	Pure Russell 3000 growth ($\times 10^{-4}$)
1	1.2306	−0.0417	1.230	−1.3637 6
2	3.6729	0.0300	−3.7262	3.6729
3	4.0153	−0.1148	4.0153	−4.3166
4	−5.6901	−0.0249	5.0587	−5.6901
5	6.1203	−2.5431	−10.2510	6.1203
6	17.7209	0.8786	−21.5255	17.7209
7	12.7347	1.5509	−14.0599	12.7347
8	−0.1632	−0.0380	0.1686	−0.1632
9	11.3998	1.4556	−12.270	11.3998 7
10	1.8036	−0.1018	1.8036	−1.5748
11	6.3168	0.9435	−6.3031	6.3168
12	13.8661	−1.7868	13.8661	−14.7742
13	−2.1714	−0.2851	1.9350	−2.1714
14	−0.4850	−0.0225	0.5308	−0.4850
15	18.0944	−0.9831	18.0944	−18.4459
Mean	5.8977	−0.0722	−1.4289	0.5987
	(0.0183)	(0.0028)	(0.0261)	(0.0245)
Std	7.3348	1.0969	10.3850	9.6750
	(0.0099)	(0.0022)	(0.0182)	(0.0169)
Mean/Std	8040.7842	−658.2883	−1375.9094	618.8129
	(26.1277)	(28.2625)	(28.1958)	(28.4250)

Numbers inside the parentheses are standard errors

Table 11 Appraisal ratio (AR) for four investment strategies using 15 intervals

Period	Switching strategy	Mixed strategy	Pure Russell 3000 value	Pure Russell 3000 growth
1	0.1539	−0.1761	−0.1571	0.1539
2	−0.0239	−0.0252	−0.0239	0.0233
3	0.0863	−0.1354	−0.0921	0.0863
4	−0.1729	−0.0946	−0.1729	0.1781
5	−0.0130	−0.2037	−0.0130	−0.0227
6	0.1810	0.1199	0.1810	−0.1746
7	0.2562	0.1838	0.2562	−0.2658
8	−0.0528	−0.0547	−0.0528	0.0529
9	0.3399	0.3356	0.3399	−0.3478
10	0.0251	−0.0042	−0.0206	0.0251
11	0.1213	0.1122	0.1213	−0.1222
12	0.3355	−0.3366	−0.3338	0.3355
13	−0.0718	−0.0686	−0.0718	0.0703
14	0.0260	0.0307	0.0260	−0.0250
15	0.3435	−0.3505	−0.3388	0.3435
Mean	0.1023	−0.0445	−0.0235	0.0207
	(4.03×10^{-4})	(4.66×10^{-4})	(4.88×10^{-4})	(4.90×10^{-4})
Std	0.1633	0.1869	0.1926	0.1953
	(2.14×10^{-4})	(3.19×10^{-4})	(3.17×10^{-4})	(3.27×10^{-4})
Mean/Std	0.6263	−0.2381	−0.1220	0.1060
	(27.1×10^{-4})	(29.41×10^{-4})	(29.01×10^{-4})	(28.63×10^{-4})

Numbers inside the parentheses are standard errors

Table 12 *p*-values for the Jarque-Bera test of normality on data given in Tables 9–11

	Switching strategy	Mixed strategy	Pure Russell value	Pure Russell growth	Pure Russell index
Sharpe ratio	0.0026	0.0350	0.3136	0.0111	0.0322
Jensen's alpha	0.4092	0.3608	0.5000	0.5000	–
Appraisal ratio	0.4929	0.5000	0.5000	0.5000	–

higher mean for switching strategy in each performance measure. A *t*-test is carried out to assess whether the means of portfolio under various performance measures are statistically different. In order to run a *t*-test, each of the datasets (i.e., columns of observations in Tables 9 through 11) being compared must be checked for normality. Table 12 presents the results of Jarque-Bera normality tests for all three portfolio measures applied to the five different portfolio selection approaches. The *p*-values for Jensen's alpha and Appraisal ratio of all portfolios are high which suggests there is no sufficient evidence to indicate that these data sets are coming from a non-normal distribution. Moreover, at 0.05 significance level, we can reject the null hypothesis that under the Sharpe ratio, the data on switching, mixed, pure growth and pure index strategies are from normal distribution. We test the difference between each pair of portfolios for the two measures. It has to be noted that the measures or criteria for comparison make use of the same experiment data set. The inherent problem with multiple comparison is the increase in type I error (i.e., probability of falsely rejecting the null hypothesis of no significance) that occurs when statistical tests are used repeatedly. To control this familywise error rate, which is the probability of making one or more false discoveries associated with multiple hypotheses tests, *p*-values are adjusted. The adjustment employed the Dunn-Šidák procedure, which is less conservative and more powerful than the Bonferroni method. Table 13 shows the Dunn-Šidák-corrected *p*-values for a one-tailed paired *t*-tests of significance assuming unequal variances. Comparing the switching and mixed strategies, the *p*-values in the first column are very small. This tells us that the difference in means under these performance measures of these two strategies is highly significant. The same can be said for the comparison of switching and pure growth strategies under the Jensen's alpha criterion. We recall that the switching strategy has the best performance for the period considered. For the switching versus pure value strategies, the *p*-values are no longer small so that we cannot reject the null hypothesis, i.e., we cannot reject that the two means are equal. Similar conclusion can be made when we compare the mixed and pure value strategies as well as the mixed and pure growth strategies where the *p*-values are extremely large. In addition to the *t*-test, we use the Wilcoxon rank sum test in assessing the significance of the differences. The *p*-values, also corrected based on Dunn-Šidák's method for multiple comparison, are

Table 13 Dunn-Šidák-adjusted *p*-values for a one-tailed significance test on the performance results shown in Tables 10–11

	Switching vs Mixed	Switching vs Pure growth	Switching vs Pure value	Mixed vs Pure growth	Mixed vs Pure value
Jensen's alpha	0.0070	0.0345	0.1004	0.5257	0.6360
Appraisal ratio	0.0296	0.2127	0.0632	0.3261	0.6182

Table 14 Dunn-Šidák-adjusted *p*-values for a Wilcoxon rank sum test on the performance results shown in Tables 9–11

	Switching vs Mixed	Switching vs Pure growth	Switching vs Pure value	Mixed vs Pure growth	Mixed vs Pure value
Sharpe ratio	0.9712	0.9712	0.9977	1.0000	1.0000
Jensen's alpha	0.0475	0.1821	0.2840	0.9446	0.9667
Appraisal ratio	0.0997	0.6419	0.2065	0.6567	0.9667

reported in Table 14. Wilcoxon test does not rely on the normality assumption and so it complements our use of the *t*-test. For switching strategy versus mixed strategies, the result suggests that the Jensen's alpha are significantly different but not for the appraisal ratio. Note that it may appear counter intuitive that the Jensen's alpha and the appraisal ratios are actually better for the switching than for the mixing strategy. Keep in mind, however, that we did not optimise the strategies for either ratio but rather for raw return; these tests were done after the fact to compare return-optimal strategies. Next, we give a simulation analysis in conjunction with the three portfolio measures.

We are interested in the statistical inference of the above portfolio measures for each portfolio strategy. The bootstrap is a way of finding sampling distribution from one sample path. Introduced by Efron and Tibshirani [12], it is a technique allowing estimation of the sample distribution of almost any statistics. This method can be implemented when the sample could be assumed to be drawn from an independent

Table 15 Performance evaluation for 10,000 bootstrapped datasets with 5 bps transaction cost

Sharpe ratio	Mean ($\times 10^{-2}$)	Std ($\times 10^{-2}$)	95 % Con.Int. ($\times 10^{-2}$)
Switching	3.1617	0.1711	[3.1498 3.1737]
Mixed	2.8455	0.1321	[2.8363 2.8548]
Pure value	2.8137	0.1430	[2.8038 2.8238]
Pure growth	1.7502	0.0829	[1.7445 1.7561]
Pure index	2.2978	0.1056	[2.2904 2.3052]
Jensen's α	Mean ($\times 10^{-3}$)	Std ($\times 10^{-3}$)	95 % Con.Int. ($\times 10^{-3}$)
Switching	0.4037	0.0607	[0.3995 0.4080]
Mixed	0.2259	0.0483	[0.2225 0.2293]
Pure Value	0.2673	0.0370	[0.2647 0.2699]
Pure growth	−0.0430	0.0506	[−0.0465 −0.0395]
AR	Mean ($\times 10^{-3}$)	Std ($\times 10^{-3}$)	95 % Con.Int. ($\times 10^{-3}$)
Switching	1.4270	0.4287	[1.3970 1.4570]
Mixed	−2.3559	0.3149	[−2.3779 −2.3338]
Pure value	−7.0522	0.5095	[−7.0879 −7.0165]
Pure growth	−6.2734	0.3956	[−6.3010 −6.2457]
Mean & Std return	Mean ($\times 10^{-3}$)	Std ($\times 10^{-3}$)	95 % Con.Int. ($\times 10^{-3}$)
Switching	0.9250	0.0429	[0.9220 0.9280]
Mixed	0.8682	0.0244	[0.8665 0.8699]
Pure value	0.7770	0.0065	[0.7765 0.7774]
Pure growth	0.7829	0.0074	[0.7824 0.7834]
Pure index	0.8967	0.0051	[0.8964 0.8971]

and identically distributed population. Bootstrap method constructs a number of resamples of the observation datasets with equal size by random sampling with replacement from the original dataset. The datasets used for the bootstrapping are the Russell 3000 index, the growth subindex and the value subindex. Each of the original datasets contains 784 data points. The procedure of the simulation is as follows:

1. We divide the datasets into 16 intervals and each interval contains 49 weeks.
2. A resample is created by repeatedly sampling with replacement from these 16 intervals. This means that we randomly pick one interval for the resample path and put the interval back for drawing again. As a result, any interval can be drawn more than once, or not at all. The resample has the same size as the original data.
3. The three measures are calculated for the new resample path.

The construction is repeated 10,000 times and the statistics of the three classical measures are obtained. Table 15 shows the statistics of the portfolios, Sharpe ratio, Jensen's alpha and AR with a transaction cost amounting to 5 bps. Table 16 presents the same analysis with transaction cost of 30 bps. When 5 bps transaction costs are introduced, the HOHMM switching and mixed strategies generate higher mean Sharpe ratio and AR from the 10,000 bootstrap sample paths than those from other strategies. Only the pure growth strategy leads to a negative mean in the Jensen's alpha measure. The standard deviation of switching strategy from the bootstrapped samples is higher than that in the mixed strategy in all cases. The estimated 95 %

Table 16 Performance evaluation for 10,000 bootstrapped datasets with 30 bps transaction cost

Sharpe ratio	Mean ($\times 10^{-2}$)	Std ($\times 10^{-2}$)	95 % Con.Int. ($\times 10^{-2}$)
Switching	3.7352	0.3169	[3.7130 3.7574]
Mixed	2.4876	0.1220	[2.4791 2.4962]
Pure value	2.8781	0.1380	[2.8684 2.8877]
Pure growth	1.5679	0.1173	[1.5597 1.5761]
Pure index	2.1326	0.1176	[2.1244 2.1408]
Jensen's α	Mean ($\times 10^{-3}$)	Std ($\times 10^{-3}$)	95 % Con.Int. ($\times 10^{-3}$)
Switching	5.0273	0.7980	[4.9714 5.0832]
Mixed	1.6992	0.5069	[1.6637 1.7347]
Pure value	2.6856	0.4149	[2.6565 2.7146]
Pure growth	−0.2362	0.5838	[−0.2770 −0.1953]
AR	Mean ($\times 10^{-3}$)	Std ($\times 10^{-3}$)	95 % Con.Int. ($\times 10^{-3}$)
Switching	7.5903	0.8554	[7.5304 7.6501]
Mixed	1.3412	0.3479	[1.3169 1.3656]
Pure value	3.4708	0.2721	[3.4517 3.4898]
Pure growth	−5.8516	0.5834	[−5.8925 −5.8108]
Mean & Std return	Mean ($\times 10^{-3}$)	Std ($\times 10^{-3}$)	95 % Con.Int. ($\times 10^{-3}$)
Switching	1.0447	0.0673	[1.0400 1.0494]
Mixed	0.7382	0.0193	[0.7369 0.7396]
Pure value	0.7669	0.0058	[0.7665 0.7674]
Pure growth	0.6304	0.0108	[0.6297 0.6312]
Pure index	0.7236	0.0070	[0.7231 0.7241]

confidence intervals for the mixed strategy are smaller than those in the switching strategy. Apparently, the HOHMM mixed strategy is more stable than the HOHMM switching strategy in terms of the standard deviation and 95 % confidence interval under the 5 bps transaction cost. A comparison of the mean and variance of the portfolio returns is also presented. The switching strategy outperforms other strategies with the highest variance nonetheless. On the other hand, the pure Russell 3000 index strategy has the lowest variance, but the 95 % confidence interval is bigger than those in both HOHMM strategies. It indicates that both HOHMM strategies are more stable than the benchmark in terms of the confidence interval. When transaction costs are set to 30 bps, the switching strategy produces the highest mean in all cases. Since the mixed strategy is the most costly strategy, it has a lower mean than pure value strategy under all measures. However, it still has positive means in both Jensen's alpha and AR measures. Both HOHMM-based strategies outperform the benchmark with 30 bps transaction cost in terms of higher Sharpe ratio, positive Jensen's alpha and positive AR. The switching strategy is the most risky judging the standard deviation of the measures. The mixed strategy shows a lower variance than the pure sub-indices strategies in Sharpe ratio and Jensen's alpha, and a lower variance than the pure growth strategy in the AR. Furthermore, the smaller 95 % confidence interval of mixed strategy indicates that it is more stable than the switching strategy in the case of 30 bps transaction cost. Therefore, we could conclude that the HOHMM switching strategy gives a higher mean return, however the HOHMM mixed strategy is less risky and more stable.

7 Conclusion

This paper examined asset allocation strategies for growth and value stocks under a weak hidden Markovian regime-switching setting. We suppose that the mean and volatilities of the price indices returns are modulated by a discrete-time multivariate HOHMM process. Recursive optimal estimates by filtering multidimensional observations are given for the state and various processes related to the underlying second-order Markov chain. The parameters of the model, including the transition probabilities, the drift and the variance parameters in the multidimensional observations, can be re-estimated and the forecasts can be obtained using the estimates. We investigated two investment strategies: a switching strategy and a mixed strategy, using the weekly Russell 3000 growth and value indices data from 1995 to 2010. The switching strategy made use of the one-step ahead forecasted return for both indices and invested into the index with higher risk-adjusted forecasted return for each time interval. The mixed strategy lead to a mean-variance optimisation problem, in which the optimal weights for each index were calculated using the estimated drifts and variance.

We compared both HOHMM strategies with the HMM-based approach. For certain levels of transaction costs, the HOHMM-based strategies outperform the HMM-based strategies in terms of the higher differences of log-return between the tested strategy and the pure strategies. The HOHMM switching strategy never gives the worst performance in the time interval considered. The evolution of the optimal weights represents the investors' reaction to regime-switching in the market. And thus the mixed strategy has a lower variance of the return. Performance comparisons of the four portfolio strategies with the benchmark using Sharpe ratio, Jensen's alpha and AR were presented. When compared to the benchmark, which is the pure Russell 3000 index strategy, both HOHMM strategies have higher risk-adjusted return. The switching strategy has higher marginal and relative returns than the benchmark. Furthermore, the bootstrap analysis with different transaction costs demonstrates that with 5 bps transaction cost, both HOHMM-based strategies have higher return and are more stable than the benchmark in terms of higher values in the performance measures and smaller confidence intervals. In the case of 30 bps transaction costs, the HOHMM strategies still have higher returns, but the switching strategy is less stable and the mixed strategy is more stable than the benchmark.

References

1. Abell, M.L., Braslton, J.P.: Introductory Differential Equations: With Boundary Value Equations. Academic Press, California (2009)
2. Ammann, M., Verhofen, M.: The effect of market regimes on style allocation. Financ. Markets Portf. Mgmt **20**, 309–337 (2006)
3. Ang, A., Bekaert, G.: International asset allocation with regime shifts. Rev. Financ. Stud. **15**, 1137–1187 (2002)
4. Ang, A., Bekaert, G.: Timing and diversification: a state-dependent asset allocation approach. Financ. Anal. J. **60**, 1137–1187 (2004)
5. Bauer, R., Haerden, R., Molenaar, R.: Timing and diversification: a state-dependent asset allocation approach. J. Invest. **13**, 72–80 (2004)
6. Bulla, J., Bulla, I.: Stylised facts of financial time series and hidden semi-Markov models. Comput. Stat. Data Anal. **51**, 2192–2209 (2006)
7. Bulla, J., Mergner, S., Bulla, I., Sesboüé, A., Chesneau, C.: Markov switching asset allocation: do profitable strategies exist? Journal of Asset Management **12**, 310–321 (2011)
8. Ching, W., Ng, M.K., Fung, E.S.: Higher-order multivariate Markov chains and their applications. Linear Algebra Appl. **428**, 492–507 (2008)
9. Ching, W.K., Siu, T.K., Li, L.M.: Pricing exotic options under a higher-order hidden Markov model. J. Appl. Math. Decis. Sci. **2007**, 1–15 (2007). Article ID 18014
10. Couillard, M., Davison, M.: A comment on measuring the Hurst exponent in financial time series. Physica. A, Statistical Mechanics and Its Applications **348**, 404–418 (2005)
11. Dajcman, S.: Time-varying long-range dependence in stock market returns and financial market disruptions—a case of eight European countries. Appl. Econ. Lett. **19**, 953–957 (2012)
12. Efron, B., Tibshirani, R.J.: An Introduction to the Bootstrap. CRC Press, Boca Raton, FL (1994)
13. Elliott, R.J., Aggoun, L., Moore, J.B.: Hidden Markov Models: Estimation and Control. Springer, New York (1995)
14. Elliott, R.J., Siu, T.K., Badescu, A.: On mean-variance portfolio selection under a hidden Markovian regime-switching model. Econ. Model. **27**, 678–686 (2010)
15. Elliott, R.J., vander Hoek, J.: An application of hidden markov models to asset allocation problems. Finance Stoch. **1**, 229–238 (1997)
16. Erlwein, C., Mamon, R., Davison, M.: An examination of HMM-based investment strategies for asset allocation. Appl. Stoch. Models Bus. Ind. **27**, 204–221 (2011)
17. Graflund, A., Nilsson, B.: Dynamic portfolio selection: the relevance of switching regimes and investment horizon. Eur. Financ. Manag. **9**, 179–00 (2003)
18. Guidolin, M., Timmermann, A.: Asset allocation under multivariate regime switching. J. Econ. Dyn. Control **31**, 3503–3544 (2007)
19. Hess, M.K.: Timing and diversification: a state-dependent asset allocation approach. Eur. J. Financ. **12**, 189–204 (2006)
20. Hurst, H.E.: Long-term storage capacity of reservoirs. Trans. Am. Soc. Civ. Eng. **116**, 770–770 (1951)
21. Lobato, I.N., Savin, N.E.: Real and spurious long memory in stock market data. J. Bus. Econ. Stat. **16**, 261–266 (1988)
22. MacDonald, I.L., Zucchini, W.: Hidden Markov Model and other Models for Discrete-Valued Time Series. Chapman & Hall, London (1997)

23. Maheu, J.: Can GARCH models capture long-range dependence? Stud. Nonlinear Dyn. Econom. **9**, 1 (2005)
24. Markowitz, H.M.: Portfolio selection. J. Finance **7**, 77–91 (1952)
25. McCarthy, J., DiSario, R., Saraoglu, H., Li, H.C.: Tests of long-range dependence in interest rates using wavelets. Q. Rev. Econ. Finance **44**, 180–189 (2004)
26. Merton, R.C.: Optimum consumption and portfolio rules in a continuous-time model. J. Econ. Theory **3**, 373–413 (1971)
27. Du Preez, J.A., Barnard, E., Weber, D.M.: Efficient high-order hidden Markov modeling. In: Proceedings of the International Conference on Spoken Language Processing, pp. 2911–2914 (1998)
28. Ray, B., Tsay, R.: Long range dependence in daily stock volatilities. J. Bus. Econ. Stat. **18**, 254–262 (2000)
29. Rydén, T., Teräsvirta, T., Åsbrink, S.: Stylised facts of daily return series and the hidden Markov model. J. Appl. Econ. **13**, 217–244 (1998)
30. Samuelson, P.: Lifetime portfolio selection by dynamic stochastic programming. Rev. Econ. Stat. **51**, 239–46 (1969)
31. Siu, T.K., Ching, W.K., Fung, E., Ng, M., Li, X.: A high-order Markov-switching model for risk measurement. Comput. Math. Appl. **58**, 1–10 (2009)
32. Solberg, J.: Modelling Random Processes for Engineers and Managers. Wiley & Sons, Inc., New Jersey (2009)
33. Xi, X., Mamon, R.: Parameter estimation of an asset price model driven by a weak hidden Markov chain. Econ. Model. **28**, 36–46 (2011)
34. Yu, S., Liu, Z., Squillante, M.S., Xia, C.H., Zhang, L.: A hidden semi-Markov model for web workload self-similarity. In: 21st IEEE International Performance, Computing, and Communications Conference, pp. 65–72 (2002)