A MARKOV REGIME-SWITCHING ARMA APPROACH FOR HEDGING STOCK INDICES

CHAO-CHUN CHEN
WEN-JEN TSAY*

This study considers the hedging effectiveness of applying the N-state Markov regime-switching autoregressive moving-average (MRS-ARMA) model to the S&P-500 and FTSE-100 markets. The distinguishing feature of this study is to incorporate the observations of serially correlated stock returns into the hedging analysis. To resolve the problem of N^T possible routes induced by the presence of MA parameters associated with the algorithm of Hamilton JD (1989) and a sample of size T, we propose an algorithm by combining the ideas of Hamilton JD (1989) and Gray SF (1996). We find that the hedging performances of the three proposed MRS-MA(1) strategies herein are superior to their corresponding MRS counterparts considered in Alizadeh A and Nomikos N (2004) over the out-of-sample periods, even when we realistically track the transaction costs generated from rebalancing the hedged portfolios. © 2010 Wiley Periodicals, Inc. Jrl Fut Mark 31:165–191, 2011

We thank the editor and referee for their valuable comments and suggestions. The first author also thanks the financial support from the Social Science Research Center of Taiwan. The current version is a substantial revision of our previous study entitled "Estimating Regime-Switching ARMA Models with Extended Algorithms of Hamilton."

*Correspondence author, The Institute of Economics, Academia Sinica, Taipei, Taiwan. Tel: (886-2) 2782-2791, Fax: (886-2) 27853946, e-mail: wtsay@econ.sinica.edu.tw

Received September 2009; Accepted March 2010

■ *Chao-Chun Chen is an Assistant Professor at the Department of Finance, Tunghai University, Taichung, Taiwan.*

■ *Wen-Jen Tsay is a Research Fellow at the Institute of Economics, Academia Sinica, Taipei, Taiwan.*

The Journal of Futures Markets, Vol. 31, No. 2, 165–191 (2011)

Published online June 28, 2010 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/fut.20465

INTRODUCTION

Futures markets allow portfolio managers to hedge their risk exposures by shorting stock index futures contracts. The critical problem for hedging centers on the determination of the hedge ratio, i.e., the number of futures contracts to sell for each unit of the underlying asset on which the short hedger bears the price risk. Based on the results in Ederington (1979) and Figlewski (1984), the minimum-variance hedge ratio, β, is equivalent to the ratio of the unconditional covariance between spot and futures price changes over the variance of futures price changes:

$$\beta = \frac{Cov(\Delta S_t, \Delta F_t)}{Var(\Delta F_t)}$$

where ΔS_t and ΔF_t denote the time t price changes in spot and futures prices, respectively. In other words, a constant hedge ratio can be estimated with the following regression:

$$\Delta S_t = \mu + \beta \Delta F_t + u_t \tag{1}$$

based on the historical spot returns and futures returns.

Because ΔS_t and ΔF_t are well-known to follow time-varying distributions, Cecchetti, Cumby, and Figlewski (1988) and Kroner and Sultan (1993) suggest that the hedge ratio should be time-varying as well. This point of view has induced many scholars to employ multivariate generalized autoregressive conditional heteroscedasticity (GARCH) models to compute time-varying hedge ratios, including Park and Switzer (1995), Gagnon and Lypny (1995), and Kavussanos and Nomikos (2000). Another way of estimating time-varying hedge ratios is via Markov regime-switching (MRS) models where the relationship between ΔS_t and ΔF_t is regime-dependent. Sarno and Valente (2000) document the rationale behind the use of the MRS model for hedging analysis, where a regime-switching relationship between spot and futures returns is found in both FTSE-100 and S&P-500 stock index futures contracts. Alizadeh and Nomikos (2004) provide additional evidence that the MRS model is useful in improving hedging performance for both FTSE-100 and S&P-500 stock index futures contracts. Alizadeh, Nomikos, and Pouliasis (2008) document that the regime-dependent hedge ratios can result in significant risk reduction in energy commodities by using an MRS vector error-correction model with GARCH errors.

The common feature among the aforementioned hedging studies is that significant autocorrelations beyond lag 0 inherent in the stock index returns are not taken into account at all. Essentially, Cohen, Hawawini, Maier, Schwartz, and Whitcomb (1980) document that the observations for stock index returns containing significant serial correlations are well established in the literature. Scholes and Williams (1977), Cohen et al. (1980), Atchison, Butler, and Simonds (1987), and Lo and Mackinlay (1990) further address the causes of serial correlations. On the other hand, LeBaron (1992) considers the relationship between serial correlations and volatility in stock returns, while Morse (1980) and Campbell, Grossman, and Wang (1993) investigate the interaction between trading volume and serial correlations in stock returns as well. Particularly, to capture the presence of autocorrelations in financial asset returns, the usual moving-average (MA) model of order 1, MA(1), is often added for empirical applications, for example, Hamao, Masulis, and Ng (1990), Bollerslev (1987), and French, Schwert, and Stambaugh (1987), to name a few.

Howard and D'Antonio (1991) first discuss the optimal multiperiod hedge ratio when spot returns are autocorrelated and possess a form of MA process, but they do not address the important empirical hedging performance. Since autocorrelations and regime-switching behaviors are found in ΔS_t and ΔF_t, an MRS autoregressive-moving average (MRS-ARMA) model is a natural candidate to improve hedging effectiveness. However, the presence of MA parameters in the N-state MRS framework makes the possible routes of states running from time 1 to time T expand exponentially to be N^T under Hamilton's (1989) approach. The first contribution of this study is to provide an easily used algorithm to estimate the N-state MRS-ARMA(p, q) models by modifying the algorithm of Hamilton (1989). The simulations show that the bias of the estimation by the EHG algorithm is very close to zero and the associated RMSE decreases with the increasing values of the sample size, revealing that the likelihood-based estimator based on our algorithm possesses a well-defined asymptotic behavior. The second contribution of this study presents that the out-of-sample hedging performances of the three proposed MRS-MA(1) models are superior to their corresponding MRS counterparts in both S&P-500 and FTSE-100 markets even when we track the transaction costs arising from the opening, rebalancing, and closing of futures contracts.

The remaining parts of this study are arranged as follows. The following section introduces the MRS-ARMA hedging model and develops an algorithm for a general class of N-state MRS-ARMA models. The later section investigates the finite sample performance of the proposed algorithm. Empirical Hedging analysis section demonstrates that the hedging effectiveness of the MRS-MA models is better than that of their MRS counterparts considered in Alizadeh and Nomikos (2004). The penultimate section shows that the performance of the MRS3-MA(1) model is promising even when we take the transaction cost associated with the opening, rebalancing, and closing the futures contracts into account. The final section provides the conclusion.

THE MRS-ARMA MODEL AND THE EHG ALGORITHM

The MRS-ARMA hedging model considered in this study is mainly based on the findings in Alizadeh and Nomikos (2004), where they employ two-state MRS models to address the following hedging problem:

$$\Delta S_t = \mu_{s_t} + \beta_{s_t}\Delta F_t + u_{t,s_t} \tag{2}$$

where the error term u_{t,s_t} is normally, independently and identically distributed (nid) across different regimes, or $u_{t,s_t} \sim nid(0, \sigma^2_{s_t})$, $s_t = 1, 2$ displays the unobserved market regime at time t, and the parameters μ_{s_t} and β_{s_t} are regime-dependent. Specifically, the relationship between index returns and futures returns in (2) is different under different regimes. The associated transition probability matrix is:

$$\mathrm{P} \equiv \begin{bmatrix} p_{11} & p_{21} \\ p_{12} & p_{22} \end{bmatrix} \tag{3}$$

where $p_{ij} = P(s_t = j|s_{t-1} = i)$ and $\sum_{j=1}^{N} p_{ij} = 1$ for all i. This is the first MRS (hereafter, MRS1) model studied by Alizadeh and Nomikos (2004). The notable characteristic is that the transition probabilities, p_{11}, p_{12}, p_{21}, and p_{22}, and the regime-dependent variances, σ_1^2 and σ_2^2, are time-invariant.

The second model of Alizadeh and Nomikos (2004), MRS2, relaxes the time-invariant assumption on transition probabilities. To ensure that the estimated transition probabilities are always within 0 and 1, Alizadeh and Nomikos (2004) employ a logistic function to model these time-varying probability measures:

$$p_{12,t} = \frac{1}{1 + \exp(\varphi_{0,1} + \varphi_{1,1}AB_{t-1})}, \quad p_{21,t} = \frac{1}{1 + \exp(\varphi_{0,2} + \varphi_{1,2}AB_{t-1})} \tag{4}$$

where $AB_t = (\sum_{i=0}^{3} \mathrm{Basic}_{t-i})/4$ denotes the average basis over the last four weeks. Accordingly, the transition probability matrix becomes:

$$\mathrm{P} \equiv \begin{bmatrix} p_{11,t} & p_{21,t} \\ p_{12,t} & p_{22,t} \end{bmatrix}.$$

Note that the variances, σ_1^2 and σ_2^2, remain time-invariant in the second model.
The third and the most flexible MRS model of Alizadeh and Nomikos (2004) further modifies the time-varying information set into the original σ_1^2 and σ_2^2 as follows:

$$\sigma^2_{s_t,t} = \exp(\lambda_{0,s_t} + \lambda_{1,s_t}AB_{t-1}). \tag{5}$$

This model is named as the MRS3 model.

The major contribution of this study is to incorporate the widely observed autocorrelations within stock index returns into the MRS hedging analysis. We model the relationship between ΔS_t and ΔF_t with the following MRS-ARMA model:

$$\Phi_{s_t}(L)(w_t - \mu_{s_t}) = \Theta_{s_t}(L)\sigma_{s_t}v_t = \Theta_{s_t}(L)\varepsilon_t,\ w_t = \Delta S_t - \beta_{s_t}\Delta F_t \tag{6}$$

where $\varepsilon_t \sim nid(0, \sigma_{s_t}^2)$, and L is the usual lag operator. The state variable s_t can assume only an integer value of 1, 2, . . . , N, and its transition probability matrix is:

$$P \equiv \begin{bmatrix} p_{11} & p_{21} & \cdots & p_{N1} \\ p_{12} & p_{22} & \cdots & p_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ p_{1N} & p_{2N} & \cdots & p_{NN} \end{bmatrix} \tag{7}$$

where $p_{ij} = P(s_t = j \,|\, s_{t-1} = i)$ and $\sum_{j=1}^{N} p_{ij} = 1$ for all i. We also impose stationarity and invertibility constraints on the AR and MA polynomials within each regime, respectively:

$$\Phi_{s_t}(L) = 1 - \phi_{1,s_t}L - \cdots - \phi_{p,s_t}L^p,\ \Theta_{s_t}(L) = 1 + \theta_{1,s_t}L + \cdots + \theta_{q,s_t}L^q. \tag{8}$$

These conditions are summarized in the following Assumption 1.

Assumption 1. *For each* $s_t = 1, \ldots, N$, *(i) the roots of the polynomial* $\Phi_{s_t}(L)$ *and those of* $\Theta_{s_t}(L)$ *in (8) are all outside the unit root circle; (ii)* $\Phi_{s_t}(L)$ *and* $\Theta_{s_t}(L)$ *share no common roots; (iii)* $\sigma_{s_t} > 0$; *(iv)* v_τ *is independent of* s_t *for all* τ *and t; and (v)* $\varepsilon_t \sim nid(0,\sigma_{s_t}^2)$.

The MRS-ARMA(p, q) model is a natural extension to the Markov regime-switching autoregressive (MRS-AR) model proposed in the seminal study of Hamilton (1989). The MRS-AR model itself has been widely used in the financial data. Particularly, Engel and Hamilton (1990), Engel (1994), and Bollen, Gray, and Whaley (2000) find Markov-switching behavior in foreign exchange data. Pagan and Schwert (1990) adopt MRS models for stock returns. The above-mentioned studies are all based on the algorithm of Hamilton (1989). This implies that they cannot consider the potential presence of MA parameters in the data-generating process (DGP) if Hamilton's (1989) approach is employed, because the possible routes of states running from time 1 to T expand exponentially to be N^T when we want to filter out the sequence $\{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_T\}$ for conducting the associated maximum likelihood estimation (MLE). Consequently, before discussing the empirical usefulness of applying the MRS-ARMA model to the hedging analysis, we need to resolve this estimation issue. An algorithm is thus developed by combining the ideas of Hamilton (1989) and

Gray (1996), and we name it as the extended Hamilton-Gray (EHG) algorithm in the following.

Before illustrating the details of the EHG algorithm, let us define the notation used throughout this study. Denote $W_t \equiv (w_1, w_2, \ldots, w_t)^T$ as a column vector containing the observations in (6) from time 1 to time t. The column vector ñ $= (\mu_1, \ldots, \mu_N, \sigma_1, \ldots, \sigma_N, \phi_{1,1}, \ldots, \phi_{p,1}, \phi_{1,2}, \ldots, \phi_{p,2}, \ldots, \phi_{1,N}, \ldots, \phi_{p,N}, \phi_{2,1}, \ldots, \phi_{p,N}, \theta_{1,1}, \ldots, \theta_{q,1}, \theta_{1,2}, \ldots, \theta_{q,2}, \ldots, \theta_{1,N}, \ldots, \theta_{q,N})^T$ and the transition probabilities p_{ij} consist of the parameters characterizing the conditional density function (cdf) of w_t. The parameters ñ and the transition probabilities p_{ij} are stacked into one column vector ζ.

Let $l = \text{Max}(p, q)$ and define a state variable s_t^* to characterize the regime path from time $t-1$ to t as follows:

$$\begin{array}{lll} s_t^* = 1 & \text{if} & s_t = 1, s_{t-1} = 1, \ldots \text{ and } s_{t-l} = 1 \\ s_t^* = 2 & \text{if} & s_t = 2, s_{t-1} = 1, \ldots \text{ and } s_{t-l} = 1 \\ \vdots & & \vdots \\ s_t^* = N^{l+1} & \text{if} & s_t = N, s_{t-1} = N, \ldots \text{ and } s_{t-l} = N. \end{array} \tag{9}$$

The $(N^{l+1} \times N^{l+1})$ transition probability matrix of s_t^*, P^*, is composed of the transition probabilities p_{ij} in (7):

$$P^* \equiv \begin{bmatrix} p_{11}^* & p_{21}^* & \cdots & p_{N^{l+1}1}^* \\ p_{12}^* & p_{22}^* & \cdots & p_{N^{l+1}2}^* \\ \vdots & \vdots & \ddots & \vdots \\ p_{1N^{l+1}}^* & p_{2N^{l+1}}^* & \cdots & p_{N^{l+1}N^{l+1}}^* \end{bmatrix} \tag{10}$$

where $p_{ij}^* = P(s_t^* = j \mid s_{t-1}^* = i)$. In other words, we do not trace the whole past history of s_t to extract ε_t in order to conduct the MLE. Instead, we only trace up to l lagged observations of w_t to compute the conditional expectation of the associated lagged errors. The choice of $l = \text{Max}(p, q)$ ensures that we have enough observations to compute these q conditional expectations.

The accuracy of our approximation method can be improved with a larger value of l. For example, $l = \text{Max}(p, q) = 4$ is chosen for an MRS-ARMA(4, 2) model, but we may use $l = 5$ or other larger values to implement the estimation procedure. The method of choosing l allows us to deal with the N^{l+1} possible regime paths based on the recursive algorithm of Hamilton (1989). See Hamilton (1994b, p. 3067) for the illustrations of s^* and P^* under the set-up, where $N = p = 2$ and $q = 0$.

As mentioned previously, we cannot exactly extract ε_t to conduct the MLE given that we only trace up to l lagged observations of w_t. One strategy is to follow the idea of Gray (1996) by replacing $\{\varepsilon_t, \ldots, \varepsilon_{t-q+1}\}$ with their corresponding conditional expectations. In this study, to accommodate the presence of MA parameters, we modify the idea of Gray (1996) by replacing the sequence of $\{\varepsilon_t, \ldots, \varepsilon_{t-q+1}\}$ with their path-dependent conditional expectation. The information set employed in our algorithm is:

$$\Omega_t^\dagger \equiv (W_t, \Pi_t^\dagger, \zeta), \ \Pi_t^\dagger = \begin{bmatrix} \hat{\varepsilon}_{t|s_t^*=1,\Omega_{t-1}^\dagger} & \hat{\varepsilon}_{t|s_t^*=2,\Omega_{t-1}^\dagger} & \cdots & \hat{\varepsilon}_{t|s_t^*=N^{l+1},\Omega_{t-1}^\dagger} \\ \hat{\varepsilon}_{t-1|s_t^*=1,\Omega_{t-1}^\dagger} & \hat{\varepsilon}_{t-1|s_t^*=2,\Omega_{t-1}^\dagger} & \cdots & \hat{\varepsilon}_{t-1|s_t^*=N^{l+1},\Omega_{t-1}^\dagger} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\varepsilon}_{t-q+1|s_t^*=1,\Omega_{t-1}^\dagger} & \hat{\varepsilon}_{t-q+1|s_t^*=2,\Omega_{t-1}^\dagger} & \cdots & \hat{\varepsilon}_{t-q+1|s_t^*=N^{l+1},\Omega_{t-1}^\dagger} \end{bmatrix} \tag{11}$$

where the matrix $\Pi_t^\dagger$ contains the conditional expectation of the sequence $\{\varepsilon_t, \ldots, \varepsilon_{t-q+1}\}$ based on the path consistent with regime $s_t^* = j$ ($j = 1, 2, \ldots, N^{l+1}$) and the information set $\Omega_{t-1}^\dagger$. Each column in $\Pi_t^\dagger$ represents these conditional expectations under a specific value of s_t^*. The information set in (11) implies that the conditional expectation of the sequence $\{\varepsilon_t, \ldots, \varepsilon_{t-q+1}\}$ is updated whenever a new observation arrives.

For the calculation of $\Pi_t^\dagger$ in (11), we first note that the value of s_t^* in (9) represents a sequence of states $\{s_t, s_{t-1}, \ldots, s_{t-l}\}$. We then define $s_{t-k}(s_t^* = j)$ as the value of s_{t-k} when the regime s_t^* is j. Following the idea in page 35 of Gray (1996), the value of $\hat{\varepsilon}_{t|s_t^*=j,\Omega_{t-1}^\dagger}$ in (11) can be calculated recursively as:

$$\begin{aligned} \hat{\varepsilon}_{t|s_t^*=j,\Omega_{t-1}^\dagger} &= w_t - E(w_t|s_t^* = j, \Omega_{t-1}^\dagger) \\ &= w_t - \mu_{s_t(s_t^*=j)} - \sum_{k=1}^{p} \phi_{k,s_t(s_t^*=j)}(w_{t-k} - \mu_{s_{t-k}(s_t^*=j)}) \\ &\quad - \sum_{k=1}^{q} \theta_{k,s_t(s_t^*=j)} \hat{\varepsilon}_{t-k|s_t^*=j,\Omega_{t-1}^\dagger}, \quad \forall j = 1, 2, \ldots, N^{l+1} \end{aligned} \tag{12}$$

where

$$\begin{aligned} \hat{\varepsilon}_{t-k|s_t^*=j,\Omega_{t-1}^\dagger} &= \frac{\sum_{i=1}^{N^{l+1}} P(s_t^* = j, s_{t-1}^* = i|\Omega_{t-1}^\dagger) \times \hat{\varepsilon}_{t-k|s_{t-1}^*=i,\Omega_{t-2}^\dagger}}{P(s_t^* = jid\Omega_{t-1}^\dagger)} \\ &= \frac{\sum_{i=1}^{N^{l+1}} P(s_t^* = j|s_{t-1}^* = i) \times P(s_{t-1}^* = i|\Omega_{t-1}^\dagger) \times \hat{\varepsilon}_{t-k|s_{t-1}^*=i,\Omega_{t-2}^\dagger}}{P(s_t^* = j|\Omega_{t-1}^\dagger)} \end{aligned} \tag{13}$$

$$\forall k = 1, 2, \ldots, q; \quad j = 1, 2, \ldots, N^{l+1}$$

such that

$$P(s_t^* = j|\Omega_{t-1}^\dagger) = \sum_{i=1}^{N^{l+1}} P(s_t^* = j, s_{t-1}^* = i|\Omega_{t-1}^\dagger). \tag{14}$$

The term $P(s^*_{t-1} = i|\Omega^\dagger_{t-1})$ in (13) denotes the inference about the probability that $s^*_{t-1} = i$ based on the information set $\Omega^\dagger_{t-1}$. Specifically, the conditional expectation of lagged error terms, $\hat{\varepsilon}_{t-k|s^*_t=j,\Omega^\dagger_{t-1}}$, in (13) is calculated by integrating over all the corresponding values of possible paths based on the N^{l+1}-state Markov chain with the transition matrix defined in (10). Furthermore, all the elements in $\Pi^\dagger_t$ can be recursively calculated by (12) and (13) provided that we have $P(s^*_{t-1} = i|\Omega^\dagger_{t-1})$ and $\Pi^\dagger_{t-1}$. The Value of $P(s^*_{t-1} = i|\Omega^\dagger_{t-1})$ across i is collected into one vector, $\hat{\xi}_{t-1|t-1}$:

$$\hat{\xi}_{t-1|t-1} = \begin{bmatrix} P(s^*_{t-1} = 1\,|\,\Omega^\dagger_{t-1}) \\ P(s^*_{t-1} = 2\,|\,\Omega^\dagger_{t-1}) \\ \vdots \\ P(s^*_{t-1} = N^{l+1}\,|\,\Omega^\dagger_{t-1}) \end{bmatrix}. \tag{15}$$

Moreover, $\hat{\xi}_{t|t}$ can be found by iterating on (22.4.5) and (22.4.6) in Hamilton (1994a, p. 692) as follows:

$$\hat{\xi}_{t|t} = \frac{\hat{\xi}_{t|t-1} \odot \eta_t}{\mathbf{1}^T(\hat{\xi}_{t|t-1} \odot \eta_t)} \tag{16}$$

$$\hat{\xi}_{t+1|t} = P^* \times \hat{\xi}_{t|t} \tag{17}$$

where $\mathbf{1}$ represents an $(N^{l+1} \times 1)$ vector of ones, the symbol $\odot$ denotes an element-by-element multiplication, and η_t is the conditional density of w_t given s^*_t and $\Omega^\dagger_{t-1}$:

$$\eta_t = \begin{bmatrix} f(w_t|s^*_t = 1, \Omega^\dagger_{t-1}) \\ f(w_t|s^*_t = 2, \Omega^\dagger_{t-1}) \\ \vdots \\ f(w_t|s^*_t = N^{l+1}, \Omega^\dagger_{t-1}) \end{bmatrix} \tag{18}$$

such that

$$f(w_t|s^*_t = j, \Omega^\dagger_{t-1}) = \frac{1}{\sqrt{2\pi}\sigma_{s_t(s^*{}_t=j)}} \exp\left\{\frac{-(\hat{\varepsilon}_{t|s^*_t=j,\Omega^\dagger{}_{t-1}})^2}{2\sigma^2_{s_t(s^*_t=j)}}\right\} \tag{19}$$
$$\forall j = 1, 2, \ldots, N^{l+1}.$$

The starting value $\hat{\xi}_{1|0}$ can be set to be the vector of unconditional probabilities described in (22.2.26) of Hamilton (1994a, p. 684).

It follows that the parameters ζ can be estimated by maximizing the following log-likelihood function with respect to these unknown parameters:

$$L(\zeta) = \sum_{t=1}^{T} \log(f(w_t|\Omega^\dagger_{t-1}))$$

where

$$f(w_t|\Omega^{\dagger}_{t-1}) = \mathbf{1}^T(\hat{\xi}_{t|t-1} \odot \eta_t).$$

See (22.4.7) and (22.4.8) of Hamilton (1994a, p. 692) for the details.

With $\hat{\xi}_{t-1|t-1}$, we have a simple formula to compute $\hat{\varepsilon}_{t-k|s^*_t=j,\Omega^{\dagger}_{t-1}}$ in (13) as follows:

$$\hat{\varepsilon}_{t-k|s^*_t=j,\Omega^{\dagger}_{t-1}} = \frac{e_k\Pi^{\dagger}_{t-1} \times (P^{*T}_j \odot \hat{\xi}_{t-1|t-1})}{P^*_j \times \hat{\xi}_{t-1|t-1}}, \forall j = 1, 2, \ldots, N^{l+1} \qquad k = 1, 2, \ldots, q \tag{20}$$

where P^*_j denotes the jth row of P* in (10), and $\mathbf{e}_k$ indicates the k-th row of the $(q \times q)$ identity matrix $\mathbf{I}_q$. Therefore, the computational cost of our algorithm is almost identical to that of Hamilton's (1989) algorithm, whereby the conditional expectation of the lagged error terms is succinctly calculated with the formula in (20).

MONTE CARLO EXPERIMENT

This section illustrates the finite sample performance of the EHG algorithm for the MRS-ARMA model. Adopting the experimental design of Psaradakis and Sola (1998, p. 377) in evaluating the finite sample performance of Hamilton's (1989) algorithm when the DGP are MRS-AR(1) processes, we focus on the following 2-state MRS-ARMA(1, 1) model:

$$w_t = \mu_{s_t} + \phi(w_{t-1} - \mu_{s_{t-1}}) + \varepsilon_t + \theta\varepsilon_{t-1},\ \varepsilon_t \sim i.i.d.N(0, \sigma^2_{s_t}). \tag{21}$$

The parameters employed are as follows:

$$(\sigma^2_1, \sigma^2_2) = (1, 1.5) \tag{22a}$$

$$(\mu_1, \mu_2) = (1, 5) \tag{22b}$$

$$\phi \in \{0.6, 0.9\} \tag{22c}$$

$$(p_{11}, p_{22}) \in \{(0.95, 0.95), (0.5, 0.5)\} \tag{22d}$$

$$T \in \{100, 200, 400, 800\} \tag{22e}$$

$$\theta \in \{0.5\}. \tag{22f}$$

The parameters in (22), except for the ones in (22f), are employed in Psaradakis and Sola (1998, p. 377).

All the computations are performed with GAUSS. Two hundred replications are conducted for each specification. For each sample size T, 200 additional

TABLE I
The Finite Sample Performance of the EHG Algorithm: Bias

DGP				MLE							
p_{11}	P_{22}	ϕ	T	μ_1	μ_2	ϕ	θ	σ_1	σ_2	p_{11}	P_{22}
							$\theta = 0.5$				
0.95	0.95	0.6	100	−0.074	0.035	−0.048	0.018	−0.057	−0.028	−0.020	−0.017
			200	−0.009	0.001	−0.021	0.003	−0.022	−0.006	−0.007	−0.005
			400	0.004	0.013	−0.007	−0.003	−0.007	−0.002	−0.003	−0.003
			800	−0.004	0.005	−0.002	−0.006	−0.002	0.003	0.000	−0.002
		0.9	100	−0.151	−0.048	−0.042	0.013	−0.054	−0.005	−0.023	−0.028
			200	−0.036	0.024	−0.021	0.001	−0.023	−0.004	−0.008	−0.011
			400	−0.045	−0.005	−0.010	−0.005	−0.011	0.003	−0.005	−0.002
			800	−0.045	−0.032	−0.004	−0.007	−0.003	0.006	−0.001	−0.002
0.5	0.5	0.6	100	−0.003	0.000	−0.048	0.024	−0.006	−0.034	0.002	−0.016
			200	0.003	−0.002	−0.020	0.009	0.000	−0.012	0.002	−0.009
			400	0.005	0.003	−0.009	−0.001	0.009	−0.003	0.000	−0.011
			800	0.005	−0.001	−0.005	−0.004	0.012	0.002	0.002	−0.007
		0.9	100	−0.097	−0.096	−0.036	0.007	−0.013	−0.029	0.001	−0.014
			200	−0.036	−0.041	−0.019	0.004	−0.005	−0.012	0.002	−0.008
			400	−0.014	−0.016	−0.010	0.001	0.003	−0.005	−0.001	−0.012
			800	−0.008	−0.013	−0.005	−0.004	0.007	0.000	0.001	−0.006

Notes. Simulations are based on 200 replications. The DGP is the MRS-ARMA(1,1) model defined in (21) with $\theta = 0.5$. Other parameters are set as $\mu_1 = 1$, $\mu_2 = 5$, $\sigma_1^2 = 1$, and $\sigma_2^2 = 1.5$ as shown in (22). Bias is defined by the mean of estimated values minus the corresponding true parameter.

values are generated in order to obtain random starting values. The true parameters are used as the initial values for the Constrained Maximum Likelihood (CML) GAUSS program. The maximum number of iterations for each replication is 100. We confine the search of the parameters μ_1 and μ_2 within the range of (−20, 20) to ensure that the resulting estimates of these parameters are not completely unreasonable. The simulation results remain intact when this range becomes (−50, 50).

Define bias as the average estimated value minus the corresponding true parameter. Table I demonstrates that the bias is very close to zero (especially when the sample size is larger) for all specifications considered in the table. The associated root-mean-squared error (RMSE) contained in Table II also decreases with the increasing values of sample size. These observations togeth-er reveal the potential of our algorithm in estimating the MRS-ARMA models.

EMPIRICAL HEDGING ANALYSIS

This section applies the MRS-ARMA hedging model to both S&P-500 and FTSE-100 stock index futures contracts. The data comprise weekly spot and

TABLE II
The Finite Sample Performance of the EHG Algorithm: RMSE

DGP				*MLE*							
p_{11}	P_{22}	ϕ	T	μ_1	μ_2	ϕ	θ	σ_1	σ_2	p_{11}	P_{22}
							$\theta = 0.5$				
0.95	0.95	0.6	100	0.573	0.659	0.135	0.135	0.184	0.191	0.061	0.089
			200	0.392	0.340	0.083	0.099	0.087	0.099	0.027	0.028
			400	0.257	0.259	0.051	0.060	0.060	0.070	0.019	0.020
			800	0.171	0.174	0.039	0.039	0.038	0.051	0.012	0.013
		0.9	100	1.775	1.692	0.074	0.136	0.173	0.445	0.079	0.105
			200	1.199	1.133	0.045	0.089	0.101	0.116	0.032	0.047
			400	0.891	0.890	0.028	0.056	0.064	0.072	0.028	0.021
			800	0.631	0.602	0.017	0.035	0.043	0.053	0.012	0.014
0.5	0.5	0.6	100	0.438	0.462	0.141	0.164	0.133	0.165	0.083	0.084
			200	0.283	0.310	0.084	0.098	0.078	0.106	0.057	0.063
			400	0.220	0.224	0.053	0.068	0.063	0.064	0.040	0.045
			800	0.144	0.153	0.040	0.045	0.043	0.045	0.026	0.029
		0.9	100	1.723	1.744	0.074	0.110	0.116	0.148	0.075	0.078
			200	1.118	1.140	0.041	0.075	0.070	0.098	0.055	0.059
			400	0.838	0.842	0.027	0.054	0.059	0.061	0.038	0.043
			800	0.574	0.580	0.018	0.034	0.041	0.041	0.026	0.028

Notes. Simulations are based on 200 replications. The DGP is the MRS-ARMA(1,1) model defined in (21) with $\theta = 0.5$. Other parameters are set as $\mu_1 = 1$, $\mu_2 = 5$, $\sigma_1^2 = 1$, and $\sigma_2^2 = 1.5$ as shown in (22).

futures prices. The spot and futures price data are Wednesday's closing prices, ranging from May 9, 1984 to July 29, 2009 and listed in the Datastream Database. When a holiday occurs on Wednesday, Tuesday's closing price is taken instead. We follow the method of Alizadeh and Nomikos (2004) to start the in-sample estimation on May 9, 1984 and reserve one year for an out-of-sample comparison. In other words, the data used for the in-sample estimation range from May 9, 1984 to July 30, 2008 (1,264 observations), while the data for the out-of-sample hedging comparison start on August 6, 2008 and end on July 29, 2009 (52 observations). Moreover, ΔS_t and ΔF_t are calculated as the differences in the logarithms of prices multiplied by 100.

Extending the idea behind the aforementioned MRS1, MRS2, and MRS3 models, we propose the following MRS-MA(1) hedging model:

$$\Delta S_t = \mu_{s_t} + \beta_{s_t}\Delta F_t + u_{t,s_t} + \theta_{s_t}u_{t-1,s_{t-1}}. \tag{23}$$

Following the MRS models studied by Alizadeh and Nomikos (2004), the transitional probabilities and variances of the corresponding three MRS-MA(1) models are set as follows:

$$\text{MRS1-MA(1) model: } u_{t,s_t} \sim nid(0, \sigma_{s_t}^2), \quad P_t \equiv \begin{bmatrix} p_{11} & p_{21} \\ p_{12} & p_{22} \end{bmatrix}$$

$$\text{MRS2-MA(1) module: } u_{t,s_t} \sim nid(0, \sigma_{s_t}^2), \quad P_t \equiv \begin{bmatrix} p_{11,t} & p_{21,t} \\ p_{12,t} & p_{22,t} \end{bmatrix}$$

$$\text{MRS3-MA(1) module: } u_{t,s_t} \sim nid(0, \sigma_{t,s_t}^2), \quad P_t \equiv \begin{bmatrix} p_{11,t} & p_{21,t} \\ p_{12,t} & p_{22,t} \end{bmatrix}.$$

The time-varying transitional probabilities, $P_{12,t}$ and $P_{21,t}$, and variances, σ_{t,s_t}^2, are specified in (4) and (5). Note that all these MRS-MA(1) models can be easily estimated with the EHG algorithm illustrated in the earlier section.

Tables III and IV present parameter estimates from the MRS-MA(1) models for the S&P-500 and FTSE-100 stock indices, respectively. Clearly, the parameters of the MA(1) terms, θ_1 and θ_2, for all three models are significantly different from zero in both the S&P-500 and FTSE-100 markets. This explains the necessity for including an MA(1) term into the empirical hedging model so as to control the autocorrelations found in the index returns.

We note that hedgers sell a specific number of futures contracts for each unit of the underlying asset in order to reduce their price risk exposure. This implies that at any time t, hedgers possess a portfolio with a return of $(\Delta S_t - \beta_{t|t-1}\Delta F_t)$, where $\beta_{t|t-1}$ is the time t hedge ratio calculated from the time $t - 1$ information set. In other words, the hedge ratio, $\beta_{t|t-1}$, is chosen ex ante in practice. The major concern of hedgers is how well they can hedge their positions in the future. We thus employ out-of-sample hedging performance as the criterion for model comparison.

Following Alizadeh and Nomikos (2004), Lee and Yoder (2007a,b), and Alizadeh et al. (2008), we measure the relative performances among various hedging strategies generated by the OLS, constant OLS, GARCH, MRS, MRS-MA(1), and unhedged strategies by calculating the variance of the returns for the portfolios as:

$$Var(\Delta S_t - \hat{\beta}_{t|t-1}^{\text{Model}_i}\Delta F_t) \tag{24}$$

where $\hat{\beta}_{t|t-1}^{\text{Model}_i}$ denotes the estimate of the time t hedge ratio estimated from Model i based on the time $t - 1$ information set. The hedging strategy producing the smallest variance is the most desirable one which effectively helps hedgers avoid future price movements. In the following empirical studies, the time t hedge ratio for the OLS hedging strategy is estimated based on the OLS model in Equation (1) and the information set at time t, whereas the constant OLS hedging strategy adopts a fixed hedge ratio estimated based on the information set in the beginning of the hedging period. In other words, the constant OLS hedging strategy requires no rebalancing once the hedged portfolio is constructed.

TABLE III
Estimated Parameters of Various Models for S&P-500

	MRS1	*MRS2*	*MRS3*	*MRS1-MA(1)*	*MRS2-MA(1)*	*MRS3-MA(1)*
α_1	0.0501	0.0504	0.0683	0.0494	0.0647	0.0622
	(0.0076)	(0.0076)	(0.0090)	(0.0057)	(0.0056)	(0.0052)
α_2	−0.0854	−0.0771	−0.3905	−0.1734	−0.4134	−0.2719
	(0.0335)	(0.0297)	(0.0574)	(0.0565)	(0.0400)	(0.0244)
β_1	0.9893	0.9890	0.9766	0.9843	0.9793	0.9777
	(0.0044)	(0.0045)	(0.0037)	(0.0038)	(0.0038)	(0.0026)
β_2	0.9173	0.9251	0.9154	0.9081	0.9050	0.9569
	(0.0144)	(0.0127)	(0.0166)	(0.0166)	(0.0179)	(0.0075)
θ_1				−0.3130	−0.4338	−0.4715
				(0.0328)	(0.0305)	(0.0362)
θ_2				−0.7012	−1.5459	−1.3781
				(0.1447)	(0.1375)	(0.1133)
p_{11}	0.8209			0.8383		
	(0.0249)			(0.0255)		
p_{22}	0.6331			0.3274		
	(0.0725)			(0.0946)		
$\varphi_{0,1}$		1.2887	−3.8259		1.1181	0.4387
		(0.2898)	(2.4613)		(0.2480)	(0.2121)
$\varphi_{1,1}$		−0.5890	−4.5058		−1.9617	−2.7606
		(0.5605)	(1.8880)		(0.4477)	(0.3805)
$\varphi_{0,2}$		−0.6043	−1.9916		−2.5389	−1.6482
		(0.5179)	(0.6011)		(0.5999)	(0.3482)
$\varphi_{1,2}$		−2.6695	0.6456		−1.5172	−1.5437
		(1.0047)	(0.3836)		(1.0503)	(0.8329)
σ_1	0.1759	0.1721		0.1951	0.2244	
	(0.0097)	(0.0089)		(0.0099)	(0.0056)	
σ_2	0.5284	0.5196		0.5598	0.4129	
	(0.0262)	(0.0243)		(0.0337)	(0.0326)	
$\lambda_{0,1}$			−1.6559			−3.7686
			(0.4596)			(0.0913)
$\lambda_{1,1}$			−1.8971			−1.0366
			(0.1781)			(0.1170)
$\lambda_{0,2}$			−0.8775			−2.8581
			(0.1178)			(0.2003)
$\lambda_{1,2}$			−3.4742			−3.3659
			(0.0942)			(0.4986)

Notes. Figures in parentheses are standard errors. The sample period ranges from May 9, 1984 to July 30, 2008.

The hedge ratio for the GARCH hedging strategy is derived from Equation (12) of Myers (1991) with a bivariate GARCH framework. Furthermore, $\hat{\beta}_{t|t-1}^{\text{Unhedged}}$ is set to be 0 for the unhedged strategy.

We now discuss how to compute the hedge ratio for the 2-state MRS framework (without MA(1) term) when the estimates of β_{s_t} are at hand, i.e., $\hat{\beta}_1$ and $\hat{\beta}_2$. Note that the estimates of β_1 and β_2 under the MRS models are calculated with Hamilton's (1989) algorithm. Since the market regime s_t is unobserved, we follow the idea of Lee, Yoder, Mittelhammer, and McCluskey (2006)

TABLE IV
Estimated Parameters of Various Models for FTSE-100

	MRS1	*MRS2*	*MRS3*	*MRS1-MA(1)*	*MRS2-MA(1)*	*MRS3-MA(1)*
α_1	0.0294	0.0264	0.0268	0.0279	0.0309	0.0551
	(0.0093)	(0.0091)	(0.0097)	(0.0117)	(0.0077)	(0.0080)
α_2	−0.0223	−0.0196	0.0371	−0.0185	−0.0177	−0.2763
	(0.0311)	(0.0366)	(0.0378)	(0.0292)	(0.0226)	(0.0311)
β_1	0.9673	0.9669	0.9685	0.9697	0.9755	0.9589
	(0.0048)	(0.0045)	(0.0046)	(0.0048)	(0.0047)	(0.0052)
β_2	0.8628	0.8585	0.8643	0.8806	0.8865	0.9262
	(0.0127)	(0.0133)	(0.0116)	(0.0137)	(0.0112)	(0.0130)
θ_1				−0.2102	−0.2329	−0.2975
				(0.0370)	(0.0342)	(0.0351)
θ_2				−0.2271	−0.2293	−0.9443
				(0.0539)	(0.0519)	(0.4677)
p_{11}	0.9321			0.9281		
	(0.0182)			(0.0354)		
p_{22}	0.8837			0.8743		
	(0.0341)			(0.0647)		
$\varphi_{0,1}$		3.0922	3.3949		1.9364	1.2434
		(0.3438)	(0.4228)		(0.2921)	(0.2240)
$\varphi_{1,1}$		1.2429	1.3705		−0.6300	−0.7070
		(0.3617)	(0.4618)		(0.4594)	(0.2646)
$\varphi_{0,2}$		1.7003	1.7913		−0.4707	−0.2356
		(0.3383)	(0.3891)		(0.3550)	(0.2727)
$\varphi_{1,2}$		−0.0345	−0.4450		−2.1453	0.5882
		(0.1846)	(0.3976)		(0.5128)	(0.3730)
σ_1	0.2248	0.2263		0.2203	0.2016	
	(0.0098)	(0.0093)		(0.0149)	(0.0104)	
σ_2	0.6580	0.6662		0.6473	0.6244	
	(0.0276)	(0.0283)		(0.0293)	(0.0235)	
$\lambda_{0,1}$			−2.8660			−3.4169
			(0.1029)			(0.1586)
$\lambda_{1,1}$			0.2481			−0.9426
			(0.1663)			(0.1106)
$\lambda_{0,2}$			−0.5878			−1.0412
			(0.1305)			(0.2064)
$\lambda_{1,2}$			0.3034			−0.5532
			(0.1049)			(0.1229)

Notes. Figures in parentheses are standard errors. The sample period ranges from May 9, 1984 to July 30, 2008.

by determining the expected hedge ratio from the weighted average of $\hat{\beta}_1$ and $\hat{\beta}_2$ as:

$$\begin{aligned}\hat{\beta}^{MRS}_{t|t} &= E\{\hat{\beta}^{MRS}_{s_t} \mid \Omega_t\} \\ &= \hat{\beta}_1 P(s_t = 1 \mid \Omega_t) + \hat{\beta}_2 P(s_t = 2 \mid \Omega_t)\end{aligned}$$

where the weight, $P(s_t = i \mid \Omega_t)$, denotes the inference about the probability that $s_t = i$ based on the information set Ω_t. Note that Ω_t is defined as:

$$\Omega_t \equiv (W_t, \zeta)$$

in which W_t contains the observations from time 1 to time t, and ζ is a column vector including all corresponding parameters. However, Ω_t does not include the lagged innovation sequence which is required for the EHG algorithm. Given this set-up, the one-step-ahead forecasts of the hedge ratio under the MRS models are:

$$\begin{aligned}\hat{\beta}^{MRS}_{t+1|t} &= E\{\hat{\beta}^{MRS}_{s_{t+1}}|\Omega_t\} \\ &= \hat{\beta}_1 P(s_{t+1} = 1|\Omega_t) + \hat{\beta}_2 P(s_{t+1} = 2|\Omega_t).\end{aligned} \tag{25}$$

When computing the hedge ratios $\hat{\beta}^{MRS-MA(1)}_{t|t-1}$ from the 2-state MRS-MA(1) models, the state variable s^*_t characterizing the regime path from time $t-1$ to time t becomes:

$$\begin{aligned}
s^*_t &= 1 \quad \text{if} \quad s_t = 1, s_{t-1} = 1 \\
s^*_t &= 2 \quad \text{if} \quad s_t = 2, s_{t-1} = 1 \\
s^*_t &= 3 \quad \text{if} \quad s_t = 1, s_{t-1} = 2 \\
s^*_t &= 4 \quad \text{if} \quad s_t = 2, s_{t-1} = 2.
\end{aligned}$$

It follows that the inference for the probability of $s_t = 1$ based on the information set $\Omega^\dagger_{t-1}$ is the sum of $P(s^*_t = 1|\Omega^\dagger_{t-1})$ and $P(s^*_t = 3|\Omega^\dagger_{t-1})$, and the inference for the probability of $s_t = 2$ is the sum of $P(s^*_t = 2|\Omega^\dagger_{t-1})$ and $P(s^*_t = 4|\Omega^\dagger_{t-1})$. Accordingly, the one-step-ahead out-of-sample forecast of the MRS-MA(1) strategy is:

$$\begin{aligned}\hat{\beta}^{MRS-MA(1)}_{t+1|t} &= E\{\hat{\beta}^{MRS-MA(1)}_{s_{t+1}}|\Omega^\dagger_t\} \\ &= \hat{\beta}_1[P(s^*_{t+1} = 1|\Omega^\dagger_t) + P(s^*_{t+1} = 3|\Omega^\dagger_t] \\ &\quad + \hat{\beta}_2[P(s^*_{t+1} = 2|\Omega^\dagger_t) + P(s^*_{t+1} = 4|\Omega^\dagger_t)].\end{aligned} \tag{26}$$

Herein, $P(s^*_{t+1} = i|\Omega^\dagger_t)$, $i = 1, 2, 3, 4$, are elements in $\hat{\xi}_{t+1|t}$ and can be found in (17).

Table V displays the empirical out-of-sample performance of various hedging strategies. In addition to portfolio variances, this table also presents the variance reduction of various hedging strategies with respect to the unhedged portfolio. Based on the one-year out-of-sample hedging performance, the MRS3-MA(1) hedging strategy produces the smallest portfolio variances on an ex ante basis and outperforms its competing models in terms of portfolio variance reduction for both S&P-500 and FTSE-100 markets, although these improvements are limited. As expected, the MRS1-MA(1) model provides the poorest variance reduction among the MRS-MA(1) models due to the extra

TABLE V
One-Year Out-of-Sample Hedging Effectiveness of MRS-MA(1) Models

	S&P500		*FTSE100*	
	Variance[a]	*Variance Reduction wrt Unhedged Position*[b]	*Variance*[a]	*Variance Reduction wrt Unhedged Position*[b]
Panel A: One-year out-of-sample hedging effectiveness[c]				
Unhedged	22.7881	–	16.9723	–
OLS	0.1872	99.179%	0.1790	98.945%
Constant OLS[d]	0.1877	99.176%	0.1849	98.911%
GARCH(1,1)	0.1999	99.123%	0.1385	99.184%
MRS1	0.1998	99.123%	0.1641	99.033%
MRS2	0.1917	99.159%	0.1591	99.063%
MRS3	0.1866	99.181%	0.1531	99.098%
MRS1-MA(1)	0.1861	99.183%	0.1537	99.094%
MRS2-MA(1)	0.1834	99.195%	0.1458	99.141%
MRS3-MA(1)	0.1826	99.199%	0.1365	99.196%
Panel B: Two-year out-of-sample hedging effectiveness[e]				
Unhedged	14.0275	–	12.9667	–
OLS	0.1499	98.931%	0.1755	98.647%
Constant OLS[d]	0.1508	98.925%	0.1873	98.556%
GARCH(1,1)	0.1540	98.902%	0.1217	99.061%
MRS1	0.1557	98.890%	0.1521	98.827%
MRS2	0.1511	98.923%	0.1494	98.848%
MRS3	0.1489	98.939%	0.1431	98.896%
MRS1-MA(1)	0.1481	98.944%	0.1394	98.925%
MRS2-MA(1)	0.1490	98.938%	0.1379	98.937%
MRS3-MA(1)	0.1458	98.961%	0.1304	98.994%

[a]Variance stands for the variance of the hedged portfolio calculated based on Equation (24).

[b]Variance reduction wrt unhedged position is calculated by: [Var(unhedged position) − Var(Model$_i$)]/Var(unhedged position). It shows the variance reduction of various strategies relative to the unhedged position.

[c]The one-year out-of-sample period is from August 6, 2008 to July 29, 2009 (52 observations).

[d]The constant OLS hedging strategy adopts a fixed hedge ratio estimated based on the OLS model in Equation (1). This strategy requires no rebalancing once it has been put in place.

[e]The two-year out-of-sample period is from August 8, 2007 to July 29, 2009 (104 observations).

restrictions imposed on the parameter space. This phenomenon is shared within the three MRS models as well.

As shown in Panel A of Table V, the variance of the hedged position is 0.1877 in the S&P-500 market when a constant OLS hedging strategy is adopted. It follows that the standard deviation is 0.433. For the MRS3-MA(1) model, the variance of the hedged portfolio is 0.1826, which means that the standard deviation is 0.427. One anonymous referee suggests another way to explain the economic significance of our findings. Suppose that the returns for the hedged portfolio follow a symmetrically normal distribution, and the initial value of the spot position to be hedged is 1 million. This implies that the 95% confidence interval of the price change in the hedged portfolio from the MRS3-MA(1) model is about \$117.6 narrower than that from the constant OLS strategy at either end per week. For the FTSE-100 market, the 95% confidence interval of the price change in the hedged portfolio from the MRS3-MA(1) hedging strategy is \$1,195.6 narrower than that from the constant OLS hedging method at either end per week.

To check whether the results in Panel A of Table V are sensitive to the choice of a one-year out-of-sample period, we proceed with robust checking by conducting a two-year out-of-sample hedging comparison in Panel B. Particularly, the data used in Panel B of Table V for the in-sample estimation range from May 9, 1984 to August 1, 2007 (1,212 observations), while the data for the out-of-sample hedging comparison start on August 8, 2007 and end on July 29, 2009 (104 observations). As we find in Panel A, all the MRS-MA(1) models outperform their corresponding MRS counterparts in Panel B, respectively. In addition, the MRS3-MA(1) method remains to generate the largest variance reduction among the three MRS-MA(1) strategies.

For the S&P-500 market, Panel B of Table V reveals that the MRS3-MA(1) strategy outperforms all its competitors. For the FTSE-100 market, however, the GARCH(1, 1) strategy performs the best, and the MRS3-MA(1) method is the second best of all approaches. This finding indicates that the MRS-MA model performs well even under longer out-of-sample periods. It also shows that incorporating the serially correlated returns, regime-switching behaviors, and GARCH properties into one empirical hedging model might be a fruitful agenda for future study.

MEASURING HEDGING PERFORMANCE BY TRACKING TRANSACTION COSTS

Earlier section employs portfolio variance reduction to compare the relative performance among various hedging strategies. Since the transaction cost generated from rebalancing the hedged portfolio is different under different hedging strategies, this section realistically compares the relative performance among different strategies by tracking the transaction costs throughout the one-year out-of-sample periods. For this reason, we analyze the economic benefits of hedging methods via the mean-variance utility function used in Kroner and Sultan (1993), Lafuente and Novales (2003), Alizadeh and Nomikos (2004), and Alizadeh et al. (2008) as follows:

$$E_tU(x_{t+1}) = E_t(x_{t+1}) - \kappa Var_t(x_{t+1}) \tag{27}$$

where κ is the degree of risk aversion ($\kappa > 0$) of the individual and x_{t+1} represents the return from the hedged portfolio. Following Alizadeh and Nomikos (2004) and Alizadeh et al. (2008), we set κ to be 4.

To calculate the real-time return, x_{t+1} in (27), from the hedged portfolio across out-of-sample periods, we first calculate the dollar return from the portfolio based on the trading practice of futures markets and then adjust it for the transaction costs. The last step converts the after-cost dollar return to the rate of return. As mentioned above, the weekly data used for the one-year out-of-sample hedging comparison range from August 6, 2008 to July 29, 2009. Thus, the out-of-sample hedging strategy involves opening a futures position on July 30, 2008, rebalancing the hedged portfolio in the following 51 weeks, and closing out all futures contracts on July 29, 2009. For the ease of reference, we denote July 30, 2008 as time 0, August 6, 2008 as time 1, and the end of out-of-sample hedging comparison as time 52.

Suppose that at time 0 a hedger has a well-diversified equity portfolio worth MV_0, and he plans to hedge the spot position by using its associated stock index futures contracts. Given the value of MV_0, the time t value of the equity portfolio is determined by:

$$MV_t = MV_{t-1} \times \frac{S_t}{S_{t-1}}, \quad \forall t = 1, 2, \ldots, 52 \tag{28}$$

where S_t is the level value of the equity index at time t. Based on Equation (3.5) of Hull (2008), the appropriate number of futures contracts to be short for hedging the equity portfolio is calculated as:

$$\mathcal{Q}_t = \frac{\hat{\beta}_{t+1|t}^{\mathrm{Model}_i} \times MV_t}{F_t \times M}, \quad \forall t = 0, 1, \ldots, 51, \tag{29}$$

where F_t is the time t futures price, and M is the contract size or multiplier specified in the futures contracts. According to the regulations of Chicago Mercantile Exchange (CME) and London International Financial Futures and Options Exchange (LIFFE), each S&P-500 futures contract is for delivery of 250 USD (US dollar) times the index, whereas each FTSE-100 futures contract is for delivery of 10 GBP (Great Britain pound) times the index. Thus, the multipliers for S&P-500 futures contracts and FTSE-100 futures contracts are 250 USD and 10 GBP, respectively. Since the hedger closes out all futures contracts at the end of the hedge, the value of $\mathcal{Q}_{52}$ is set at zero. Note that the market participants are not allowed to trade a fractional futures contract in practice. Therefore, the number of short futures contracts to hedge the equity portfolio is Q_t, which is the nearest integer of $\mathcal{Q}_t$. In summary, at time t the hedger possesses a hedged portfolio with an equity portfolio worth MV_t and a short position in Q_t futures contracts.

Once the appropriate number of futures contracts, i.e., Q_t, for the short hedge is decided, the number of futures contracts traded to rebalance the hedged portfolio, A_t, is:

$$A_t = \begin{cases} Q_0 & \text{when } t = 0 \\ Q_t - Q_{t-1} & \text{when } t = 1, 2, \ldots, 52. \end{cases} \tag{30}$$

Assume that the transaction cost of futures contracts is $c\%$ of the contract value. The transaction fee for rebalancing the futures position is then calculated as:

$$C_t = c\% \times |A_t| \times F_t \times M, \quad \forall t = 0, 1, \ldots, 52 \tag{31}$$

where $|A_t|$ represents the absolute value of A_t. By taking the transaction cost into account, the return from the hedged portfolio, x_{t+1}, is derived as:

$$x_{t+1} = \frac{(MV_{t+1} - MV_t) - Q_t \times (F_{t+1} - F_t) \times M - C_t}{MV_t} \times 100, \quad \forall t = 0, 1, \ldots, 50. \tag{32}$$

At the end of the hedge, there is still a transaction fee arising from the closure of the futures position. The amount of this transaction cost is denoted as C_{52}, acting as a reduction of the portfolio return at the end of the hedge, i.e.,

$$x_{t+1} = \frac{(MV_{t+1} - MV_t) - Q_t \times (F_{t+1} - F_t) \times M - C_t - C_{t+1}}{MV_t} \times 100 \quad \text{when } t = 51. \tag{33}$$

To conveniently compare the economic benefits of hedging strategies, we assume that the initial value of the spot position to be hedged is 250 million USD and 250 million GBP for the S&P-500 and FTSE-100 markets, respectively. The relatively large spot position is designed for the convenience of observing an integer value of change in futures positions. In fact, the preceding assumption of the initial spot position, MV_0, will be shown not to qualitatively affect the results of the hedging performance comparison.

Figures 1 and 2 display the time series plot for the number of futures positions, Q_t, in the S&P-500 and FTSE-100 markets, respectively. We observe that the movement of Q_t associated with the GARCH strategy could be very different from the other two methods. As shown in Figure 1, the patterns of futures positions based on the MRS3-MA(1) and GARCH methods move in opposite directions around time 10 and 40. Similarly, for the FTSE-100 market, Figure 2 presents a very different movement of Q_t around time 30.

To further highlight the difference between the MRS3 and MRS3-MA(1) hedging strategies, we display the associated number of futures positions for the S&P-500 and FTSE-100 markets in Figures 3 and 4, respectively. Essentially, the futures position based on the MRS3-MA(1) strategy is less volatile than that of the MRS3 strategy in both markets. This indicates that the cost for rebalancing the portfolio is less for the MRS-MA model than that for

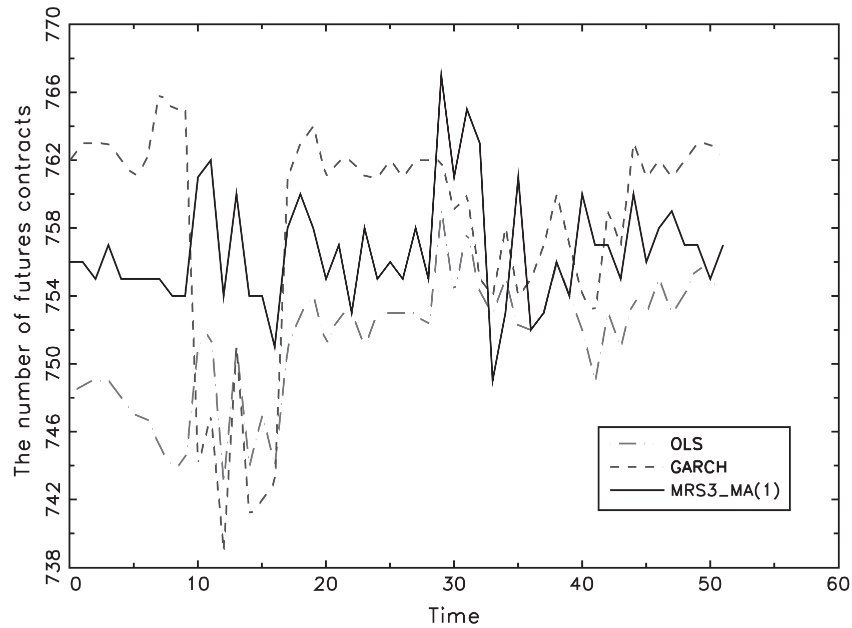

FIGURE 1

The number of short futures contracts for hedging the S&P-500 spot position based on the OLS, GARCH, and MRS3-MA(1) methods. The initial value of the spot position to be hedged, i.e., MV_0, is 250 million USD.

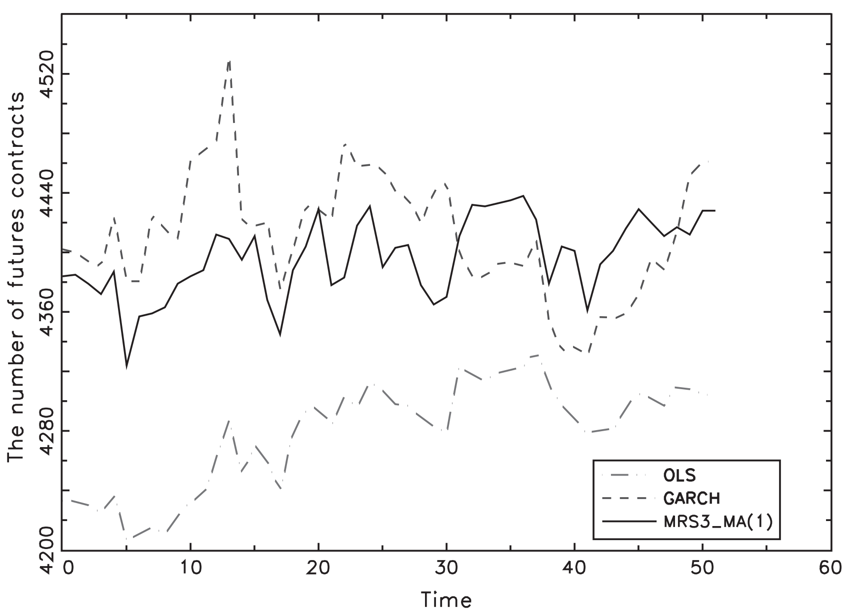

FIGURE 2

The number of short futures contracts for hedging the FTSE-100 spot position based on the OLS, GARCH, and MRS3-MA(1) methods. The initial value of the spot position to be hedged, i.e., MV_0, is 250 million GBP.

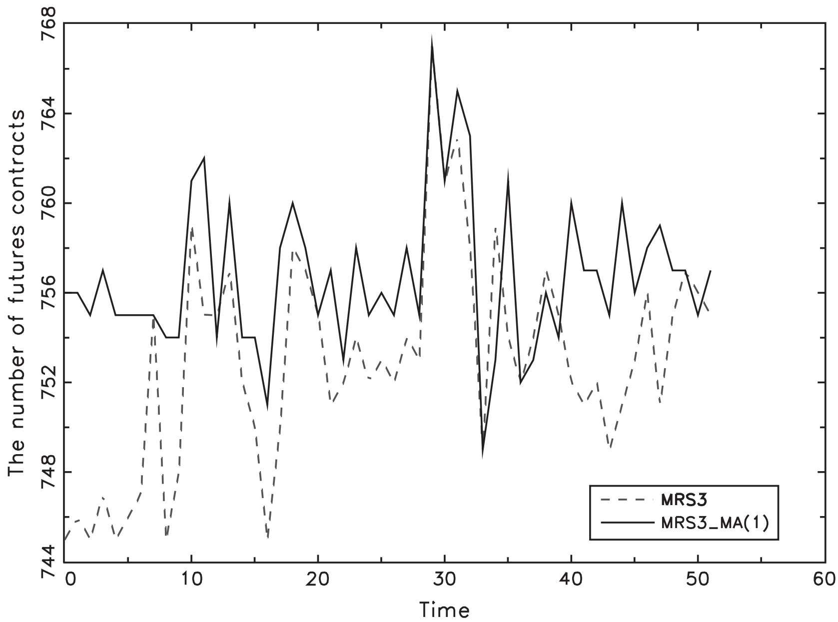

FIGURE 3
The number of short futures contracts for hedging the S&P-500 spot position based on the MRS3 and MRS3-MA(1) methods. The initial value of the spot position to be hedged, i.e., MV_0, is 250 million USD.

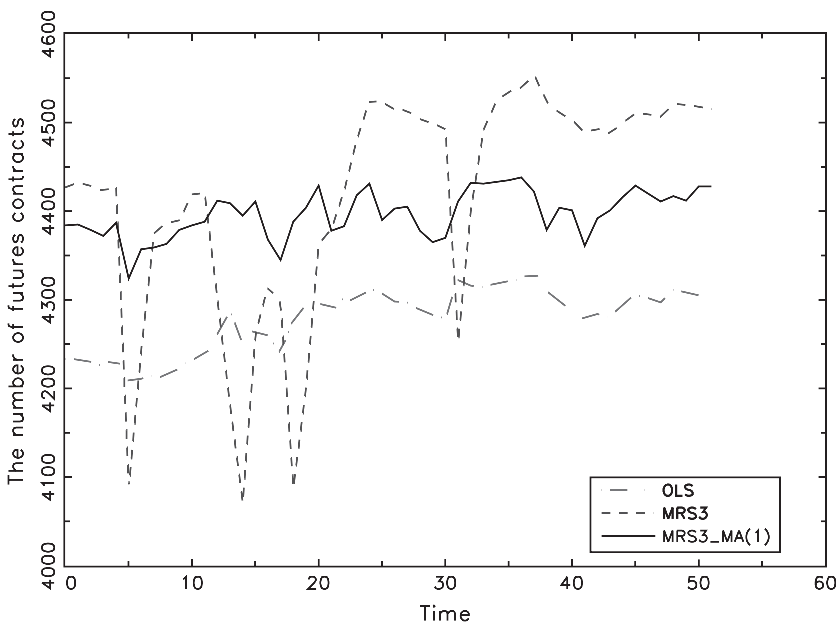

FIGURE 4
The number of short futures contracts for hedging the FTSE-100 spot position based on the OLS, MRS3, and MRS3-MA(1) methods. The initial value of the spot position to be hedged, i.e., MV_0, is 250 million GBP.

the MRS counterpart. This observation clearly implies that a fair comparison among various hedging strategies should explicitly take the transaction costs into account. Moreover, Figures 1 and 2 show that the number of futures contracts from the OLS model seems to be usually lower than those of the MRS3-MA(1) model. This phenomenon stands out very clearly in the FTSE-100 market. On the other hand, by using the MRS3 model as the benchmark, the OLS hedge ratio is not always lower than the corresponding MRS3 hedge ratio for the FTSE-100 market. This indicates that the value of the hedge ratio might change when the autocorrelations within index returns are taken into account.

In the literature, Park and Switzer (1995) and Alizadeh and Nomikos (2004) suggest that the transaction costs in the S&P-500 and FTSE-100 markets are typically between 0.010 and 0.015% of the contract value. We thus assume the transaction cost $c\%$ as 0.010, 0.015, 0.050, and 0.200% for the out-of-sample hedging comparison. The empirical results are contained in Tables VI and VII.

As expected, the average return of the hedged portfolio, $E_t(x_{t+1})$, is found to decrease with an increase in transaction cost. This implies that the weekly utility improvements of the dynamic MRS3-MA(1) hedging strategy with respect to the OLS counterpart (measured by $E_tU(x_{t+1}^{MRS3-MA(1)}) - E_tU(x_{t+1}^{OLS}))$ might deteriorate under a higher transaction cost scenario. This conjecture is clearly borne out in Table VI and VII. In particular, Table VI shows that the weekly utility improvement of the MRS3-MA(1) method with respect to the OLS strategy is 0.0329 in the S&P-500 market if the transaction cost is 0.010%. When the transaction cost increases to be 0.200%, the corresponding utility improvement changes to be 0.0319. We also observe that the constant OLS hedging model outperforms the OLS hedging strategy in the S&P-500 market, whereas it is not superior to the OLS hedging strategy in the FTSE-100 market. Nevertheless, in both Tables VI and VII, the MRS3-MA(1) hedges bring the best economic benefits among all competing models for both futures markets, although the transaction cost is as big as 0.200%.

We now investigate the impacts of the initial spot position MV_0 on the weekly utility of hedging strategies and illustrate the results in Table VIII. The values of MV_0 considered are 50, 100, and 200 million, respectively, while the transaction cost is fixed at 0.050%. Qualitatively, the size of the initial spot position does not have a strong impact on the hedging performance comparison under various configurations considered in Table VIII, mainly because the MRS3-MA(1) hedging strategy still ranks the best among all hedging strategies. This is another evidence showing the potential of the MRS3-MA(1) hedge strategy for both the S&P-500 and FTSE-100 markets.

TABLE VI

Out-of-Sample Hedging Effectiveness Under the Presence of Transaction Costs for the S&P-500 Market[a]

	$c\% = 0.010\%$			$c\% = 0.015\%$			$c\% = 0.050\%$			$c\% = 0.200\%$		
	Mean[b]	*Variance*[c]	*Weekly Utility*[d]	*Mean*[b]	*Variance*[c]	*Weekly Utility*[d]	*Mean*[b]	*Variance*[c]	*Weekly Utility*[d]	*Mean*[b]	*Variance*[c]	*Weekly Utility*[d]
Unhedged	−0.4181	21.9280	−88.1299	−0.4181	21.9280	−88.1299	−0.4181	21.9280	−88.1299	−0.4181	21.9280	−88.1299
OLS	−0.0201	0.1895	−0.7781	−0.0203	0.1895	−0.7784	−0.0217	0.1899	−0.7812	−0.0278	0.1923	−0.7969
Constant OLS	−0.0158	0.1881	−0.7684	−0.0160	0.1882	−0.7687	−0.0173	0.1885	−0.7713	−0.0229	0.1908	−0.7861
GARCH(1,1)	−0.0096	0.1992	−0.8064	−0.0098	0.1993	−0.8068	−0.0113	0.1997	−0.8100	−0.0175	0.2025	−0.8274
MRS1	−0.0143	0.2018	−0.8215	−0.0145	0.2019	−0.8219	−0.0161	0.2022	−0.8250	−0.0229	0.2048	−0.8420
MRS2	−0.0131	0.1930	−0.7851	−0.0133	0.1931	−0.7856	−0.0148	0.1935	−0.7886	−0.0212	0.1961	−0.8058
MRS3	−0.0155	0.1886	−0.7697	−0.0157	0.1886	−0.7701	−0.0172	0.1890	−0.7730	−0.0234	0.1915	−0.7893
MRS1-MA(1)	−0.0160	0.1874	−0.7654	−0.0163	0.1874	−0.7658	−0.0177	0.1878	−0.7688	−0.0241	0.1904	−0.7857
MRS2-MA(1)	−0.0155	0.1851	−0.7558	−0.0157	0.1851	−0.7562	−0.0171	0.1854	−0.7589	−0.0233	0.1878	−0.7746
MRS3-MA(1)	−0.0117	0.1834	−0.7452	−0.0119	0.1834	−0.7456	−0.0134	0.1838	−0.7485	−0.0197	0.1863	−0.7650

[a]The out-of-sample period is from August 6, 2008 to July 29, 2009 (52 observations).

[b]Mean stands for the average of the return from the hedged portfolio, i.e., $E(x_{t+1})$, where x_{t+1} is calculated based on Equations (32) and (33). The initial value of the spot position to be hedged, i.e., MV_0, is 250 million USD.

[c]Variance stands for the variance of the return from the hedged portfolio, i.e., $\mathrm{Var}(x_{t+1})$, where x_{t+1} is defined by Equations (32) and (33).

[d]Weekly utility is the average weekly utility calculated by Equation (27) based on assumptions of a coefficient of risk aversion of 4.

TABLE VII

Out-of-Sample Hedging Effectiveness Under the Presence of Transaction Costs for the FTSE-100 Market[a]

	c% = 0.010%			c% = 0.015%			c% = 0.050%			c% = 0.200%		
	Mean[b]	*Variance*[c]	*Weekly Utility*[d]	*Mean*[b]	*Variance*[c]	*Weekly Utility*[d]	*Mean*[b]	*Variance*[c]	*Weekly Utility*[d]	*Mean*[b]	*Variance*[c]	*Weekly Utility*[d]
Unhedged	−0.2550	16.4156	−65.9173	−0.2550	16.4156	−65.9173	−0.2550	16.4156	−65.9173	−0.2550	16.4156	−65.9173
OLS	−0.0110	0.1768	−0.7180	−0.0111	0.1768	−0.7182	−0.0125	0.1768	−0.7196	−0.0182	0.1777	−0.7291
Constant OLS	−0.0035	0.1886	−0.7580	−0.0036	0.1886	−0.7582	−0.0049	0.1886	−0.7594	−0.0102	0.1896	−0.7684
GARCH(1,1)	0.0064	0.1385	−0.5477	0.0062	0.1386	−0.5480	0.0048	0.1388	−0.5502	−0.0014	0.1406	−0.5636
MRS1	0.0061	0.1626	−0.6444	0.0058	0.1627	−0.6448	0.0040	0.1629	−0.6475	−0.0035	0.1649	−0.6631
MRS2	−0.0001	0.1576	−0.6307	−0.0004	0.1577	−0.6310	−0.0021	0.1579	−0.6337	−0.0095	0.1598	−0.6488
MRS3	−0.0016	0.1518	−0.6088	−0.0018	0.1518	−0.6091	−0.0035	0.1521	−0.6117	−0.0108	0.1540	−0.6269
MRS1-MA(1)	0.0072	0.1527	−0.6037	0.0070	0.1528	−0.6040	0.0053	0.1530	−0.6067	−0.0021	0.1550	−0.6220
MRS2-MA(1)	0.0033	0.1449	−0.5762	0.0031	0.1449	−0.5765	0.0014	0.1451	−0.5790	−0.0057	0.1470	−0.5937
MRS3-MA(1)	0.0027	0.1361	−0.5419	0.0025	0.1362	−0.5421	0.0011	0.1363	−0.5441	−0.0050	0.1379	−0.5565

[a]The out-of-sample period ranges from August 6, 2008 to July 29, 2009 (52 observations).

[b]Mean stands for the average of the return from the hedged portfolio, i.e., $E(x_{t+1})$, where x_{t+1} is calculated based on Equations (32) and (33). The initial value of the spot position to be hedged, i.e., MV_0, is 250 million GBP.

[c]Variance stands for the variance of the return from the hedged portfolio, i.e., $Var(x_{t+1})$, where x_{t+1} is defined by Equations (32) and (33).

[d]Weekly utility is the average weekly utility calculated by Equation (27) based on assumptions of a coefficient of risk aversion of 4.

TABLE VIII
Out-of-Sample Hedging Effectiveness Under Different Initial Spot Positions[a]

	Weekly Utility[b]: S&P-500			*Weekly Utility[b]: FTSE-100*		
MV_0 *(million)*	*50*	*100*	*200*	*50*	*100*	*200*
Unhedged	−88.1299	−88.1299	−88.1299	−65.9173	−65.9173	−65.9173
OLS	−0.7818	−0.7840	−0.7805	−0.7188	−0.7200	−0.7197
Constant OLS	−0.7612	−0.7739	−0.7739	−0.7557	−0.7604	−0.7604
GARCH(1,1)	−0.8064	−0.8092	−0.8112	−0.5498	−0.5496	−0.5504
MRS1	−0.8233	−0.8218	−0.8266	−0.6482	−0.6479	−0.6473
MRS2	−0.7923	−0.7886	−0.7886	−0.6340	−0.6331	−0.6337
MRS3	−0.7781	−0.7721	−0.7723	−0.6117	−0.6114	−0.6117
MRS1-MA(1)	−0.7726	−0.7711	−0.7699	−0.6054	−0.6066	−0.6070
MRS2−MA(1)	−0.7589	−0.7612	−0.7578	−0.5800	−0.5792	−0.5789
MRS3-MA(1)	−0.7512	−0.7448	−0.7483	−0.5443	−0.5437	−0.5445

[a]The out-of-sample period is from August 6, 2008 to July 29, 2009 (52 observations).

[b]Weekly utility is the average weekly utility calculated by Equation (27) based on assumptions of a coefficient of risk aversion of 4 and a transaction cost of 0.050%.

CONCLUSIONS

This study applies a class of MRS-ARMA models to improve the hedging effectiveness of stock index futures contracts by incorporating the autocorrelations and regime-switching behaviors into empirical hedging models. To resolve the N^T exploding regime paths problem induced by the presence of MA terms, we develop an EHG algorithm based on the method of Hamilton (1989) and the idea of Gray (1996) in order to estimate the MRS-ARMA hedging models. The simulations show that the bias of the estimation by the EHG algorithm is very close to zero and the associated RMSE decreases with the increasing values of the sample size, revealing that the likelihood-based estimator based on our algorithm possesses a well-defined asymptotic behavior.

We also apply the EHG algorithm to compute the hedge ratios of three MRS-MA(1) models in both the S&P-500 and FTSE-100 markets. The results indicate that the hedging performances of the MRS-MA(1) strategies are superior to those of their corresponding MRS counterparts considered in Alizadeh and Nomikos (2004) in terms of out-of-sample variance reduction. When we take the transaction cost associated with opening, rebalancing, and closing the futures contracts into account, the MRS3-MA(1) model is found to slightly outperform the other models in a one-year out-of-sample hedging comparison for both markets.

One potential extension of this study is to incorporate the autocorrelations, regime-switching behaviors, and GARCH properties into one empirical hedging model. The EHG algorithm needs to be modified to accommodate the presence of GARCH effects. That will be left for our future study.

BIBLIOGRAPHY

Alizadeh, A., & Nomikos, N. (2004). A Markov regime switching approach for hedging stock indices. Journal of Futures Markets, 24, 649–674.

Alizadeh, A., Nomikos, N., & Pouliasis, P. (2008). A Markov regime switching approach for hedging energy commodities. Journal of Banking & Finance, 32, 1970–1983.

Atchison, M. D., Butler, K. C., & Simonds, R. R. (1987). Nonsynchronous security trading and market index autocorrelation. Journal of Finance, 42, 111–118.

Bollen, N. P. B., Gray, S. F., & Whaley, R. E. (2000). Regime switching in foreign exchange rates: Evidence from currency option prices. Journal of Econometrics, 94, 239–276.

Bollerslev, T. (1987). A conditionally heteroskedastic time series model for speculative prices and rates of return. Review of Economics and Statistics, 69, 542–547.

Campbell, J. Y., Grossman, S. J., & Wang, J. (1993). Trading volume and serial correlation in stock returns. Quarterly Journal of Economics, 108, 905–939.

Cecchetti, S. G., Cumby, R. E., & Figlewski, S. (1988). Estimation of the optimal futures hedge. Review of Economics and Statistics, 70, 623–630.

Cohen, K. J., Hawawini, G. A., Maier, S. F., Schwartz, R. A., & Whitcomb, D. K. (1980). Implications of microstructure theory for empirical research on stock price behavior. Journal of Finance, 35, 249–257.

Ederington, L. H. (1979). The hedging performance of the new futures markets. Journal of Finance, 34, 157–170.

Engel, C. (1994). Can the Markov switching model forecast exchange rates? Journal of International Economics, 36, 151–165.

Engel, C., & Hamilton, J. D. (1990). Long swings in the dollar: Are they in the data and do markets know it? American Economic Review, 80, 689–713.

Figlewski, S. (1984). Hedging performance and basis risk in stock index futures. Journal of Finance, 39, 657–669.

French, K. R., Schwert, G. W., & Stambaugh, R. F. (1987). Expected stock returns and volatility. Journal of Financial Economics, 19, 3–29.

Gagnon, L., & Lypny, G. (1995). Hedging short-term interest risk under time-varying distributions. Journal of Futures Markets, 15, 767–783.

Gray, S. F. (1996). Modeling the conditional distribution of interest rates as a regime-switching process. Journal of Financial Economics, 42, 27–62.

Hamao, Y., Masulis, R. W., & Ng, V. (1990). Correlations in price changes and volatility across international stock markets. Review of Financial Studies, 3, 281–307.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica, 57, 357–384.

Hamilton, J. D. (1994a). Time Series Analysis. New Jersey: Princeton University Press.

Hamilton, J. D. (1994b). State-space models. In Engle, R. F., McFadden, D. L., (Eds.), Handbook of Econometrics, Vol. 4 (pp. 3039–3080). Amsterdam: North-Holland.

Howard, C. T., & D'Antonio, L. J. (1991). Multiperiod hedging using futures: A risk minimization approach in the presence of autocorrelation. Journal of Futures Markets, 11, 697–710.

Hull, J. C. (2008). Fundamentals of futures and options markets. New Jersey: Pearson Prentice Hall.

Kavussanos, M., & Nomikos, N. (2000). Hedging in the freight futures market. Journal of Derivatives, 8, 41–58.

Kroner, K., & Sultan, J. (1993). Time-varying distributions and dynamic hedging with foreign currency futures. Journal of Financial and Quantitative Analysis, 28, 535–551.

Lafuente, J., & Novales, A. (2003). Optimal hedging under departures from the cost-of-carry valuation: Evidence from the Spanish stock index futures market. Journal of Banking & Finance, 27, 1053–1078.

LeBaron, B. (1992). Some relations between volatility and serial correlations in stock market returns. Journal of Business, 65, 199–219.

Lee, H., & Yoder, J. K. (2007a). A bivariate Markov regime switching GARCH approach to estimate time varying minimum variance hedge ratio. Applied Economics, 39, 1253–1265.

Lee, H., & Yoder, J. K. (2007b). Optimal hedging with a regime-switching time-varying correlation GARCH model. Journal of Futures Markets, 27, 495–516.

Lee, H., Yoder, J. K., Mittelhammer, R. C., & McCluskey, J. J. (2006). A random coefficient autoregressive Markov regime switching model for dynamic futures hedging. Journal of Futures Markets, 26, 103–129.

Lo, A. W., & Mackinlay, A. C. (1990). An econometric analysis of nonsynchronous trading. Journal of Econometrics, 45, 181–211.

Morse, D. (1980). Asymmetrical information in securities markets and trading volume. Journal of Financial Quantitative Analysis, 15, 1129–1148.

Myers, R. J. (1991). Estimating time-varying optimal hedge ratios on futures markets. Journal of Futures Markets, 11, 39–53.

Pagan, A. R., & Schwert, G. W. (1990). Alternative models for conditional stock volatility. Journal of Econometrics, 45, 267–290.

Park, T., & Switzer, L. (1995). Bivariate GARCH estimation of the optimal hedge ratios for stock index futures: A note. Journal of Futures Markets, 15, 61–67.

Psaradakis, Z., & Sola, M. (1998). Finite-sample properties of the maximum likelihood estimator in autoregressive models with Markov switching. Journal of Econometrics, 86, 369–386.

Sarno, L., & Valente, G. (2000). The cost of carry model and regime shifts in stock index futures markets: An empirical investigation. Journal of Futures Markets, 20, 603–624.

Scholes, M., & Williams, J. (1977). Estimating betas from nonsynchronous. Journal of Financial Economics, 5, 309–327.