International Review of Economics and Finance 47 (2017) 233–254

Contents lists available at ScienceDirect

International Review of Economics and Finance

journal homepage: www.elsevier.com/locate/iref

Stock index hedging using a trend and volatility regime-switching model involving hedging cost

EnDer Su*

National Kaohsiung First University of Science and Technology, Taiwan, ROC

ARTICLE INFO

Keywords:
Stock index
Regime switch
White reality test
Hedging ratio
Hedge cost

ABSTRACT

In this study, the risk hedge between the Morgan Stanley Taiwan stock index (MSTI) and its underlying futures is analyzed regarding hedging cost under various hedge states using the trend and volatility involved regime-switching models compared with OLS and the naive method. The correlation between MSTI spot and futures is very high and the cointegration test between them is very significantly. The hedging result is evaluated using out-of-sample data and the realized variances and covariances, and is really satisfying in the subprime period or when both spot and future stay in a down state. In this stock index context, the risk aversion is determined to be 1. The optimal hedge ratio on average is around 0.878. The White test favors the regime-switching models in prediction. The trend and volatility-switching model performs particularly well in wealth increase.

1. Introduction

Stock indices fluctuate frequently because of information shocks caused by sudden events resulting from economic, political, or natural disasters. The major market crises, such as the Asian financial crisis in 1997, the U.S. subprime mortgage crisis in 2008, the 2011 earthquake in Japan, and the Greek debt crisis in 2012, have all caused chaotic aftermaths in global stock markets. Therefore, hedging a stock portfolio is necessary.

The removal of investment quotas for foreign investors in July 2003 enabled the qualified foreign investors to invest in the Taiwan Stock Exchange (TWSE) openly and feely. At the end of 2013, the total market capitalization of the Taiwan stock exchange, involving 797 listed companies, had increased up to 20 percent compared with that in 2008, amounting to around US$ 666 billion, with the daily trading value and market turnover rate reaching approximately US$ 3 billion and 119.87 percent, respectively. In this paper, the MSCI Taiwan stock index (MSTI) is regarded as a spot position owned by domestic and qualified foreign investors that is susceptible to shocks of either favorable or unfavorable news in the Taiwanese stock market. MSTI is comprised of around 70 percent Taiwanese stock market capitalization covering 100 component stocks. To hedge the MSTI, the hedge tools used are the futures of the MSTI, the STW. STW is coded and traded on the Singapore Exchange (SGX) that was formed in 1999 thanks to the merger between the Stock Exchange of Singapore (SES) and the Singapore International Monetary Exchange (SIMEX).

At present, STW is traded globally, even online, at any time. Its open interest is gradually increasing and is admired by many foreign traders, who occupy over 28 percent of the trading volume. As of the end of 2013, the averaged STW open interest and the number of trading contracts were 185,372 and 45,838, respectively, gradually increasing from 12,977 and 31,884 at the beginning of 2001. This episode indicates that STW is useful and advantageous for foreign investors. In fact, the Taiwan stock market is more

* Corresponding address: National Kaohsiung First University of Science and Technology, Department of Risk Management and Insurance, No. 2, Jhuoyue Rd., Nansih District, Kaohsiung City 811, Taiwan, ROC.

E-mail address: suender@ccms.nkfust.edu.tw.

http://dx.doi.org/10.1016/j.iref.2016.10.016
Received 25 May 2015; Received in revised form 27 October 2016; Accepted 30 October 2016
Available online 11 November 2016
1059-0560/

efficient and healthy than other Asian stock markets due to its robust economic strength and vitality, and it has survived several large financial disasters. For foreign investors, comprehending how to hedge the portfolio risk is crucial to benefit from the investment and hedge of MSTI.

The optimal hedge ratios (i.e., the best number of futures required to hedge the spot position) are calculated usually by the hedging model and the mean-variance method. However, the stock markets are nonstationary, in that anomalies, various trends, and different volatility states often occur.B, Besides, speculators who take advantage of trading signals are active in finding the momentum of price changes, while hedgers who want to insulate the risk of price changes are passive and, among them, the trading information is even asymmetrical. The aftermath can drive some hedgers to evade the futures market, as studied by Carter (1999). Besides, if bad news arrives, traders are subject to overreact more than to than news, as studied by Brooks, Henry, and Persand (2002). In consequence, the stochastic trend of returns often shifts up and down swiftly, as studied in time series momentum, and the volatility state is persistently high or low (i.e. long memory) for a period: as Hamilton (1990, 1991) has evidenced, the stock returns exhibit high and low volatility states. Hence, the conventional regression method that considers only one state equation is insufficient to describe the dynamic behavior of stock markets involving various states of trends and volatility. Since volatility is persistent in a high or low state for a while, the Markov switch becomes an appropriate way to model the phenomenon of volatility state switch. It is a nonlinear dynamic model and permits switching between two or more regimes to capture the specific dynamic patterns. Its state variables follow a first order Markov chain—i.e., the current value of the state variable depends only on its immediate past value. Hamilton (1991) and Engel (1994) have first started to apply the Markov chain to describe and predict nominal US dollar exchange rate changes based on the Bayes' theorem. Later, Hamilton and Susmel (1994) and Cai (1994) began to incorporate mixture normal likelihood function into autoregressive conditional heteroscedasticity (ARCH) model to investigate the time-varying volatility changes. McLachlan and Peel (2000) supported that the use of mixture distributions is appropriate due to their suitability and flexibility.

Evidences also show that not only correlation in an asset itself but correlation between assets is crucial to track the price returns and thus the application of vector autoregression is very popularly for studying two or more assets. Based on the aforementioned researches, this paper uses the bivariate regime-switching generalized ARCH to study the stock index hedge. Besides, as Alizadeh, Nomikos, and Pouliasis (2008), Kroner and Sultan (1993), and Lien (1996), among others, pointed out, the spread between spot and futures prices is mean reverse and a long term equilibrium relationship exists and this phenomenon is termed 'cointegration'. Hence, in our paper, the vector autoregression combining the cointegration effect becomes the vector error correction model (VECM) is incorporated into the regime-switching generalized autoregressive conditional heteroscedasticity (RS-GARCH) model to accommodate the reverse changes in the trend of returns and the different state persistences in volatility.

In this study, using the regime switch, the trend of returns with cointegration is allowed to shift between two states and the volatility can persist at either a high or low state. Once the trend of returns in cointegration and volatility in GARCH are predicted using the regime-switching hedging model, the optimal hedge ratios are subsequently computed by maximizing the expected hedging utility function while considering the hedge cost. The optimal risk preference also determined when maximizing the expected hedging utility function actually implies the amount of risk premium in market. In other words, if one can capture the true market risk premium, it follows that the return should reach its maximum with risk restricted to a constant. Meanwhile, the hedging cost would definitively reduce the hedgers' wealth, and so the use of time-varying models should not ignore the hedging cost (Lence, 1995, 1996). In addition, the realized variances and covariances are compared to the predicted variances and covariances to evaluate the hedging performance.

The contribution of this paper is exclusive compared to Alizadeh and Nomikos (2004) and Lee and Yoder (2007), because it allows for the state change of stochastic return trend when hedging and it evaluates the hedge effectiveness of the out-of-sample period instead of the in-sample period using the realized variance and covariance. Most articles evaluating the in-sample variance reduction may not reflect the hedger's true benefit. In addition, to discover the optimal risk aversion, hedge effectiveness is assessed and compared using the different risk aversions of utility function.

The remainder of the paper is structured as follows. In Section 2, the methods for obtaining optimal hedge ratios and the regime-switching models are discussed based on related articles. In Section 3, the specification of future contract and transaction costs associated with hedger behavior are described in detail. In Section 4, the sample of data, estimations of regime-switching methods for calculating hedge ratios, and measurements of hedge effectiveness are addressed in detail. In Section 5, the hedge ratios obtained using a rolling window technique and the hedge results are compared among various hedging models, measurements, periods, and states. The final section presents the concluding remarks.

2. Literature review

Investors generally want to understand the methods for obtaining the optimal hedge ratio to achieve successful hedging of portfolio risk. According to Markowitz (1952), investors wish to minimize the risk of investment. Therefore, the minimum-variance optimization method has been applied by early authors such as Ederington (1979), Johnson (1960), and Stein (1961), and subsequent authors such as Alizadeh et al. (2008), Baillie and Myers (1991), Harris and Shen (2003), Koutmos and Pericli (1998), Lien and Yang (2008), Moschini and Myers (2002), and Poomimars, Cadle, and Theobald (2003). However, this method considers neither the wealth effect, nor the hedge cost.

Investors also want to minimize risk and obtain maximum profit (Working, 1962); therefore, the mean-variance utility (MVU) method is adopted by authors such as Gagnon, Lypny, and McCurdy (1998), Heifner (1972), Kroner and Sultan (1993), and Lence (1995). The MVU method incorporates the expectation of investment return and variance into a hedging utility function involving a suitable degree of risk aversion. In addition, Frechette (2000) stated that the cost of hedging indeed affects hedging profits and ratios. Therefore, subsequent authors such as Haigh and Holt (2000, 2002), Jin and Koo (2006), Liu, Geaun, and Lei (2001), and Yun and Kim (2010) have derived hedge ratios while considering the cost of hedging. Their demonstrations of applying a hedge ratio formula verify that the hedging cost affects hedge ratios and performance levels. An alternative to the use of a utility function proposed as a prospective theory by Kahneman and Tversky (1979) is to construct a value function combining with loss probability to explain investor behavior. Mattos, Garcia, and Pennings (2008) employed this method to investigate the probability distortion in a weighting function and calculate the hedge ratios.

As Lamoureux and Lastrapes (1990) *discussed in relation to volatility persistence and structure break,* Hamilton (1990, 1991) demonstrated that stock returns exhibit high and low volatility states, which typically persist for a period. Therefore, Hamilton and Susmel (1994) proposed the implementation of an unobserved Markov chain in the ARCH model to describe the property of the volatility regime switch. However, they discovered that the regime switch involved in GARCH is not feasible because the condition variance is path-dependent. In other words, the conditional variance is the recursive solution of a previous-state random variable, and therefore the pattern of previous regimes should be considered to determine the current conditional variance. To overcome path dependence in the estimation process, Gray (1996) and Klaassen (2002) developed a method to determine the expected conditional volatility in various regimes for the regime-switching GARCH. To consider different state returns, Sarno and Valente (2005) used regime switch to model and forecast stock and futures market returns, whereas Bergman and Hansson (2004) and Ichiue and Koyama (2011) employed regime switch to explain exchange rate behavior. Instead, Lien and Tse (2001) considered the downside risk of currency exposure and applied the partial lower moments to find the optimal currency hedge ratios. They found currency futures are more appropriate in hedging than currency option.

For the estimation of conditional variance and the covariance of multivariate assets, the multivariate GARCH (MGARCH) model has been used by Bollerslev, Engle, and Wooldridge (1988), Hansson and Hordahl (1998), and Ng (1991). It was also applied to explain the spillover effect of contagion by Bae, Karolyi, and Stulz (2003), Tse and Tsui (2002), and Haqmmoudeh, Yuan, McAleer and Thomson (2010). In addition, Ang and Bekaert (2002), Haas and Mittnik (2008), Haas, Mittnik, and Paolella (2004), Honda (2003), Marcucci (2005), and Ramchand and Susmel (1998) have developed a Markov-switching MGARCH model.

The trend of returns is the time series momentum which shows the persistence of the information shocks. Speculators tend to capture the positive trends to make profit, while hedgers tend to take the opposite position to speculators, according to Moskowitz, Ooi, and Pedersen (2012). However, the trend is not permanent. It can reverse in a moment due to investors' over-reaction. In other words, structure breaks exist. Thus, considering the return trend switch in model should account for structure breaks and the alternative regime-switching model including return trend switch should provide appropriate advantage to capture the reverse change of trend.

3. Materials and hedgers

The purpose of this paper is to describe the hedge ratios and effectiveness of STW futures on Morgan Stanley Corporation's MSCI Taiwan stock index, coded MXTW in Bloomberg, by using several hedging models under various hedge states.

The MXTW index is created and constructed using the stock shares of 100 large Taiwanese companies, including most electronics companies and financial banks. The underlying futures contracts on the MSCI Taiwan index, coded as STW, began to be traded on the Singapore Exchange on January 9, 1997. Subsequently, the Taiwan Futures Exchange started to trade futures contracts on the MSCI Taiwan index on March 27, 2006, but stopped trading because of the considerable transaction cost, comprising a 0.004 percent tax rate on the contract value in addition to a NT$ 20 commission fee per contract, and less demand compared with that on the Singapore Exchange.

To hedge the risk of MXTW, the STW futures are designated as the hedge tools. They are traded based on the underlying asset of the MXTW index with a contract size of US$ 100 × the SGX MSCI Taiwan Index Futures (i.e., STW) price; contract months comprising the two nearest serial months and four quarterly months in a March, June, September, and December cycle; a tick size equal to 0.1 index points, which are equivalent to US$ 10; and a daily price limit of ± 7 percent. No trading tax is imposed for short or long STW, but the commission fee is US$ 2.85 per contract per side based on interactive brokers.

Hedgers typically consider both maximizing the expected return and minimizing the volatility of a hedged portfolio. However, the two cannot be achieved simultaneously and there must therefore be a tradeoff between them. Thus, the maximum of the mean-variance utility is established to determine the optimal hedge ratios. The utility function depends on the tradeoff between return and variance, which is the hedger's risk preference. Also, the utility function depends on the random return and unknown variance, which are necessary to be determined. Furthermore, they might exist in different states; hence the expected utility turns out to depend on the state probability, the distribution of returns, and risk aversion or preference.

Accordingly, the hedging models are used to estimate the expected conditional returns and volatilities in different states associated with a plausible risk averse utility function. The hedging model performance is evaluated based on increases in wealth or utility or decreases in variance.

4. Research method

Stock markets have been proven to be nonlinear, nonstationary, and leptokurtic in the literature. In fact, the impacts of favorable and unfavorable news are often asymmetrical in the market. Investors typically overreact to unfavorable news, whereas they are accustomed to favorable news. It implies that the assumption of different states in dynamic stock price change is more reasonable.

Noticeably, the trend of stock returns often changes reversely and quickly, and volatility does not often keep in a usual state. In fact, the trend of returns changes rapidly between up and down state and volatility is persistent for a period between high and low state. Besides, to hedge the spot price using futures, one has to consider the cointegration between both prices, as mentioned above, and use the appropriate vector error correction model to study hedging. Consequently, the major framework of two state regime-switching VECM and GARCH models (RS-VECM-GARCH) is established in this study to describe the dynamic behaviors of volatility and the stochastic trend of returns in cointegration.

Along with the main RS-VECM-GARCH model, the following three nested conjunction models are proposed: (1) regime switch for VECM only; (2) regime switch for GARCH only; and (3) no switch for VECM and GARCH. For comparisons, conventional ordinary regression and naive hedging methods are applied here, too. Hence, a total of five models, plus one naive method, are provided and evidenced in the paper. The VECM part can be used to determine the trend exhibited in the data; the GARCH part can be used to determine the volatility property of the data; and the regime-switching method can be used to identify the asymmetric effect (i.e., data properties in various states).

Nevertheless, trend or volatility estimates using regime switch are based on conditional information which reflects the temporary trend and volatility, while the conventional method reflects the expectation of the trend and volatility in the long run. The models described in (1)–(3) can be constructed accordingly, because they are the restricted versions of the RS-VECM-GARCH model.

4.1. Vector error correction model combined with the regime switch

The regime-switching VECM model for analyzing bivariate spot price (P_{S_t}) and futures price (P_{F_t}) at time t is interpreted by the trend of returns and the cointegration between $P_{S_{t-1}}$ and $P_{F_{t-1}}$. If switching between two state models is assumed to occur, it can be written as

$$\begin{aligned} \Delta\ln(P_{S_t}) &= \theta_{01}^{s_t} + \sum_{i=1}^{k}(\theta_{i1}^{s_t}\Delta\ln(P_{S_{t-i}}) + \theta_{k+i1}^{s_t}\Delta\ln(P_{F_{t-i}})) + \gamma_1^{s_t}(\ln(P_{S_{t-1}}) - \phi\ln(P_{F_{t-1}})) + \varepsilon_t^{s_t}, \\ \Delta\ln(P_{F_t}) &= \theta_{02}^{s_t} + \sum_{i=1}^{k}(\theta_{i2}^{s_t}\Delta\ln(P_{F_{t-i}}) + \theta_{k+i2}^{s_t}\Delta\ln(P_{S_{t-i}})) + \gamma_2^{s_t}(\ln(P_{S_{t-1}}) - \phi\ln(P_{F_{t-1}})) + \varepsilon_t^{s_t}, \end{aligned} \tag{1}$$

where s_t is a transition state variable at time t and k is the number of lags. The cointegration vector is $\Psi = [1 \quad \phi]$ and the speed adjustment vector is $\Upsilon^{s_t} = [\gamma_1^{s_t} \quad \gamma_2^{s_t}]$. Rewriting Eq. (1) in matrix form becomes

$$\Delta\ln(\mathbf{P}_t) = \mathbf{c}^{s_t} + \sum_{i=1}^{k}\Pi_i^{s_t}\Delta\ln(\mathbf{P}_t) + \Pi^{s_t}\ln(\mathbf{P}_{t-1}) + \varepsilon_t^{s_t}, \tag{2}$$

where Π^{s_t} is the cointegration matrix equal to $\Upsilon^{s_t'}\Psi$.

In this study, the trend of returns may remain in the "usual" state, denoted by outcome $s_t = u$, or the "down" state, denoted by outcome $s_t = d$. The state probability and the transition probability are denoted respectively as

$$\begin{aligned} \pi_{s_t t} &= P(s_t|\Omega_{t-1}) \text{ with state } s_t \in \{u, d\}, \\ p_{s_t s_{t-1}} &= P(s_t|s_{t-1}) \text{ with state } s_t \in \{u, d\}. \end{aligned} \tag{3}$$

The dynamic conditional variance covariance matrix in different states can be written as

$$\mathbf{H}_t^{s_t} = \begin{pmatrix} \sigma_{SS}^{s_t 2} & \sigma_{SF}^{s_t 2} \\ \sigma_{FS}^{s_t 2} & \sigma_{FF}^{s_t 2} \end{pmatrix}. \tag{4}$$

4.2. Regime-switching bivariate GARCH(1,1) combining with VECM

The conditional volatility cannot be directly expressed using GARCH(1,1) with regime switch, because the conditional volatility depends on the unobservable previous latent state series. Several articles—such as those authored by Gray (1996), Haas et al. (2004), Klaassen (2002), and Lee, Yoder, Mittelhammer, and McCluskey (2006)—have proposed various methods to identify the conditional switching volatility that occurs under certain latent state series. Based on these articles, the conditional switching volatility process under a latent state series is determined by integrating out the latent state variables (i.e., expecting the conditional switching volatility on the current and previous states).

In this paper, the dynamic latent state variable is assumed to have two specific state outcomes and the state probabilities are embedded within the parameters of transition probabilities. The dynamic state variable is assumed to be parallel concurrently in each period but not forward evolved, and can be determined using an iterative forward technique based on Bayes' theorem repeatedly, as described in the following section. Therefore, the sum of the log likelihood function can be derived using the parallel GARCH estimated conditional volatilities.

To simplify the estimation process, the bivariate GARCH(1,1) model is applied to estimate the conditional volatilities using the residual series as in Eq. (1), which are calculated by subtracting the estimated conditional mean from the actual return. But, establishing the multivariate GARCH model is difficult because of the positive definite requirement and is non-parsimonious because of the considerable number of parameters involved in the model. To make the model succinct, the conditional constant correlation (CCC) MGARCH proposed by Bollerslev (1990) is adopted. Practically, CCC was applied by Tse (2000), who stated that correlation in the stock market is time varying. Hence, correlation is likely to vary in different states and the switching between various state correlations can be used to describe the change in correlation structures, as suggested by Pelletier (2006). As a result, the dynamic changes of variance covariance could arise from the state changes of correlation.

Based on a Markov regime-switching GARCH model (Haas et al., 2004), the regime change in the CCC is specified as follows. Assume the state correlation between series i and j under state s_t at time t is $\rho_{ij}^{s_t}$ and the time-varying variance is $h_{ij,t}^{s_t}$; therefore,

$$h_{ij,t}^{s_t} = \rho_{ij}^{s_t}\sqrt{h_{i,t}^{s_t}h_{j,t}^{s_t}}. \tag{5}$$

Let i and j represent MXTW (i.e. S) and STW (i.e. F), respectively; thus, the dynamic conditional variances and covariances are written as

$$\begin{aligned} h_{S,t}^{s_t} &= w_S^{s_t} + \alpha_S^{s_t}\varepsilon_{t-1}^2 + \beta_S^{s_t}h_{S,t-1}^{s_t}, \\ h_{F,t}^{s_t} &= w_F^{s_t} + \alpha_F^{s_t}\varepsilon_{t-1}^2 + \beta_F^{s_t}h_{F,t-1}^{s_t}, \end{aligned} \tag{6}$$

$$\mathbf{H}_t^{s_t} = \begin{pmatrix} h_{S,t}^{s_t} & h_{S,t}^{s_t}h_{F,t}^{s_t}\rho_{SF}^{s_t} \\ h_{S,t}^{s_t}h_{F,t}^{s_t}\rho_{SF}^{s_t} & h_{F,t}^{s_t} \end{pmatrix}. \tag{7}$$

The conditional variances and covariances in matrix $\mathbf{H}_t^{s_t}$ are path dependent on the dynamic process of unknown latent random variable s_t. However, these variances and covariances are formed by two concurrent parallel sets of GARCH, each with a unique state outcome and probability. Consequently, the regime-switching conditional volatility is the aggregate of two parallel GARCH conditional volatilities weighted by the respective state probability.

4.3. The maximum likelihood estimation

Given the estimation of the trend of returns and conditional volatilities, the quasi-log likelihood function is written as

$$\ln L = \sum_{t=1}^{T}\ln f(\varepsilon_t|\Omega_{t-1}) = \sum_{t=1}^{T}\sum_{s_t\in\{u,d\}}\ln f(\varepsilon_t, s_t|\Omega_{t-1}) = \frac{1}{2}\sum_{t=1}^{T}\sum_{s_t\in\{u,d\}}(k\ \log(2\pi) + \ \log(|\mathbf{H}_t^{s_t}|) + \varepsilon_t^{s_t\prime}\mathbf{H}_t^{s_t}\varepsilon_t^{s_t})P(s_t|\Omega_{t-1}), \tag{8}$$

where $\mathbf{H}_t^{s_t}$ is the matrix of the dynamic variances and covariances, as in Eq. (7), $\varepsilon_t^{s_t}$ is the residual vector as in Eq. (1), and $P(s_t|\Omega_{t-1})$ is the conditional state probability.

If the random state variable s_t is observable, estimating the parameters using maximum likelihood estimation (MLE) is straightforward. However, s_t is a latent random variable. The plausible method of solving the likelihood function under latent state series is to adopt Bayes' theorem to determine the density of $\boldsymbol{\varepsilon}_t$ conditional on previous information Ω_{t-1}—i.e., the likelihood contribution $f(\varepsilon_t|\Omega_{t-1})$ by integrating the joint probability of $\boldsymbol{\varepsilon}_t$ and s_t with respect to $s_t = u, d$. Moreover, the joint probability of $\boldsymbol{\varepsilon}_t$ and s_t is derived from multiplying the conditional probability of $\boldsymbol{\varepsilon}_t$ given a latent state variable, either u or d, i.e., $f(\varepsilon_t|s_t, \Omega_{t-1})$, and the probability of the latent state variable, i.e., $P(s_t|\Omega_{t-1})$. Typically, the series of random state variables exhibit the property of Markov chains: the current state estimation depends on the previous state information only. Thus, the probability of the latent state variable is assumed to be derived from the multiplication of the conditional probability of the latent state variable given the previous latent state variable and the probability of the previous latent state variable as shown in Eq. (10).

Starting from the initial condition (e.g., known state variable and information set) at time $t = 0$, one of the likelihood contributions at time $t = 1$ can be constructed by jointly using only two concurrent and parallel state outcomes and probabilities, given the known condition at time $t = 0$. Subsequently, by iterating forward through time $t = 1$ to T, all likelihood contributions can be determined and the sum of the likelihood contributions is thereby acquired. The details are as follows.

To integrate out the state variable in Eq. (8), the conditional state probability must be determined accordingly. Therefore, the likelihood contribution in Eq. (8) is rewritten as

$$\begin{aligned} f(\varepsilon_t|\Omega_{t-1}) &= \textstyle\sum_{s_t\in\{u,d\}}f(\varepsilon_t, s_t|\Omega_{t-1}) = \sum_{s_t\in\{u,d\}}f(\varepsilon_t|s_t, \Omega_{t-1})P(s_t|\Omega_{t-1}) \\ &= f(\varepsilon_t|s_t = u, \Omega_{t-1})P(s_t = u|\Omega_{t-1}) + f(\varepsilon_t|s_t = d, \Omega_{t-1})P(s_t = d|\Omega_{t-1}). \end{aligned} \tag{9}$$

The joint probability of s_{t-1} and s_t is used to obtain the conditional state probability $P(s_t|\Omega_{t-1})$ which can be written as

$$P(s_t|\Omega_{t-1}) = \sum_{s_{t-1}\in\{u,d\}}P(s_t, s_{t-1}|\Omega_{t-1}) = \sum_{s_{t-1}\in\{u,d\}}P(s_t|s_{t-1})P(s_{t-1}|\Omega_{t-1}). \tag{10}$$

However, to update the likelihood contribution $f(\varepsilon_t|\Omega_{t-1})$ in Eq. (9) for the next period of time $t + 1$, $P(s_t|\Omega_{t-1})$ is first updated to obtain $P(s_t|\Omega_t)$. Hence, given ε_t is known at time t, it is

$$P(s_t|\varepsilon_t, \Omega_{t-1}) = P(s_t|\Omega_t) = \frac{f(\varepsilon_t, s_t|\Omega_{t-1})}{f(\varepsilon_t|\Omega_{t-1})} = \frac{f(\varepsilon_t, s_t|\Omega_{t-1})}{\sum_{s_t\in\{u,d\}}f(\varepsilon_t, s_t|\Omega_{t-1})} = \frac{f(\varepsilon_t|s_t, \Omega_{t-1})P(s_t|\Omega_{t-1})}{\sum_{s_t\in\{u,d\}}f(\varepsilon_t|s_t, \Omega_{t-1})P(s_t|\Omega_{t-1})}. \tag{11}$$

Then, $P(s_{t+1}|\Omega_t)$ is updated using Eq. (11). Substituting $P(s_{t+1}|\Omega_t)$ into Eq. (9), the likelihood contribution $f(\varepsilon_t|\Omega_{t-1})$ is updated to $f(\varepsilon_{t+1}|\Omega_t)$.

The unconditional state probability can be derived from the transition probability according to Hamilton (1989), as follows.

Using Eq. (10), the unconditional state probability based on the initial information is written as

$$P(s_t) = \sum_{s_{t-1}\in\{u,d\}} P(s_t, s_{t-1}) = \sum_{s_{t-1}\in\{u,d\}} P(s_t|s_{t-1})P(s_{t-1}). \tag{12}$$

As time approaches infinity, the above equation in matrix form becomes $\Pi = P\Pi$, where Π is the state probability matrix and P is the transition matrix that can be used to solve for the unconditional state probabilities. In our two state case, the result is

$$\pi_u = \frac{1 - p_{dd}}{2 - p_{uu} - p_{dd}} \text{and} \pi_d = \frac{1 - p_{uu}}{2 - p_{uu} - p_{dd}}. \tag{13}$$

To derive the MLE clearly, the following detailed steps for time t =1 are applicable.

(1) Given the information Ω_0 i.e. ε_0, $P(s_0 = u|\Omega_0) = \pi_{u0}$, and $P(s_0 = d|\Omega_0) = \pi_{d0}$, the density of the log likelihood contribution at time t =1 is

$$\begin{aligned} f(\varepsilon_1|\Omega_0) &= \textstyle\sum_{s_1\in\{u,d\}} f(\varepsilon_1, s_1|\Omega_0) = \sum_{s_1\in\{u,d\}} f(\varepsilon_1|s_1, \Omega_0)P(s_1|\Omega_0) \\ &= f(\varepsilon_1|s_1 = u, \Omega_0)P(s_1 = u|\Omega_0) + f(\varepsilon_1|s_1 = d, \Omega_0)P(s_1 = d|\Omega_0). \end{aligned} \tag{14}$$

(2) The initial state probability of the regime, i.e., $P(s_1|\Omega_0)$ is

$$P(s_1 = u|\Omega_0) = \sum_{s_0\in\{u,d\}} P(s_1 = u, s_0 = u|\Omega_0) = \sum_{s_0\in\{u,d\}} P(s_1 = u|s_0 = u)P(s_0 = u_0|\Omega_0). \tag{15}$$

Suppose that the state variable s_1 becomes u or d. $P(s_1 = u|\Omega_0)$ can be rewritten as

$$P(s_1 = u|\Omega_0) = p_{uu}\pi_{u0} + (1 - p_{dd})\pi_{d0} \text{ or } P(s_1 = d|\Omega_0) = (1 - p_{uu})\pi_{u0} + p_{dd}\pi_{d0}. \tag{16}$$

Substituting the unconditional state probability in Eq. (13) into the above equation, the result of the initial state probability becomes π_u. It means that the regime probabilities at time 1 can be set to unconditional state probabilities assuming a stationary Markov chain, because there is no information at time 0.

(3) The likelihood contribution of ε_1 is

$$\begin{aligned} f(\varepsilon_1|\Omega_0) &= \textstyle\sum_{s_1\in\{u,d\}} f(\varepsilon_1, s_1|\Omega_0) = \sum_{s_1\in\{u,d\}} f(\varepsilon_1|s_1, \Omega_0)P(s_1|\Omega_0) \\ &= f(\varepsilon_1|s_1 = u, \Omega_0)P(s_1 = u|\Omega_0) + f(\varepsilon_1|s_1 = d, \Omega_0)P(s_1 = d|\Omega_0). \end{aligned} \tag{17}$$

(4) The state probability $P(s_1 = u|\Omega_0)$ or $P(s_1 = d|\Omega_0)$ are updated using the known ε_1 as

$$P(s_1 = u|\varepsilon_1, \Omega_0) = P(s_1 = u|\Omega_1) = \frac{f(\varepsilon_1|s_1 = u, \Omega_0)P(s_1 = u|\Omega_0)}{\sum_{s_1\in\{u,d\}} f(\varepsilon_1|s_1, \Omega_0)P(s_1|\Omega_0)}. \tag{18}$$

By repeatedly iterating steps (1)–(4) from t =2 to t =T, all sample data likelihood contributions and state probabilities can be acquired using Eq. (14) to Eq. (18). Since ε_t in Eq. (18) is updated repeatedly, Eq. (18) is also known as the filter probability. Consequently, the sum of the log likelihood equation for the bivariate vector ε_t when t = 1 to T is written as

$$\begin{aligned} &\textstyle\sum_{t=1}^{T} \log f(\varepsilon_t|\Omega_{t-1}) = \sum_{s_t\in\{u,d\}} \sum_{t=1}^{T} f(\varepsilon_t, s_t|\Omega_{t-1}) = \sum_{s_t\in\{u,d\}} \sum_{t=1}^{T} \log f(\varepsilon_t|s_t, \Omega_{t-1})P(s_t|\Omega_{t-1}) \\ &= \textstyle\sum_{t=1}^{T} \log(f(\varepsilon_t|s_t = u, \Omega_{t-1})P(s_t = u|\Omega_{t-1}) + f(\varepsilon_t|s_t = d, \Omega_{t-1})P(s_t = d|\Omega_{t-1})). \end{aligned} \tag{19}$$

Once the path dependence of the recursive conditional volatility is derived and determined, the Brendt-Hall-Hall-Hausman (BHHH) or Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm can be applied to solve the optimization problem by maximizing the log likelihood function in Eq. (19) and determine the optimal parameters, including the VECM parameters$(\theta_{ij}^{s_t}, \gamma_j^{s_t})$for i=0 to 4 and j=1,2, the GARCH parameters $(w_j^{s_t}, \alpha_j^{s_t}, \beta_j^{s_t})$ for j=S, F, the constant correlations $\rho_{SF}^{s_t}$, and the transition parameters which are established to estimate the transition probabilities p_{uu} and p_{dd}. The cumulative probabilities of the standard normal distribution of the transition parameters are equal to the transition probabilities$p_{uu}$and.

Because the cointegration matrix is the multiplication of the error correction term ϕ by the speed adjustment, it does not appear possible to identify them jointly. Thus, empirically, the cointegration parameter ϕ is determined in advance by regressing the log of MXTW prices on the log of STW prices.

4.4. Hedge ratio measurement

To measure hedge effectiveness while considering hedge cost, mean-variance utility and optimization must be adopted. The utility function is used to solve the optimal hedge ratio, which should involve not only the risk aversion but also the expected utility function associated with random spot and futures prices and unknown variance that should be worked out through the regime-switching models.

Suppose that a hedged portfolio is constructed using a long position on one unit of spot that is equal to the index points of MXTW, and a short position on Q_F units (i.e., contracts) of STW futures. The hedger's expected utility at time t is determined using the expected wealth and variance of the hedged portfolio and the information set Ω_{t-1} at current time t-1. Accordingly, the hedge ratios are obtained by maximizing the hedger's utility as

$$\underset{Q_{Ft}}{Max}\ E(U_t|\Omega_{t-1}) = E(W_t|\Omega_{t-1}) - \frac{1}{2}\lambda Var(W_t|\Omega_{t-1}), \tag{20}$$

where W_t is the wealth of the hedged portfolio,$E(\cdot)$ and $Var(\cdot)$ at time t respectively represent the expectation and variance operators, and λ is the risk-aversion coefficient. According to Gagnon et al. (1988), λ is suitable to be 1 for a mildly risk-averse hedger.

The portfolio is created using long one spot and short Q_{F_t} futures, and the wealth change for the hedged portfolio at time t, concerning the cost of hedging, is written as

$$W_t = S_t - Q_{F_t}(F_t - F_{t-1}) - Q_{F_t}c_h, \tag{21}$$

where S and F denote spot and futures prices, respectively. The transaction cost of STW, which includes only the commission fee, is equal to US$ 2.85 per contract per side, as mentioned previously. Therefore, for trading STW,c_h is set to US$ 5.7 (i.e., 5.7/100 of one index point value) that would lessen the wealth of hedging portfolio.

The spot and futures prices are known at current time t-1 and, thus, the expectation and variance of wealth in Eq. (20) are written respectively as

$$E(W_t) = E(S_t) - Q_{F_t}E(F_t) + Q_{F_t}F_{t-1} - Q_{F_t}c_n, \tag{22}$$

$$Var(W_t) = \sigma^2(S_t) + Q_{F_t}^2\sigma^2(F_t) - 2Q_{F_t}\sigma(S_t, F_t), \tag{23}$$

where $\sigma^2(\cdot)$ and $\sigma(\cdot,\cdot)$ respectively denote the operators of variance and covariance. Because returns are the first order difference of prices, using returns instead of prices to express the expectation and variance as $S_t = S_{t-1}\cdot r_{S_t}$ and $F_t = F_{t-1}\cdot r_{f_t}$ is more appropriate. Thus, Eq. (23) is rewritten as

$$\begin{aligned}Var(W_t) &= \sigma^2(S_{t-1}r_{S_t}) + Q_{F_t}^2\sigma^2(F_{t-1}r_{F_t}) - 2Q_{F_t}\sigma(S_{t-1}r_{S_t}, F_{t-1}r_{F_t})\\ &= S_{t-1}^2\sigma^2(r_{S_t}) + F_{t-1}^2Q_{F_t}^2\sigma^2(r_{F_t}) - 2S_tF_tQ_{F_t}\sigma(r_{S_t}, r_{F_t}).\end{aligned} \tag{24}$$

To simplify this expression,$\sigma^2(r_{S_t})$ is replaced by $\sigma_{r_{S_t}}^2$ and the equation becomes

$$Var(W_t) = S_{t-1}^2\sigma_{r_{S_t}}^2 + F_{t-1}^2Q_{F_t}^2\sigma_{r_{F_t}}^2 - 2S_{t-1}F_{t-1}Q_{F_t}\sigma_{r_{S_t},r_{F_t}}. \tag{25}$$

Substituting Eqs. (22) and (25) into Eq. (20) gives

$$\begin{aligned}&\underset{Q_{F_t}}{Max}\ E(U_t|\Omega_t)=\\ &E(S_t) - Q_{F_t}E(F_t) + Q_{F_t}F_{t-1} - Q_{F_t}c_n - \tfrac{1}{2}\lambda\left(S_{t-1}^2\sigma_{r_{S_t}}^2 + F_{t-1}^2Q_{F_t}^2\sigma_{r_{F_t}}^2 - 2S_{t-1}F_{t-1}Q_{F_t}\sigma_{r_{S_t},r_{F_t}}\right).\end{aligned} \tag{26}$$

To maximize Eq. (26) and solve for Q_{F_t}, the first order derivative of Eq. (26) with respect to Q_{F_t} is derived as

$$\frac{dU_t}{dQ_{F_t}} = -E(F_t) + F_{t-1} - c_n - \frac{1}{2}\lambda(2Q_{F_t}F_{t-1}^2\sigma_{r_{F_t}}^2 - 2S_{t-1}F_{t-1}\sigma_{r_{S_t},r_{F_t}}). \tag{27}$$

Set Eq. (27) equal to zero, and the solution is the optimal hedge ratio as

$$Q_{F_t}^* = \frac{R_t + \lambda S_{t-1}F_{t-1}\sigma_{r_{S_t},r_{F_t}}}{\lambda F_{t-1}^2\sigma_{r_{F_t}}^2}\text{and}R_t = F_{t-1} - E(F_t) - c_n. \tag{28}$$

At current time t-1, the spot and futures prices are known in Eq. (28), but the expectation of futures price at time t is not. Therefore, the futures price should be predicted using the regime-switching VECM model, as expressed in Eq. (1) and the covariance $\sigma_{r_{S_t},r_{F_t}}$ and variance $\sigma_{r_{F_t}}^2$ are estimated using regime-switching GARCH(1,1) to acquire the optimal hedge ratio $Q_{F_t}^*$.

According to Eq. (28), the hedge position increases either as the futures price and transaction cost decrease, or as the correlation and the relative volatility between spot and futures prices increases, and vice versa for the hedge position decrease. The former is the return effect and the latter is the volatility effect. If market volatility feedback effect exists (French, Schwert, & Stambaugh, 1987), the increase in $Q_{F_t}^*$ due to the former effect should be accompanied by a decrease due to the latter effect until the optimal hedging position is reached. The risk aversion also reveals the tradeoff magnitude between the two effects, implying that one unit of risk deduction is for a λ unit of return compensation.

4.5. Hedge evaluation

In fact, to reflect the hedger's real benefit, the hedge performance should be evaluated through the out-of-sample period but not the in-sample period. However, to evaluate the utility-related hedge performance, the realized price, returns, variance, and covariance in the out-of-sample period should be substituted into Eqs. (22) and (25) to compute the realized wealth and its variance. According to Andersen, Bollerslev, Diebold, and Labys (2003), the h-day realized variance (RV) (i.e., quadratic variation) is

$$RV_t = \sum_{i=1}^{h/\Delta} \mathbf{r}'_{t-h+i\Delta}\mathbf{r}^{t-h+i\Delta}, \tag{29}$$

and according to Hayashi and Yoshida (2005), the h-day realized covariance (RC) (i.e., cross variation) is

$$RC_{r_{St},r_{Ft}} = \sum_{j=1}^{h/\Delta} r_{S,t-h+j\Delta} r_{F,t-h+j\Delta}, \tag{30}$$

where, for MXTW and STW, $\mathbf{r}_{t-h+i\Delta}$ denotes a 2×1 vector of logarithmic returns in one $1/\Delta$ period, RV_t denotes a 2×1 vector of return variances in h period, and$RC_{r_{St},r_{Ft}}$ denotes logarithmic return covariance in h period. Because of our weekly data, h is equal to 5 and Δ is equal to 1.

Using Eqs. (29) and (30), the h-day realized variances and covariances of MXTW and STW at each week over the out-of-sample period are computed. Substituting the computations and the estimated hedge ratio in Eq. (28) into Eq. (20), the realized wealth, variance of wealth, and utility are hence available for hedge evaluation. Three types of hedge evaluation are addressed as follows.

The hedge effectiveness is measured using the incremental wealth increase (IWI), augmented variance decrease (AVD), and incremental utility increase (IUI), which are written as follows:

$$\text{IWI} = \frac{W_t^{Hg} - W_t^{UHg}}{W_t^{UHg}} = \frac{W_t^{Hg}}{W_t^{UHg}} - 1, \tag{31}$$

$$\text{AVD} = \frac{V_t^{Hg} - V_t^{UHg}}{V_t^{UHg}} = \frac{V_t^{Hg}}{V_t^{UHg}} - 1, \tag{32}$$

$$\text{IUI} = \frac{U_t^{Hg} - U_t^{UHg}}{U_t^{UHg}} = \frac{U_t^{Hg}}{U_t^{UHg}} - 1, \tag{33}$$

where W, V, U, Hg, and UHg respectively denote wealth, variance, utility, hedge, and no hedge. V_t^{Hg} is the variance of the hedged portfolio and the other notations such as V_t^{UHg},W_t^{Hg},W_t^{UHg},U_t^{Hg}, and U_t^{UHg} are expressed accordingly.

5. Empirical results

Five models and one naive method are established for trend of returns and volatility estimations and hedge evaluation. The basic regime-switching model is designated as Model 1, which involves switching for both the trend of returns and volatility. The other three models are the restricted forms of Model 1, as represented by Models 2, 3, and 4: Model 2 is used for the trend of returns switch only; Model 3 is used for volatility or GARCH switch only; and Model 4 has no switch, which is analogous to the VECM combined with GARCH. The conventional hedge method is the ordinary least squares (OLS) regression, whereas the naive method that uses the one to one position for spot and futures is also applied for comparisons.

5.1. Data description

The correlation between spots and near-month futures prices is definitively higher compared with that between spots and far-month futures prices. Therefore, to hedge MXTW successfully, the time of STW maturity is selected as the nearby month. Both the

Table 1
Data description.

	Min	Max	Mean	Stdev	Skewness	Kurtosis	Jarque-Bera Test
In-Sample Data							
MXTW	-0.1407	0.1617	1.475E-03	0.0360	0.0516	5.7153	93.146
P value			0.4869		0.7205	0.000	0.000
STW	-0.1503	0.1797	1.475E-03	0.0400	0.1399	5.6533	89.797
P value			0.5323		0.3325	0.000	0.000
Correlation	0.9782						
Out-of-Sample Data							
MXTW	-0.1274	0.1327	-5.173E-04	0.0331	-0.3508	4.7474	46.880
P value			0.7858		0.0128	0.000	0.000
STW	-0.1298	0.1376	-5.479E-04	0.0346	-0.4648	4.9424	61.071
P value			0.7836		0.0009	0.000	0.000
Correlation	0.9745						

Note: The in-sample (08/09/2001–05/31/2007) and the out-of-sample (06/07/2007–06/26/2013) data comprise 288 and 302 weekly observations, respectively.

MXTW index and nearby STW futures price data are collected from the Taiwan Economic Journal (TEJ) database. The STW data in the TEJ database are continued by simply connecting the first nearby contract price following maturity (i.e., the last settlement date) to the second nearby contract price, as the futures prices roll over between the first two nearby contracts.

Because of the transaction cost considered in the hedge position, hedging the spot position daily is not recommended. Therefore, the hedge horizon is set to a week. The in-sample period is set from August 9, 2001 to May 31, 2007, yielding a total of 288 weekly observations that are subsequently used to estimate the trend and volatility by using regime-switching models. The out-of-sample period is set from June 7, 2007 to June 26, 2013, yielding a total of 302 weekly observations. To account for the probable structure break, the models are estimated rolling week by week to analyze the model estimations and forecasts for the out-of-sample period. The window is fixed at the same size as that used in the in-sample period. The fixed size window is then rolled one step forward (i.e., a week) repeatedly throughout the entire out-of-sample period and, thus, 302 forecasts are created.

Furthermore, to determine and compare the hedge ratios and performances during high-risk events such as the U.S. subprime mortgage crisis and the Greek debt crisis, the out-of-sample data are separated into three datasets: the subprime data spanning from June 1, 2007 to March 10, 2009; the Greek debt data spanning from December 7, 2009 to March 9, 2011; and the post risk data spanning from November 11, 2011 to June 26, 2013. These three periods are denoted as periods 1, 2, and 3 for later recall, while 'all periods' covers periods 1 to 3.

Table 1 reports the basic statistical properties of the weekly logarithmic returns of the MXTW and STW datasets for the in-sample and out-of-sample periods. The standard deviations of the STW futures are higher than those of the MXTW spots during either the in-sample or out-of-sample period. The futures are observed to have higher volatility than that of the spots. For both datasets and periods, the results of the excess kurtosis tests and non-normally distributed Jarque-Bera tests are found to be strongly significant, with a near zero P value. The skewness test results are significant, with a negative value for both spots and futures (i.e., −0.3508 and −0.4648) during the out-of-sample period, but the results are not significant and with a positive value during the in-sample period. It indicates that both datasets are fat tailed and exhibit increased peakedness in the out-of-sample period. The mean of returns in the in-sample period is positive, while it is negative in the out-of-sample period because of covering the huge risk events, but the test results for the mean equal to zero are not rejected when large P values at least greater than .4869 are presented. Accordingly, the trend of returns is changeable.

Also, the minimum, maximum, and standard deviation of logarithmic returns indicate that the returns are not invariable and should be taken into account in modeling. Therefore, regime switch is used to identify the trend and volatility changes between various states. Note that the correlation between MXTW and STW is gauged to reach around 0.9745. The high correlation indicates that STW is an effective tool for tracking and hedging MXTW.

Fig. 1 shows the movements of indices and the basis risk between MXTW and STW. The basis risk is the MXTW index minus the STW index and is observed to fluctuate between -10.64 and 17.39. This range of basis risk is equal to 28.03 × \$100/point = \$2,803 for a futures contract, indicating that profit or loss might reach \$2,803 × the number of futures contracts. If the futures position is sufficiently large, the basis risk becomes extremely high. Therefore, the hedge ratios should be optimized cautiously.

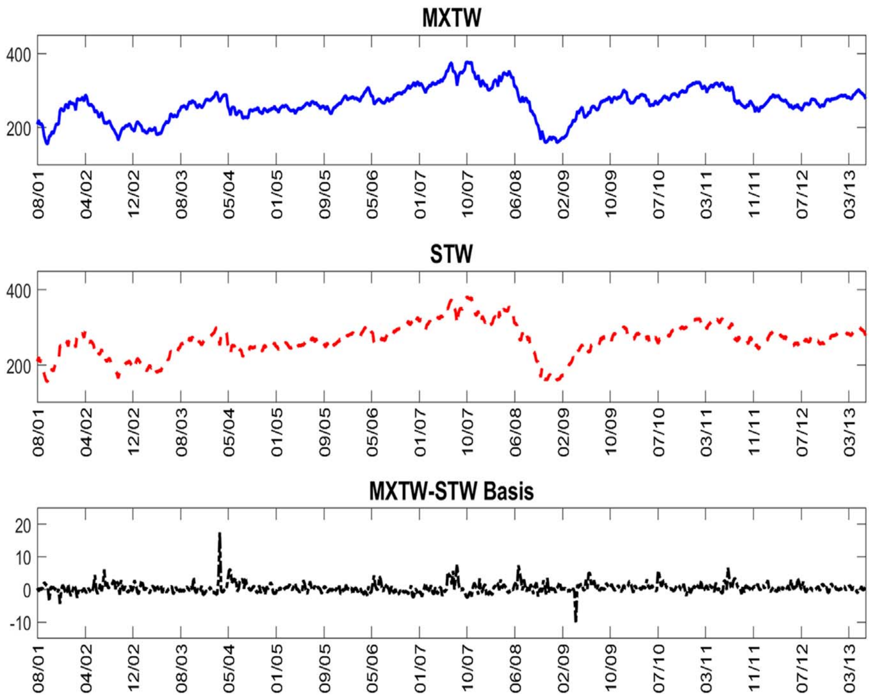

Fig. 1. MXTW and STW prices and spreads.

Table 2
The tests of MXTW and STW return series.

	Autocorrelation Test						Unit Root Test						ARCH Test	
	MXTW			STW			MXTW			STW			MXTW	STW
Lag	AC	Q-Stat	*P* value	AC	Q-Stat	*P* value	ADF(Return)			ADF(Return)			*F* test	*F* test
1	-0.003	0.0045	0.947	-0.045	1.1984	0.274	-25.336			-24.292			21.233	22.824
2	0.008	0.040	0.980	-0.029	1.7007	0.427	Critical level			Critical level			*P* value	*P* value
3	0.037	0.8707	0.832	0.038	2.5389	0.468	1%	5%	10%	1%	5%	10%	0.000	0.000
4	0.073	4.0262	0.402	0.065	5.0409	0.283	-3.441	-2.866	-2.569	-3.441	-2.866	-2.569	*LM*	*LM*
5	-0.015	4.1561	0.527	-0.033	5.6838	0.338	ADF(Price)			ADF(Price)			20.56226	22.04214
6	0.025	4.5381	0.604	0.006	5.7055	0.457	-1.630			-1.850			*P* value	*P* value
7	0.037	5.3464	0.618	0.057	7.6389	0.366	Critical level			Critical level			0.000	0.000
8	0.043	6.4556	0.596	0.052	9.288	0.319	1%	5%	10%	1%	5%	10%		
9	-0.046	7.7175	0.563	-0.059	11.405	0.249	-3.453	-2.871	-2.572	-3.453	-2.871	-2.572		

Note: The tests are performed using all 590 observations, including in-sample and out-of-sample data.

Table 2 reports the properties of the logarithmic returns of the MXTW and STW series. Neither return series has a significant autocorrelation or unit root. The unit root test examines whether the data series is a dynamic stationary process and the augmented Dickey Fuller test (ADF; Dickey and Fuller, 1981) is used typically to do the work. However, both price series are tested and found having a significant unit root. Thus, to maintain a stationary price series, the price series should be differenced once to transform them into a return series for modeling. According to the significant *F* and *LM* statistics in Table 2, both return series have presented a significant ARCH effect, suggesting that conditional time-varying volatility exists.

5.2. Model estimation and prediction

The spread between the MXTW and STW prices is suspected to have a cointegration effect. Like most pair spot and futures index behavior, the prices of MXTW and STW must have some specific pattern of co-movement in stationary distance. Before testing the cointegration, the lag structure of the error correction model should be determined. Panel A of Table 3 reports the lag exclusion Wald test. The test result for joined MXTW and STW lags of two endogenous variables in the model indicates the second lag is the

Table 3
The cointegration test between MXTW and STW.

Panel A. Lag structure Wald test

	dLag 1	dLag 2	dLag 3	dLag 4	dLag 5
MXTW	4.570767	6.158975	7.817064	7.108982	11.75789
	0.101	0.046	0.02	0.029	0.003
STW	5.507816	7.049617	8.727365	8.240045	11.40652
	0.064	0.029	0.013	0.016	0.003
Joint Test	7.00341	10.71661	9.250699	9.295753	11.92845
	0.136	0.03	0.055	0.054	0.018

Panel B. Cointegration test

Trace Test				
Cointegration nos.	Eigenvalue	Statistic	Critical value	*P* value
In-Sample Data				
None *	0.1733	61.08	20.26	0.0000
At most 1	0.0118	3.59	9.16	0.4767
Maximum Eigenvalue				
None *	0.1733	57.49	15.89	0.0000
At most 1	0.0118	3.59	9.16	0.4767
Out-of-Sample Data				
None *	0.1729	60.90	15.49	0.0000
At most 1	0.0138	4.14	3.84	0.0418
Maximum Eigenvalue				
None *	0.1729	56.76	14.26	0.0000
At most 1	0.0138	4.14	3.84	0.0418

Note: The test model adopts the vector error correction with first two lags of two endogenous variables.

Table 4
The estimations of regime-switching and no-switching models.

	Model 1				Model 2				Model 3				Model 4	
	State u		State d		State u		State d		State u		State d		No Switch	
	MXTW	STW	MXTW	STW	MXTW	STW	MXTW	STW	MXTW	STW	MXTW	STW	MXTW	STW
ρ_{SF}	.9883		.9999		.9895				.9900		.9862		.9881	
Stderr	.0010		.0020		.0014				.0016		.0029		.0014	
P value	.0000		.0000		.0000				.0000		.0000		.0000	
θ_{0j}	.004	.004	-.044	-.058	.004	.005	-.044	-.034	.006	.006			.005	.701
Stderr	.002	.002	.004	.003	.002	.002	.013	.014	.002	.002			.002	.711
P value	.030	.021	.000	.000	.009	.018	9.1E-04	.008	3.4E-04	3.3E-04			.008	.325
θ_{1j}	-.926	.415	1.191	1.252	-.450	-.573	.088	1.436	-.647	-1.087			.124	-.807
Stderr	.790	.710	.975	1.180	.637	.605	.751	1.255	4.586	11.909			.208	.709
P value	.242	.560	.223	.290	.481	.172	.907	.127	.888	.927			.553	.256
θ_{2j}	.187	.609	-.439	-1.258	.123	.511	.406	-.667	.087	.434			-.136	-.906
Stderr	.207	.238	2.388	2.631	.199	.234	1.018	1.141	.259	.278			.169	.712
P value	.368	.011	.854	.633	.538	.015	.690	.280	.738	.119			.423	.204
θ_{3j}	.855	-.497	.028	.006	.361	.470	-.227	-.509	.559	.999			.005	-1.090
Stderr	.795	.716	1.101	1.035	.630	.599	.688	1.181	4.583	11.918			.002	.699
P value	.283	.488	.980	.995	.567	.217	.741	.333	.903	.933			.010	.120
θ_{4j}	-.206	-.639	.480	1.167	-.173	-.563	-.386	-.019	-.147	-.508			.460	.977
Stderr	.181	.212	2.051	2.427	.170	.206	.907	.960	.240	.259			.238	.698
P value	.256	.003	.815	.631	.308	.003	.671	.492	.542	.051			.054	.163
γ_j	.815	.031	.472	.472	.219	.767	-1.087	-2.119	.585	1.537			-.474	1.309
Stderr	.783	.714	1.669	1.867	.632	.611	.764	.954	4.538	12.002			.198	.703
P value	.299	.965	.777	.801	.729	.105	.156	.014	.897	.898			.018	.064
w_j	3.1E-05	3.6E-05	6.5E-04	2.7E-05	2.8E-04	.191			5.1E-04	1.3E-04	.002	.003	4.9E-04	.323
Stderr	1.5E-05	1.7E-05	1.3E-04	3.8E-05	9.2E-05	.045			1.7E-04	1.4E-04	8.1E-04	1.0E-03	1.0E-04	.080
P value	.037	.036	7.2E-07	.488	.001	1.5E-05			.002	.181	.003	.006	7.8E-07	3.5E-05
α_j	.100	.105	.102	.200	.573	2.8E-04			.038	.006	.337	.418	.372	5.2E-04
Stderr	.042	.040	.131	.049	.083	8.8E-05			.021	.010	.180	.196	.088	1.0E-04
P value	.020	.009	.435	6.8E-05	2.2E-11	.001			.035	.285	.031	.017	1.5E-05	3.3E-07
β_j	.850	.838	.177	.182	.206	.592			2.2E-16	.773	.024	.066	.366	.387
Stderr	.038	.034	.289	.044	.045	.070			.317	.254	.261	.249	.074	.069
P value	.000	.000	.540	4.0E-05	3.9E-06	8.9E-16			.500	.001	.463	.395	6.2E-07	2.7E-08
p_{jj}	.949		.849		.983		.567		.921		.161			
Stder	.194		.597		.041		.090		.026		.043			
P value	.000		.085		1.8E-15		.754		.000		1.5E-04			
LogL	1701.10				1646.65				1657.67				1611.06	
AIC	-3322.20				-3227.31				-3259.34				-3184.12	
BIC	-3175.68				-3106.43				-3156.78				-3114.52	
π_j	.748		.252		.961		.039		.914		.086			

Notes: 1. Model 1 is switch for trend and volatility; Model 2 is switch for trend only; Model 3 is switch for volatility only; and Model 4 is no-switch. 2. The estimated parameters are the coefficient of correlation ρ_{SF} in Eq. (7); the coefficients of VECM θ_{ij} (i=0 to 4, j=1,2) and the cointegration speed adjustment γ_j (j=1,2) in Eq. (1); and the coefficients of GARCH (w_j, α_j, β_j) (j=S, F) in Equation (6). 3. $\alpha_j + \beta_j$ in both u and d states is restricted to less than 1 to avoid iterative volatility explosion. 4. p_{jj} and π_j (j=u, d) are respectively the transition probability and the unconditional state probability as shown in Equation (13). 5. In Model 2, the GARCH estimates are the same for states u and d, and in Model 3 the VECM estimates are the same for states u and d. The estimates are listed in state u but not in state d. 6. The cointegration term ϕ in Eq. (1) is determined in advance to around 1.000 with P value = 0.000. The OLS regresses the logarithmic returns of STW on the logarithmic returns of MXTW without constant coefficient. The regression coefficient is estimated to 0.879 with P value =0.000.

relatively strong significant factor to include in the model. For the parsimonious reason, the lag structure should consider the first and second lags and hence the number of lags (i.e. k) in Eq. (1) is set to 2. Panel B of Table 3 reports the cointegration test between the log of MXTW and the log of STW prices. Using the two-lag error correction model, both trace and maximum eigenvalue tests indicate that MXTW and STW have cointegration and they share a common stochastic trend. As a result, our models have to apply the VECM in the regime-switching model framework.

Table 4 reports the estimations of the regime-switching model with 2 lags VECM using in-sample data from August 9, 2001 to May 31, 2007. Overall, the correlation between MXTW and STW is high and significant, with a value greater than 0.98. The intercepts of the model (i.e., θ_{0j}) for the trend of returns are significant except in Model 4 STW. The cointegration parameter ϕ in Eq. (1) is very significant but the adjustment speed γ_j is significant only in Model 4. In a comparison between state u and state d, the coefficients of GARCH are larger in state u while the coefficients of ARCH are higher in state d. It implies that volatility persistence occurs more often in state u and volatility innovation occurs more often in state d. Compared with the other coefficients between models, overall the coefficients of trend in state d are greater than those in state u. As a result, more changeful volatility and a larger trend exist in state d than in state u. Generally, the switching of volatility states between high state and low state is evidenced, but the trend switch estimation requires further investigation, since most parameters, except intercepts, are not significant.

According to the parameter settings, the probability of staying in state u (i.e., p_{uu}) or remaining in state d (i.e., p_{dd}), as shown in

Eqs. (3) and (13), is equal to the cumulative probability of standard normal distribution and the unconditional state probabilities π_u and π_d are formulated by the conditional transition probability as shown in Eq. (13). In Model 1, the conditional transition probabilities of p_{uu} and p_{dd} are equal to 0.949 and 0.849 respectively, whereas the unconditional state probabilities π_u and π_d, are equal to 0.748 and 0.252, respectively. Table 4 also presents the results for the transition and state probabilities in Models 2 and 3. Overall, the parameters of conditional transition probability in state u: p_{uu} are larger and more significant than in state d: p_{dd} and so are the unconditional state probabilities π_j. The result reveals that staying in state u: p_{uu} is more likely than remaining in state d: p_{dd}, so that the transition to state u is more probable than the transition to state d. This result explains why state u, with a higher p_{uu}, is assumed to be the usual (u) state and state d, with a lower p_{dd} is considered as the down (d) state.

Regarding Model 4, in which regime switch is not considered, the positive estimates of θ_0 reveal that the stock prices over a long period exhibit an increasing trend, and the large value of α=0.323 in MXTW reveals that volatilities vary with time. The GARCH terms 0.366 and 0.387 are not too high, indicating that the volatility persistence is not very strong.

As discussed in Haas and Paolella (2012), it is not necessary to hold the coefficient of ARCH term plus the coefficient of GARCH term to 1 in both u and d regimes (i.e. $\alpha_j + \beta_j < 1$ for j=S, F) to ensure a stationary process provided the nonstationary regime has a sufficiently small weight. However, in our case, the initial estimation has experienced the issue in that the predicted process in state d is seemingly nonstationary. Besides, the unconditional state probabilities π_d are not all small enough in Table 4. It appears necessary that the restrictions of $\alpha_j + \beta_j < 1$ in both states be added in model estimation to avoid the iterative and unusual volatility explosion during recurrent rolling window.

5.3. The return and volatility predictions

The feature of our autoregression models gives the comparative prediction capability. Once the model parameters are estimated, according to Eq. (1), one can simply use the week's returns (for autoregression) and the log of prices (for cointegration vector) to predict the next week's spot or futures returns, and even prices. Similarly, according to Eqs. (6) and (7), the variances and covariances of spot and futures returns are predicted for next week. Subsequently, using the rolling window technique, 302 one-week predictions are calculated. With the available predictions, the optimal hedge ratios are solved according to Eq. (28) and the hedge performance is evaluated afterward based on Eqs. (31)–(33).

Panel A of Table 5 reports the MXTW and STW return predictions using four models plus the naive method. Note that the predicted prices use a logarithm to evaluate the predicted price errors (difference between the log of predicted prices and the log of actual prices) to avoid the effects of outliers and price scale. Examining from the mean absolute error (MAE) of the predicted errors, overall the OLS and naive method that both use the present price to predcit next week's price seem to have the best prediction and it approaches the mean of actual return. The prediction strategy of the naive method uses present actual return to predict next week's return. No wonder it is much closer to the mean of actual returns than examining the mean of return predicted errors. Model 1, Model 3 and Model 4 follow the naive method to trace the returns, while Model 2 follows behind. All models tend to overestimate the returns, except Model 2, which has a negative mean of errors but is very close to zero, like the naive method. In fact, because the actual returns are very changeful in terms of the wide range between minimums and maximums and large standard errors, it appears not easy to predict the changeful returns and outperform the naive method.

For the volatility prediction evaluation, the realized variance and covariance calculated from Eqs. (29) and (30) are considered suitable as the proxy of next week's actual volatility. Panel B in Table 5 reports the volatility predictions. Examining the MAE of volatility predicted error reveals that Model 2 procures the best volatility prediction and Model 1, Model 3, and Model 4 follow closely, while the naive method is clearly left behind. The naive method, which uses present volatility proxy to predict next week's volatility, is also found much closer to the mean of volatility proxy, and so are Model 3 and Model 2. Examining the mean of volatility prediction errors reveals that all models underestimate the volatility. Although the volatility proxy is very volatile looking at the extensive spread between minimums and maximums and large standard errors, most models still appear to implement the volatility prediction well enough.

Before computing the hedge ratios using the predictions, the degree of risk aversion λ should be set appropriately. Though λ is suitable to be 1 for a mildly risk-averse hedger, there are likely other scenarios of risk aversion that can be discussed later. Fig. 2 exhibits the predictions of indices, volatility, covariance, and hedge ratios for the out-of-sample period. During the subprime period, around the second half of 2008, the volatilities and covariances are observed to be much higher and the hedge ratios fluctuate more compared with those during the out-of-sample period. For all models, most extracted index predictions are higher than the actual index from August 2008 to December 2008 within the subprime period.

5.4. The effect of various risk aversions

The setting of 1 for risk aversion λ seems suitable, as illustrated in Fig. 2. However, what if the different risk aversions are used to cause the changes in hedge ratio? And what is the optimal risk aversion? According to Eq. (28), it is apparent that the hedger's risk aversion in the market should determine the optimal hedge ratio. If the hedgers dislike the risk, the variance of their hedging portfolio should be confined to some smaller value, since they wish to offset their risk with greater return, and vice versa for hedgers who prefer risk. In this regard, the hedger's risk aversion is referred to as the marginal cost of a unit of risk increase, known as the shadow price or the risk premium.

Panel A of Table 6 reports the changes of hedge ratio (HR) in models due to the various risk aversions of hedgers. As shown in the table in Model 3 and λ=1, the mean and standard deviation of HR are 0.98 and 0.060, respectively. Overall, the switching models

Table 5
MXTW and STW prediction statistics.

Panel A. Predicted Returns

		Actual Return (%)				Predicted Return (%)				Predicted Return Error (%)				
		Min	Max	Mean	Stdev	Min	Max	Mean	Stdev	Min	Max	Mean	MAE	Stdev
Model 1	MXTW	-11.96	14.194	2.51E-03	3.289	-1.75	2.568	.158	.483	-11.78	12.547	.208	2.424	3.296
	STW	-12.17	14.757	4.73E-03	3.440	-3.55	4.373	.144	.662	-11.86	12.347	.196	2.508	3.441
Model 2	MXTW	-11.96	14.194	2.51E-03	3.289	-8.22	2.073	-.106	.876	-12.63	13.343	-.058	2.463	3.355
	STW	-12.17	14.757	4.73E-03	3.440	-7.87	3.039	-.140	.951	-12.57	13.235	-.090	2.553	3.494
Model 3	MXTW	-11.96	14.194	2.51E-03	3.289	-1.85	2.539	.342	.497	-12.27	13.304	.392	2.430	3.291
	STW	-12.17	14.757	4.73E-03	3.440	-3.02	4.305	.384	.695	-12.07	13.345	.436	2.501	3.433
Model 4	MXTW	-11.96	14.194	2.51E-03	3.289	-7.25	3.563	.148	.806	-11.51	13.250	.196	2.453	3.313
	STW	-12.17	14.757	4.73E-03	3.440	-6.56	5.585	.152	.918	-11.28	13.405	.202	2.535	3.464
OLS & Naive	MXTW	-11.96	14.194	2.51E-03	3.289	0	0	0	0	-13.27	12.741	.052	2.412	3.305
	STW	-12.17	14.757	4.73E-03	3.440	0	0	0	0	-13.76	12.978	.055	2.499	3.465

Panel B. Predicted Variance

		Proxy Variance (in 1/1000)				Predicted Variance (in 1/1000)				Predicted Variance Error (%)				
		Min	Max	Mean	Stdev	Min	Max	Mean	Stdev	Min	Max	Mean	MAE	Stdev
Model 1	σ_S^2	.028	8.132	1.135	1.304	.331	5.954	.944	.590	-.677	.169	-.019	.066	.108
	σ_F^2	.014	9.723	1.207	1.538	.360	5.949	.995	.614	-.818	.204	-.021	.074	.129
	σ_{SF}	.019	20.067	1.464	2.250	.360	6.071	1.048	.669	-1.776	.209	-.042	.095	.198
Model 2	σ_S^2	.028	8.132	1.135	1.304	.326	5.301	.868	.576	-.680	.153	-.027	.065	.109
	σ_F^2	.014	9.723	1.207	1.538	.390	5.198	.940	.612	-.823	.296	-.027	.074	.132
	σ_{SF}	.019	20.067	1.464	2.250	.358	5.099	.939	.627	-1.751	.223	-.053	.093	.198
Model 3	σ_S^2	.028	8.132	1.135	1.304	.397	5.034	1.053	.675	-.653	.188	-8.24E-03	.066	.107
	σ_F^2	.014	9.723	1.207	1.538	.432	5.830	1.164	.758	-.790	.248	-4.26E-03	.076	.127
	σ_{SF}	.019	20.067	1.464	2.250	.460	6.504	1.198	.850	-1.714	.239	-.027	.093	.192
Model 4	σ_S^2	.028	8.132	1.135	1.304	.275	7.192	1.057	.749	-.623	.276	-7.78E-03	.067	.108
	σ_F^2	.014	9.723	1.207	1.538	.288	7.454	1.110	.823	-.758	.213	-9.67E-03	.074	.126
	σ_{SF}	.019	20.067	1.464	2.250	.313	7.921	1.207	.939	-1.628	.275	-.026	.094	.190
OLS & Naive	σ_S^2	.028	8.132	1.135	1.304	.028	8.132	1.133	1.305	-.509	.690	-2.22E-04	.073	.122
	σ_F^2	.014	9.723	1.207	1.538	.014	9.723	1.204	1.539	-.699	.860	-2.05E-04	.084	.146
	σ_{SF}	.019	20.067	1.464	2.250	.019	20.067	1.462	2.251	-1.239	1.388	-1.99E-04	.109	.219

Panel A Notes: 1. MAE is the mean absolute error. 2. The error of predicted return is measured by (log(Predicted Price)-log(Actual Price)) ×100%. Predicted Price is Predicted Return×(1+previous week Actual Price). 3. Naive or OLS prediction that uses the present price to predcit next week's price so that the Predicted Return is 0 in Table. Panel B Notes: 1. Variance proxy is the squared actual return of the next week. 2. Predicted variance error is (predicted variance−proxy variance) × 100%. 3. Naive and OLS use the current week's realized volatility to predict next week's volatility. 4. σ_S^2, σ_F^2, and σ_{SF} denote the MXTW variance, the STW variance, and their covariance, respectively.

tend to overestimate HR, especially Model 2, while OLS tends to underestimate HR at a steady 0.879 with nearly 0 standard deviation. As λ increases, the mean of HR follows to increase, except in Model 2 and Model 4, but all models' standard deviations of HR follow to decrease, and vice versa. The reason is that the large λ causes the risk to be more restricted. When λ decreases to 0.01, all models have very high standard deviation, having the maximum HR rising to over 3, as in Model 2, and the minimum of HR dropping to negative, as in Model 1 and Model 3. This result implies that it is not appropriate to set λ too small and to restrict risk too much.

In a comparison of volatility related models, such as Model 1 and Model 4, HR is not necessarily positively or negatively related to the risk aversion, all other things being equal. In the light of Eq. (28), if the market is efficient and follows martingale (i.e., next expected futures price is equal to present futures price and no trading cost), indeed HR is not relevant to risk aversion. However, concerning the expected futures prices, the larger futures price prediction could influence HR negatively, and vice versa the lower futures price prediction. That is why HR is found to be proportional to λ in Model 1 and Model 3 with higher futures price prediction

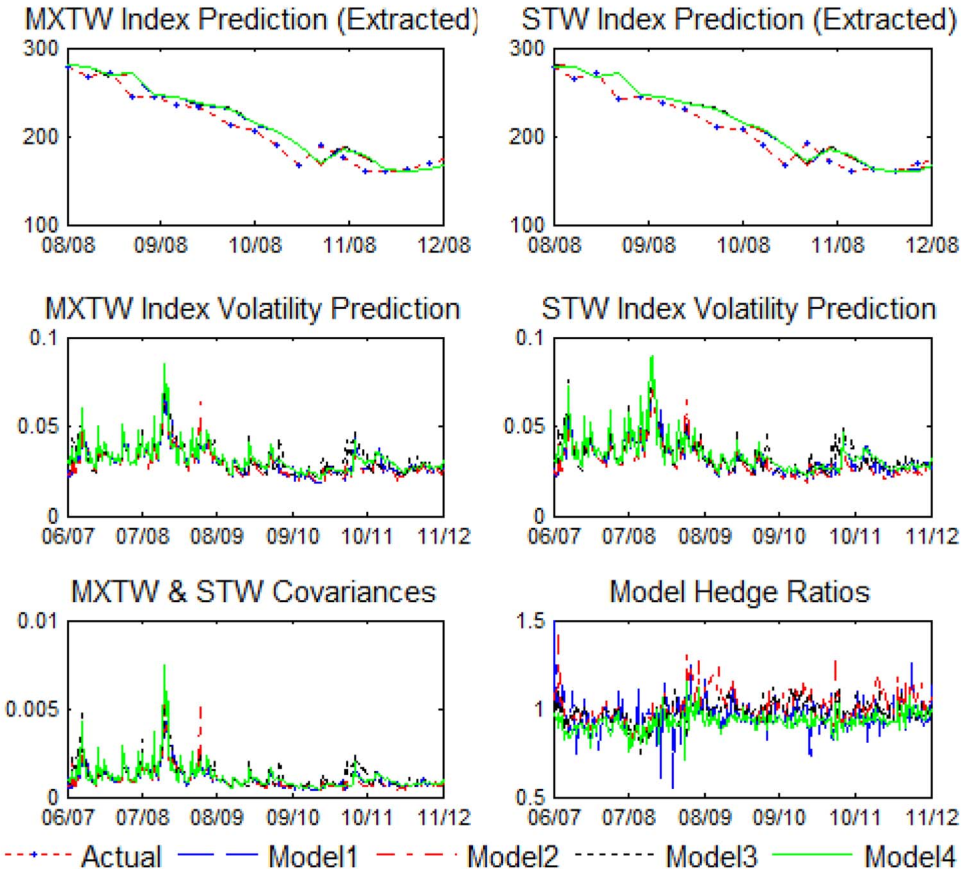

Fig. 2. The predicted indices, volatilities, co-variances, and hedged ratios.

and inversely proportional to λ in Model 2 and Model 4 with lower futures price prediction.

Panel B of Table 6 reports the optimal HRs as the expected price and variance covariance are estimated using realized price and variance covariance in Eqs. (29) and (30) and substituted into Eq. (28) to solve for the optimal HR. Notice that the standard deviations of the optimal HRs are relatively large compared with those of estimated HRs in the models. It implies that the optimal HRs are indeed very changeful and difficult to estimate in most models. Fig. 3 shows the change process of the optimal HRs. It reveals that the estimated HRs in models including Model 1 to Model 4 could not pursue the optimal HRs closely. The optimal HRs do indeed fluctuate more than the estimate HRs. On the other hand, the HRs of OLS are almost constant at 0.879 and the naive method is fixed at 1, as shown in Panel A of Table 6. Clearly, the OLS and the naive method try to estimate HR to approximate the mean of HRs in the long run. Hence, they do not pertain to a predicted model, but to a variance analysis model. However, as λ=1, the mean of the optimal HRs is 0.878, which is very close to the mean of OLS HRs, equal to 0.879. It can be seen from the realized variance of wealth and the realized utility that λ=0.5 to λ =3 is pertinent to evaluate the hedge performance, because if λ=0.1, there are negative estimated HRs and if λ=5, there is negative realized utility. In this stock index context, the middle value of λ being equal to 1—i.e., a mild risk aversion—appears more appropriate than λ=0.5 and λ=3.

Fig. 4 illustrates the change process of the out-of-sample HRs. Clearly, as λ is 0.01, it appears that the HRs have a higher fluctuation compared with the other risk aversions, within the same 0 to 2 scale. The maximum HR is 3.143 in Model 2 and the minimum HR is -0.229 in Model 1, but similar occurrences are infrequently observed from the figure.

5.5. *Different period prediction comparisons*

One issue considered in prediction is data snooping, since the out-of-sample period is divided into four sub-periods and used for several selected models. Thus, the White reality (WR) check (White, 2000) is applied to detect whether one of the models is really superior to the others. The simulated finite distribution is used to conduct the test. The principle of WR assumes that the difference between the loss function (e.g., MAE or RMSE) of the asserted superiors and the benchmark is expected to be zero and it should converge to some finite distribution provided that the dataset is sizable enough. Hence, the null hypothesis follows that the asserted superior model is no different from the benchmark model. If the null is rejected, there exists a superior model outperforming the benchmark. In our case, since the finite distribution of difference between loss functions cannot be known or solved analytically, one has to use the bootstrapping technique (Efron, 1979; Freedman, 1981) to simulate the finite distribution numerically. Once the finite

Table 6
Hedge ratios in various risk aversions.

Panel A. Estimated HR in models

	λ=0.01				λ=0.05				λ=1				λ=3				λ=5			
	Min	Max	Mean	Stdev	Min	Max	Mean	Stdev	Min	Max	Mean	Stdev	Min	Max	Mean	Stdev	Min	Max	Mean	Stdev
Model 1	-0.229	1.848	0.897	0.251	0.549	1.507	0.951	0.096	0.556	1.497	.958	.089	.561	1.490	.962	.087	.562	1.489	.963	.087
Model 2	0.145	3.143	1.070	0.352	0.767	1.477	1.032	0.110	0.808	1.418	1.027	.091	.828	1.382	1.024	.084	.826	1.375	1.023	.083
Model 3	-0.061	1.579	0.848	0.206	0.701	1.158	0.966	0.066	0.717	1.124	.980	.060	.694	1.144	0.990	.060	.687	1.151	.992	.060
Model 4	0.429	2.124	0.928	0.187	0.663	1.220	0.927	0.060	0.650	1.107	.927	.051	.643	1.069	0.927	.048	.641	1.060	.927	.047
OLS	0.879	0.879	0.879	0.000	0.879	0.879	0.879	0.000	0.879	.879	.879	.000	.879	.879	0.879	.000	.879	.879	.879	.000
Naive	1	1	1	0	1	1	1	0	1	1	1	0	1	1	1	0	1	1	1	0

Panel B. Optimal HR in realized context

	λ=0.01				λ=0.05				λ=1				λ=3				λ=5			
	Min	Max	Mean	Stdev	Min	Max	Mean	Stdev	Min	Max	Mean	Stdev	Min	Max	Mean	Stdev	Min	Max	Mean	Stdev
HR*	-4.536	9.850	.649	1.636	-.354	3.267	.853	.391	-.094	2.445	.878	.261	.005	1.900	.895	.200	.014	1.867	.898	.193
RV	159.1	395.1	287.5	45.8	159.1	376.3	280.4	44.4	159.1	376.3	279.5	44.3	159.1	376.3	278.9	44.3	159.1	376.3	278.8	44.3
RU	158.2	377.8	282.7	44.8	154.8	375.1	277.9	44.4	135.6	373.9	275.9	44.5	54.9	370.5	269.4	46.6	-25.5	367.4	263.2	50.6

Note: HR* is the optimal hedge ratio, RV is the realized variance of wealth, and RU is the realized utility.

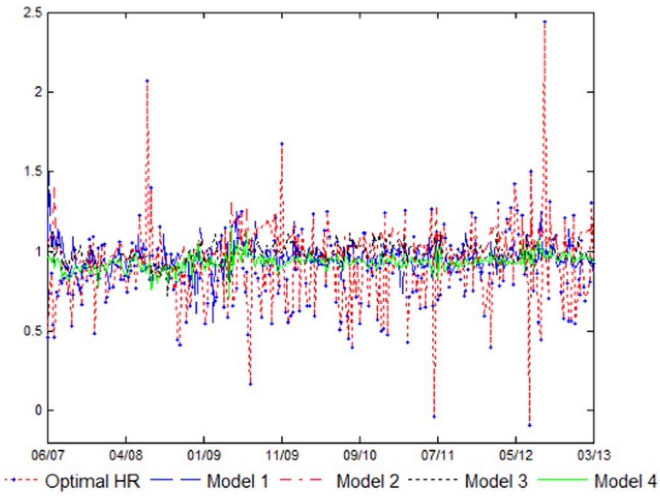

Fig. 3. The tracing of the optimal hedge ratios.

distribution is simulated, it is straightforward to compute the sample statistics of MAE or RMSE of the price prediction errors or the volatility prediction errors between several models. The *P* value of the test is simply equal to the ratio of the number of out limited simulations and the number of simulations. The number of out limited simulations is obtained by counting the number of occurrences when the simulated minimum of MAE or RMSE differences between models is less than the sample average of the minimums of MAE and RMSE differences between models. To date, the bootstrap method is considered as an applicable tool in the WR test.

The log of price differences between the predicted and the actual prices at next week (i.e., the predicted price error) through the out-of-sample period (period 1 to period 3) are the source for conducting the WR price test. The variance or covariance differences between the predicted variance and the proxy variance are used to conduct the WR volatility test. Those models without a volatility estimation, such as OLS and naive method, are evaluated using the previous week's volatility proxy. The outperformance might be due to the chance of data selection using the same data for different selected models. Thus, the WR test is needed to examine if data snooping exists in particular periods.

Table 7 reports the results of the White reality test. This table takes each model listed in a row as the benchmark and tests whether any of the other models could outperform the benchmark model. Examining *P* values of either WR MAE or RMSE tests in

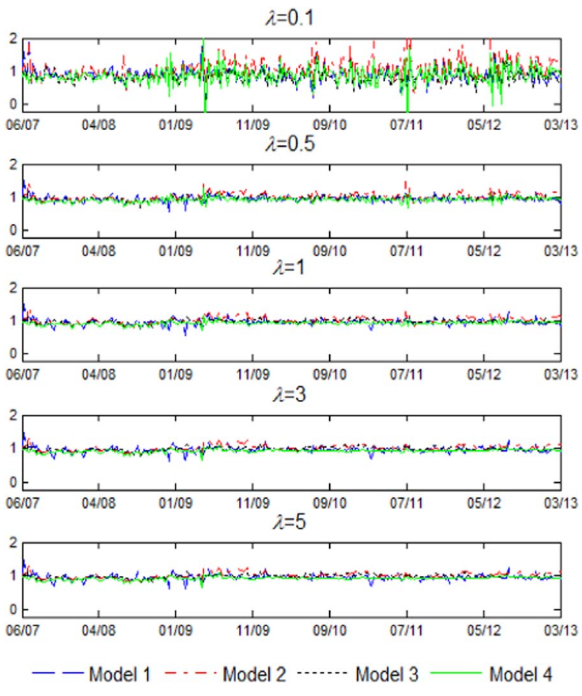

Fig. 4. The change process of hedge ratios in different risk aversions.

Table 7
The White reality prediction test *P* values.

Panel A. WR price test *P* value

	Period All		Period 1		Period 2		Period 3	
	MXTW	STW	MXTW	STW	MXTW	STW	MXTW	STW
	MAE							
Model1	0.215$^{\text{No}}$	0.223$^{\text{No}}$	0.086*	0.092*	0.048**	0.057*	0.945$^{\text{No}}$	0.940$^{\text{No}}$
Model2	0.168$^{\text{No}}$	0.173$^{\text{No}}$	0.933$^{\text{No}}$	0.936$^{\text{No}}$	0.017**	0.015**	0.067*	0.059*
Model3	0.213$^{\text{No}}$	0.243$^{\text{No}}$	0.064*	0.081*	0.040**	0.056*	0.069*	0.071*
Model4	0.177$^{\text{No}}$	0.186$^{\text{No}}$	0.068*	0.082*	0.044**	0.046**	0.058*	0.059*
OLS & Naive	0.810$^{\text{No}}$	0.790$^{\text{No}}$	0.085*	0.064*	0.968$^{\text{No}}$	0.959$^{\text{No}}$	0.076*	0.083*
	RMSE							
Model1	0.211$^{\text{No}}$	0.225$^{\text{No}}$	0.081*	0.081*	0.044**	0.052*	0.943$^{\text{No}}$	0.940$^{\text{No}}$
Model2	0.157$^{\text{No}}$	0.167$^{\text{No}}$	0.921$^{\text{No}}$	0.929$^{\text{No}}$	0.015**	0.017**	0.074*	0.064*
Model3	0.208$^{\text{No}}$	0.237$^{\text{No}}$	0.071*	0.079*	0.040**	0.050**	0.074*	0.077*
Model4	0.179$^{\text{No}}$	0.187$^{\text{No}}$	0.071*	0.081*	0.041**	0.050**	0.055*	0.051*
OLS & Naive	0.815$^{\text{No}}$	0.790$^{\text{No}}$	0.091*	0.069*	0.960$^{\text{No}}$	0.962$^{\text{No}}$	0.077*	0.077*

Panel B. WR volatility test *P* value

	Period All			Period 1			Period 2			Period 3		
	σ_S^2	σ_F^2	σ_{SF}	σ_S^2	σ_F^2	σ_{SF}	σ_S^2	σ_F^2	σ_{SF}	σ_S^2	σ_F^2	σ_{SF}
	MAE											
Model 1	0.122$^{\text{No}}$	0.150$^{\text{No}}$	0.120$^{\text{No}}$	0.049**	0.050**	0.054*	0.008***	0.004***	0.022**	0.010***	0.003***	0.004***
Model 2	0.922$^{\text{No}}$	0.894$^{\text{No}}$	0.922$^{\text{No}}$	0.974$^{\text{No}}$	0.974$^{\text{No}}$	0.977$^{\text{No}}$	0.023**	0.006***	0.050*	0.091*	0.031**	0.964$^{\text{No}}$
Model 3	0.077*	0.076*	0.089*	0.025**	0.023**	0.028**	0.000***	0.000***	0.002***	0.006***	0.001***	0.003***
Model 4	0.083*	0.091*	0.088*	0.024**	0.026**	0.027**	0.000***	0.000***	0.001***	0.001***	0.001***	0.001***
OLS & Naive	0.005***	0.012**	0.000***	0.000***	0.000***	0.000***	0.992$^{\text{No}}$	0.999$^{\text{No}}$	0.972$^{\text{No}}$	0.962$^{\text{No}}$	1.000$^{\text{No}}$	0.093*
	RMSE											
Model 1	0.158$^{\text{No}}$	0.199$^{\text{No}}$	0.156$^{\text{No}}$	0.052*	0.057*	0.055*	0.040**	0.059*	0.045**	0.016**	0.021**	0.020**
Model 2	0.905$^{\text{No}}$	0.884$^{\text{No}}$	0.913$^{\text{No}}$	0.969$^{\text{No}}$	0.972$^{\text{No}}$	0.970$^{\text{No}}$	0.978$^{\text{No}}$	0.078*	0.978$^{\text{No}}$	1.000$^{\text{No}}$	1.000$^{\text{No}}$	1.000$^{\text{No}}$
Model 3	0.102$^{\text{No}}$	0.103$^{\text{No}}$	0.117$^{\text{No}}$	0.038**	0.029**	0.035**	0.004***	0.000***	0.001***	0.019**	0.017**	0.019**
Model 4	0.086*	0.099*	0.104$^{\text{No}}$	0.028**	0.024**	0.033**	0.001***	0.007***	0.002***	0.012**	0.016**	0.014**
OLS & Naive	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.031**	0.963$^{\text{No}}$	0.006***	0.000***	0.001***	0.000***

Note: Superscript $^{\text{No}}$ means no difference, i.e., no rejection of the null. ***, **, and * denote the rejection of the null at a 1%, 5%, and 10% significance level, respectively.

Table 8
Hedge evaluation of models as λ=1.

	Mean(W_0)	Mean(Var_0)	Mean(U_0)	Mean(HR_1)	Std(HR_1)	IWI	AVD	IUI
	Period All							
Model 1	278.50	81.27	237.87	0.958	0.089	0.099%	-82.46%	78.20%
Model 2	278.50	81.27	237.87	1.027	0.091	0.076%	-80.93%	76.79%
Model 3	278.50	81.27	237.87	0.980	0.060	0.076%	-82.51%	78.32%
Model 4	278.50	81.27	237.87	0.931	0.052	0.083%	-84.23%	78.78%
OLS	278.50	81.27	237.87	0.879	0.000	0.076%	-85.17%	78.77%
Naive	278.50	81.27	237.87	1.000	0.000	0.087%	-82.34%	73.06%
	Period 1							
Model 1	287.62	152.67	211.28	0.945	0.118	0.793%	-79.75%	225.07%
Model 2	287.62	152.67	211.28	0.960	0.080	0.740%	-80.57%	220.42%
Model 3	287.62	152.67	211.28	0.949	0.066	0.755%	-81.86%	224.83%
Model 4	287.62	152.67	211.28	0.899	0.054	0.731%	-83.31%	226.57%
OLS	287.62	152.67	211.28	0.879	0.000	0.701%	-84.58%	226.60%
Naive	287.62	152.67	211.28	1.000	0.000	0.798%	-79.87%	206.34%
	Period 2							
Model 1	287.49	43.61	265.68	0.961	0.066	-0.161%	-82.13%	7.84%
Model 2	287.49	43.61	265.68	1.038	0.062	-0.182%	-79.43%	7.71%
Model 3	287.49	43.61	265.68	1.017	0.049	-0.192%	-80.18%	7.71%
Model 4	287.49	43.61	265.68	0.934	0.027	-0.169%	-83.46%	7.89%
OLS	287.49	43.61	265.68	0.879	0.000	-0.165%	-84.65%	7.92%
Naive	287.49	43.61	265.68	1.000	0.000	-0.187%	-81.49%	7.76%
	Period 3							
Model 1	271.11	39.59	251.32	0.956	0.069	-0.009%	-85.02%	7.81%
Model 2	271.11	39.59	251.32	1.065	0.073	-0.025%	-82.57%	7.70%
Model 3	271.11	39.59	251.32	0.973	0.044	-0.035%	-84.92%	7.76%
Model 4	271.11	39.59	251.32	0.955	0.038	-0.029%	-85.47%	7.81%
OLS	271.11	39.59	251.32	0.879	0.000	-0.020%	-86.05%	7.77%
Naive	271.11	39.59	251.32	1.000	0.000	-0.023%	-84.79%	7.77%

Notes: 1. W_0, Var_0, and U_0 represent no hedge wealth, no hedge variance of wealth, and no hedge utility, respectively. HR_1, W_1, Var_1, and U_1 represent hedge ratio, hedged wealth, hedged variance of wealth, and hedged utility, respectively. 2. IWI, AVD, and IUI are defined in Eqs. (31), (32), and (33), respectively.

panel A of Table 7, the result of the price WR test indicates that over all periods, all models accept the null hypothesis—that is, the benchmark model—(i.e., each model in a row) cannot be outperformed under large enough scenarios or simulations. However, if the data are divided into three small periods, it appears that Model 2 in period 1, the naive method in period 2, and Model 1 in period 3 cannot be outperformed for price prediction in mostly scenarios. For the WR volatility test, panel B in Table 7 shows that over all periods, only Model 1 and Model 2 cannot be outperformed based on the values of WR MAE and RMSE tests. If the data are divided into three small periods, clearly period 1 favors Model 2 for volatility prediction, while generally period 2 and period 3 prefer Model 2 or the naive method.

5.6. Hedge evaluation

As indicated in Eq. (28), in our case, the hedge ratio is related to the expected price, expected variance and covariance, and risk aversion, and the hedge ratio also influences the hedge performance, as indicated in Eq. (20). Eq. (31)–(33) are used to evaluate the hedge performance but the expected price, variance, and covariance should be deemed as being realized and hence the realized variance and covariance in Eqs. (29) and (30) are adopted to replace the expected variance and covariance.

Table 8 reports the hedge evaluation based on the 302 one-week ahead predictions for the out-of-sample period. The risk aversion λ is set to 1, as discussed in Section 5.4. In a comparison of the three periods—period 1, the U.S. subprime crisis; period 2, the Greek debt crisis; and period 3, the post risk period—the hedge performance in period 2 is the most favorable time to conduct the hedge, with IWIs surprisingly over 0.701 percent and IUIs up to 206.34 percent in all models, far better than the other two periods. Though AVDs on average have an 80 percent reduction, which is slightly less than that in all period (i.e., the period as a whole), this amount of variance reduction is still actually acceptable. Compared with period 1, periods 2 and 3 are not preferable because of the lower IWI and IUI increments though their AVDs are maintained at a low level. In the entire out-of-sample period, Model 1 performs the best in IWI and OLS and Model 4 performs the best in AVD and IUI. Among three switching models, Model 1 appears to perform more favorably than the naive method.

5.7. Hedge evaluation in various states

It could be useful to evaluate the hedge performance among the models in various hedge states. Considering a univariate series, a hedge state is classified as usual (i.e., state *u*) if the state probability is less than or equal to 0.5; otherwise it can be deemed as down (i.e., state *d*). Thus, for MXTW and STW two series, a total of four hedge states are constructed accordingly and are denoted as follows: *uu*, *ud*, *du*, and *dd*. Under each of the four hedge states, the hedge performance is evaluated and one can explore in which

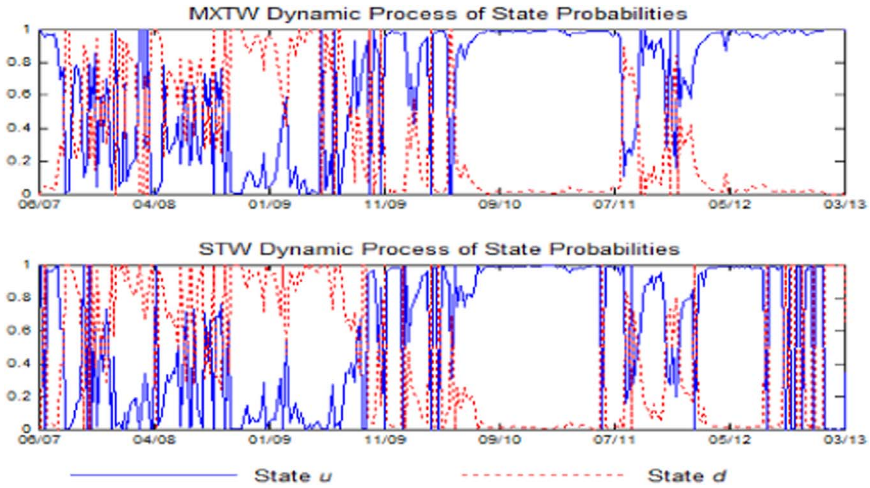

Fig. 5. The dynamic process of state probabilities.

hedge state the hedge performance is more favorable.

The state probabilities for univariate MXTW or STW are estimated in a manner similar to that estimated by the bivariate series, but the residual ε in Eq. (1) should pertain to either the MXTW or STW equation. Model 3 is applied to classify the hedge state, because in Model 1 the part of GARCH is more significant than the part of return trend. Then the rolling window technique is used to

Table 9
Hedge evaluation in different hedge states.

State	Nos.	Mean(W_0)	Mean(Var_0)	Mean(U_0)	Mean(HR_1)	Std(HR_1)	IWI	AVD	IUI
		Model 1							
all	302	278.50	81.27	237.87	0.958	0.089	0.099%	-82.46%	78.20%
uu	169	286.49	61.75	255.61	0.962	0.078	-0.063%	-82.01%	34.83%
ud	42	291.23	59.28	261.59	0.963	0.086	-0.188%	-82.33%	10.89%
du	10	293.98	116.78	235.59	0.926	0.067	-0.704%	-92.87%	200.62%
dd	81	253.35	129.04	188.83	0.950	0.110	0.687%	-82.16%	188.49%
		Model 2							
all	302	278.50	81.27	237.87	1.027	0.091	0.076%	-80.93%	76.79%
uu	169	286.49	61.75	255.61	1.035	0.085	-0.073%	-80.05%	34.85%
ud	42	291.23	59.28	261.59	1.073	0.084	-0.232%	-79.14%	10.60%
du	10	293.98	116.78	235.59	0.972	0.066	-0.833%	-92.39%	199.66%
dd	81	253.35	129.04	188.83	0.995	0.096	0.658%	-82.30%	183.46%
		Model 3							
all	302	278.50	81.27	237.87	0.980	0.060	0.076%	-82.51%	78.32%
uu	169	286.49	61.75	255.61	0.993	0.061	-0.077%	-81.14%	34.90%
ud	42	291.23	59.28	261.59	0.971	0.046	-0.213%	-82.19%	10.85%
du	10	293.98	116.78	235.59	0.974	0.026	-0.775%	-93.42%	204.65%
dd	81	253.35	129.04	188.83	0.960	0.060	0.649%	-84.17%	188.29%
		Model 4							
all	302	278.50	81.27	237.87	0.931	0.052	0.083%	-84.23%	78.78%
uu	169	286.49	61.75	255.61	0.935	0.045	-0.073%	-83.66%	35.10%
ud	42	291.23	59.28	261.59	0.939	0.036	-0.204%	-83.55%	10.92%
du	10	293.98	116.78	235.59	0.901	0.029	-0.702%	-93.66%	200.32%
dd	81	253.35	129.04	188.83	0.920	0.070	0.655%	-84.60%	190.12%
		OLS							
all	302	278.50	81.27	237.87	0.879	0.000	0.076%	-85.17%	78.77%
uu	169	286.49	61.75	255.61	0.879	0.000	-0.064%	-84.51%	35.14%
ud	42	291.23	59.28	261.59	0.879	0.000	-0.214%	-84.46%	10.85%
du	10	293.98	116.78	235.59	0.879	0.000	-0.716%	-93.26%	198.51%
dd	81	253.35	129.04	188.83	0.879	0.000	0.618%	-85.91%	190.26%
		Naive							
all	302	278.50	81.27	237.87	1.000	0.000	0.087%	-82.34%	73.06%
uu	169	286.49	61.75	255.61	1.000	0.000	-0.073%	-82.00%	34.09%
ud	42	291.23	59.28	261.59	1.000	0.000	-0.244%	-81.58%	10.63%
du	10	293.98	116.78	235.59	1.000	0.000	-0.815%	-93.12%	206.60%
dd	81	253.35	129.04	188.83	1.000	0.000	0.703%	-82.10%	170.24%

Note: "all" represents the all out-of-sample data in all periods covering all different hedge states.

Table 10
The number of hedge states in different periods.

		uu		*ud*		*du*		*dd*	
Period	Subtotal Nos.	Nos.	Percentage	Nos.	Percentage	Nos.	Percentage	Nos.	Percentage
Subprime	87	23	26.44%	11	12.64%	3	3.45%	50	57.47%
Greek Debt	64	58	90.63%	2	3.13%	2	3.13%	2	3.13%
Post Risk	79	55	69.62%	20	25.32%	1	1.27%	3	3.80%
Other	72	33	45.83%	9	12.50%	4	5.56%	26	36.11%
Total	302	169	55.96%	42	13.91%	10	3.31%	81	26.82%

compute the total 302 state probabilities and identify the total 302 hedge states for the out-of-sample period.

Fig. 5 illustrates the dynamic processes of the state probabilities for both MXTW and STW. Clearly, the state probabilities switch frequently for both series from June 2007 to November 2009, covering the subprime period, and the state probabilities for both series are stable from July 2010 to June 2011.

Table 9 reports the hedge performance evaluated in the four different hedge states. Surprisingly, for all six models, the hedges in *dd* state present the best performance, having IWIs over 0.618 percent and IUIs over 170.24 percent, compared with the average values having an IWI around 0.082 percent and an IUI around 78 percent for the entire out-of-sample period. Though AVDs in *dd* state are not so unique, they still catch up with the level of average variance reduction in the entire out-of-sample period, which is around 83 percent variance reduction.

The *uu* states are the most frequently observed (169/302=56%) during the out-of-sample period. However, all the models performed below the average in the *uu* state, having IWIs, AVDs, and IUIs around −0.071 percent,−82 percent, and 35 percent, respectively. Noticeably, all models in the *du* state produce the greatest IUIs of 198 percent and AVD of−92 percent, but the IWI is negative at around−0.76 percent. All the models in the *ud* state perform worst regarding IWI and IUI, calculated below−0.188 percent and less than 11 percent, respectively. So, hedgers must avoid the *ud* states to prevent the undesirable IWI and IUI. Hence, the critical hedge opportunities are in *dd* states to achieve a larger IWI or in the *du* state to achieve a desirable IUI and AVD.

Table 10 reports the number of hedge states in a certain hedge period with a total of 302 sample observations. The number of *uu* states has occupied around 55.96 percent of the total observations. The largest frequencies of the *uu* state arise in the Greek debt crisis and the post risk period, having 90.63 percent and 69.62 percent observations respectively in the sub-period. The number of *dd* states occupies around 26.82 percent of the total observations and the most intensive *dd* state occurs in the subprime period, having around 57.47 percent *dd* state observations in this period. The *ud* states occur less frequently and occupy only 13.91 percent of total observations, while the *du* states are observed rarely, at a 3.31 percent rate over all periods.

Most of the hedge states occur primarily in the *uu* state, as hedgers encounter the negative IWI and poor IUI. The Greek debt period comprises 90.63 percent *uu* states, which might be suggested as the less favorable hedging period if hedgers disapprove of negative IWI. The second most observed hedge states are the *dd* states, as hedgers procure a superior performance. The subprime period consists of up to 57.47 percent *dd* states, which are also discernible in Fig. 2, and thus it is regarded as the most demanded hedging period.

5.8. Model discussion

In fact, based on the price or volatility prediction result, most models perform very closely. The WR price test indicates that all models cannot be outperformed over entire periods. For the divided periods, it suggests that period 1 should select Model 2, period 2 should select the naive method, and period 3 should select Model 1 as the superior model. The WR volatility test indicates that only Model 2 and Model 1 cannot be outperformed unambiguously. Model 2 is the most favorable model in period 1. Overall, Model 2 appears to have the best prediction based on the WR test.

Regarding the hedge evaluation, Model 1 has the better IWI, OLS has the better AVD, and Model 4 has the better IUI. Note that the result of hedge evaluation does not correspond to the result of the prediction. The reason for this controversy is that the hedge ratio in our case is a function of expected prices, expected variance and covariance, and risk aversion λ, which is not sufficiently linear to determine the change of HRs.

6. Conclusions

The logarithmic returns of MXTW and STW reveal high peakedness and correlation, and the Jarque-Bera test results indicate the nonlinear property of the series. The high correlation points out that the hedge is effective. Cointegration is found to be significant between MXTW and STW. Based on the White reality test, the regime-switching models incorporating cointegration produce the most favorable price prediction in the trend-switching model, which also produces the most favorable volatility prediction. However, the better prediction does not guarantee the better hedge performance due to the nonlinear structure of the hedge ratio constructed by the predicted return, the predicted variance and covariance, and the risk aversion.

The stock markets are evidenced as having a volatility feedback effect between returns and variance. The risk aversion of hedgers in the market reflects the tradeoff—i.e., the risk premium—between them. If hedgers are more risk averse, the tradeoff becomes large and so does the risk premium. The hedging models aim to find the optimal hedge ratios to either minimize the hedged portfolio risk or maximize the hedged portfolio wealth, but they cannot be achieved simultaneously. Thus, the optimal hedge ratios occur when the risk aversion is confined to the marginal cost of one unit of risk increase in market. In our stock index context, the suitable setting of risk aversion is found to be 1, compared with the settings of 0.1, 0.5, 3, and 5. Further substituting the realized variances and covariances into the utility function to solve the optimal hedge ratios, the mean of optimal hedge ratios is found to be 0.878, which is very close to the OLS estimation, almost constant at 0.879, and so it is little wonder that the OLS performs the best in variance reduction. However, it cannot have the best wealth increase and variance decrease at the same time, because the restriction in risk should be relaxed to acquire higher returns. Therefore, Model 1 performs the best in wealth increase, while OLS and Model 4 perform the best in variance decrease and utility increase, respectively.

The subprime period is found to be the most effective time to conduct the futures hedge. Among all models in the subprime period, the average wealth increase is improved from 0.083 percent to 0.753 percent, an increase of more than nine times, and the average utility increase is up from 77.32 percent to 221.64 percent, an increase of more than three times. Meanwhile, the average variance decrease is attained around 81.66 percent, which is slightly less than the average of 82.94 percent in models, but remains desirable. Finally, the hedge context is divided into four types of hedge state using the volatility-switching model. It is valuable to find that the hedge performance is substantially enhanced when hedging in the *dd* state as both MXTW and STW are in down state. In fact, the subprime period is made up of 57.47 percent *dd* states. In sum, hedgers in this stock index context should maintain hedge ratios around 0.878 and grab the *dd* states or large risk periods to ensure greater hedge performance.

Acknowledgements

The author wants to express great appreciation to the Editor Dr. Chen and both anonymous referees for their constructive comments and valuable suggestions to improve the article substantially.

References

Alizadeh, A. H., & Nomikos, N. K. (2004). A Markov regime-switching approach for hedging stock indices. *Journal of Futures Markets*, *24*(7), 649–674.
Alizadeh, A. H., Nomikos, N. K., & Pouliasis, P. K. (2008). A Markov regime-switching approach for hedging energy commodities. *Journal of Banking and Finance*, *32*(9), 1970–1983.
Andersen, T., Bollerslev, T., Diebold, F., & Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, *71*, 579–625.
Ang, A., & Bekaert, G. (2002). International asset allocation with regime shifts. *Review of Financial Studies*, *11*, 1137–1187.
Bae, K. H., Karolyi, G. A., & Stulz, R. M. (2003). A new approach to measuring financial contagion. *The Review of Financial Studies*, *16*, 717–763.
Baillie, R. T., & Myers, R. J. (1991). Bivariate GARCH estimation of the optimal commodity futures hedge. *Journal of Applied Econometrics*, *6*, 109–124.
Bergman, U. M., & Hansson, J. (2004). Real exchange rates and switching regimes. *Journal of International Money and Finance*, *24*(1), 121–138.
Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rate: A multivariate generalized ARCH approach. *Review of Economics and Statistic*, *72*, 498–505.
Bollerslev, T., Engle, R. F., & Wooldridge, J. M. (1988). A capital asset pricing model with time-varying covariances. *Journal of Political Economy*, *96*(1), 116–131.
Brooks, C., Henry, O. T., & Persand, G. (2002). The effect of asymmetries on optimal hedge ratios. *Journal of Business*, *75*(2), 333–352.
Cai, J. (1994). A Markov model of switching-regime ARCH. *Journal of Business and Economic Statistics*, *12*, 309–316.
Carter, C. (1999). Commodity futures markets: A survey. *The Australian Journal of Agricultural and Resource Economics*, *43*, 209–247.
Dickey, D., & Fuller, W. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, *49*, 1057–1072.
Ederington, L. H. (1979). The hedging performance of the new futures markets. *Journal of Finance*, *34*(1), 157–170.
Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, *7*(1), 1–26.
Engel, C. (1994). Can the Markov switching model forecast exchange rates? *Journal of International Economics*, 36, 151–165.
Frechette, D. L. (2000). The demand for hedging and the value of hedging opportunities. *American Journal of Agricultural Economics*, *82*(4), 897–907.
Freedman, D. A. (1981). Bootstrapping regression models. *The Annals of Statistics*, *9*, 1218–1228.
French, K. R., Schwert, G. W., & Stambaugh, R. F. (1987). Expected stock returns and volatility. *Journal of Financial Economics*, *19*, 3–30.
Gagnon, L., Lypny, G. J., & McCurdy, T. H. (1998). Hedging foreign currency portfolios. *Journal of Empirical Finance*, *5*, 197–220.
Gray, S. (1996). Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics*, *42*, 27–62.
Haas, M. & Mittnik, S. (2008). *Multivariate regime-switching GARCH with an application to international stock markets*. CFS Working Paper Series 2008/08, Center for Financial Studies.
Haas, M., Mittnik, S., & Paolella, M. S. (2004). A new approach to Markov switching GARCH models. *Journal of Financial Econometrics*, *2*(4), 493–530.
Haas, M., & Paolella, M. S. (2012). Mixture and regime-switching GARCH models. , in: Bauwens, L., Hafner, C., & Laurent, S. (Eds.). (2012). Handbook of volatility models and their applications (Article 6). United Kingdom: Wiley, 71–102.
Haigh, M. S., & Holt, M. T. (2000). Hedging multiple price uncertainty in international grain trade. *American Journal of Agricultural Economics*, *82*(4), 881–896.
Haigh, M. S., & Holt, M. T. (2002). Hedging foreign currency, freight and commodity futures portfolios- A note. *The Journal of Futures Markets*, *22*(12), 1205–1221.
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. (1990). Analysis of time series subject to changes in regime. *Journal of Econometrics*, *45*, 39–70.
Hamilton, J. D. (1991). A quasi-Bayesian approach to estimating parameters for mixtures of normal distributions. *Journal of Business and Economic Statistics*, *9*, 27–39.
Hamilton, J. D., & Susmel, R. (1994). Autoregressive conditional heteroskedasticity and changes in regime. *Journal of Econometrics*, 64, 307–333.
Hansson, B., & Hordahl, P. (1998). Testing the conditional CAPM using multivariate GARCH. *Applied Financial Economics*, *8*, 377–388.
Haqmmoudeh, S., Yuan, Y., McAleer, M., & Thomson, M. A. (2010). Precious metals exchange rate volatility transmissions and hedging strategies. *International Review of Economics and Finance*, *19*, 633–647.
Harris, R. D. F., & Shen, J. (2003). Robust estimation of the optimal hedging ratio. *Journal of Futures Markets*, *23*(8), 799–816.
Hayashi, T., & Yoshida, N. (2005). On covariance estimation of non-synchronously observed diffusion processes. *Bernoulli*, *11*, 359–379.
Heifner, R. G. (1972). Optimal hedging levels and hedging effectiveness in cattle feeding. *Agricultural Economic Research*, *24*, 25–35.
Honda, T. (2003). Optimal portfolio choice for unobservable and regime-switching mean returns. *Journal of Economic Dynamics and Control*, *28*, 45–78.
Ichiue, H., & Koyama, K. (2011). Regime switches in exchange rate volatility and uncovered interest parity. *Journal of International Money and Finance*, *30*(7), 1436–1450.
Jin, H. J., & Koo, W. W. (2006). Offshore hedging strategy of Japan-based wheat traders under multiple sources of risk and hedging costs. *Journal of International Money and Finance*, *25*(2), 220–236.

Johnson, L. (1960). The theory of hedging and speculation in commodity futures. *Review of Economic Studies*, *27*, 139–151.
Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*, 263–291.
Klaassen, F. (2002). Improving GARCH volatility forecasts with regime-switching GARCH. *Empirical Economics*, *27*(2), 363–394.
Koutmos, G., & Pericli, A. (1998). Dynamic hedging of commercial paper with T-bill futures. *Journal of Futures Markets*, *18*(8), 925–938.
Kroner, K. F., & Sultan, J. (1993). Time varying distributions and dynamic hedging with foreign currency futures. *Journal of Financial and Quantitative Analysis*, *28*, 535–551.
Lamoureux, C. G., & Lastrapes, W. D. (1990). Heteroskedasticity in stock return data: Volume versus GARCH effects. *Journal of Finance*, 45, *221–229*.
Lee, H. T., & Yoder, J. K. (2007). A bivariate Markov regime-switching GARCH approach to estimate time varying minimum variance hedge ratios. *Applied Economics*, *39*, 1253–1265.
Lee, H. T., Yoder, J. K., Mittelhammer, R. C., & McCluskey, J. J. (2006). A random coefficient autoregressive Markov regime-switching model for dynamic futures hedging. *The Journal of Futures Markets*, *26*, 103–129.
Lence, S. H. (1995). The economic value of minimum-variance hedges. *American Journal of Agricultural Economics*, *77*(2), 353–364.
Lence, S. H. (1996). Relaxing the assumptions of minimum variance hedging. *Journal of Agricultural and Resource Economics*, *21*(1), 39–55.
Lien, D. (1996). The effect of the cointegration relationship on futures hedging: A note. *Journal of Futures Markets*, *16*, 773–780.
Lien, D., & Tse, Y. K. (2001). Hedging downside risk futures vs. options. *International Review of Economics and Finance*, *10*(2), 159–169.
Lien, D., & Yang, L. (2008). Asymmetric effect of basis on dynamic futures hedging: Empirical evidence from commodity markets. *Journal of Banking and Finance*, *32*(2), 187–198.
Liu, K. E., Geaun, J., & Lei, L. F. (2001). Optimal hedging decisions for Taiwanese corn traders on the way of liberalization. *Agricultural Economics*, *25*, 303–309.
Marcucci, J. (2005). Forecasting stock market volatility with regime-switching GARCH models. *Studies in Nonlinear Dynamics Econometrics*, 9(4) (Article 6).
Markowitz, H. M. (1952). Portfolio selection. *Journal of Finance*, *7*(1), 77–91.
Mattos, F., Garcia, P., & Pennings, J. M. E. (2008). Probability weighting and loss aversion in futures hedging. *Journal of Financial Markets*, *11*(4), 433–452.
McLachlan G. & Peel D. (2000). *Finite mixture models*. Wiley Inter-Science, New York.
Moschini, G., & Myers, R. (2002). Testing for constant hedge ratios in commodity markets: A multivariate GARCH approach. *Journal of Empirical Finance*, *9*, 589–603.
Moskowitz, T., Ooi, Y. H., & Pedersen, L. H. (2012). Time series momentum. *Journal of Financial Economics*, *104*(2), 228–250.
Ng, L. (1991). Tests of the CAPM with time-varying covariances: A multivariate GARCH approach. *The Journal of Finance*, *46*, 1507–1521.
Pelletier, D. (2006). Regime-switching for dynamic correlations. *Journal of Econometrics*, 131(1–2), 445–473.
Poomimars, P., Cadle, J., & Theobald, M. (2003). Futures hedging using dynamic models of the variance/covariance structure. *Journal of Futures Markets*, 23(3), 241–260.
Ramchand, L., & Susmel, R. (1998). Cross correlations across major international markets. *Journal of Empirical Finance*, 5, 397–416.
Sarno, L., & Valente, G. (2005). Modelling and forecasting stock returns: Exploiting the futures market, regime shifts and international spillover. *Journal of Applied Econometrics*, 20(3), 345–376.
Stein, J. (1961). The simultaneous determination of spot and futures prices. *American Economic Review*, 51, 1012–1025.
Tse, Y. K. (2000). A test for constant correlations in a multivariate GARCH model. *Journal of Econometrics*, 98(1), 107–127.
Tse, Y. K., & Tsui, K. C. (2002). A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. *Journal of Business and Economic Statistics*, 20, 351–362.
White, H. (2000). A reality check for data snooping. *Econometrica*, 68, 1097–1126.
Working, H. (1962). New concepts concerning futures markets and prices. *American Economic Review*, 52(3), 431–459.
Yun, W. C., & Kim, H. J. (2010). Hedging strategy for crude oil trading and the factors influencing hedging effectiveness. *Energy Policy*, 38(5), 2404–2408.