



Contents lists available at ScienceDirect

Energy Economics

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

Markov switching GARCH models for Bayesian hedging on energy futures markets

Monica Billio^a, Roberto Casarin^a, Anthony Osuntuyi^{a, b, *}

^a Department of Economics, University Ca' Foscari of Venice, Fondamenta San Giobbe 873, Venice 30121, Italy

^b Department of Mathematics, Obafemi Awolowo University, Ile-Ife 220005, Nigeria

ARTICLE INFO

Article history:

Received 24 February 2015

Received in revised form 29 May 2017

Accepted 3 June 2017

Available online xxxx

JEL classification:

C580

C150

C110

C130

C340

Q470

Keywords:

Energy futures

GARCH

Hedge ratio

Markov-switching

ABSTRACT

Effective hedging strategies on oil spot and future markets are relevant in reducing price volatility for investors, energy traders and companies operating in the oil markets. A new Bayesian multi-chain Markov-switching GARCH model for dynamic hedging in energy futures markets is developed. It builds on the construction of a system of simultaneous equations for the return dynamics on the hedged portfolio and the futures. The implication of our model for portfolios allocation and energy trading are manifold. First, our formulation allows for an easy identification of the different states of the discrete processes as volatility regimes. Secondly, the use of regime-switching models with multiple chains allows for an effective way to reduce risk. Furthermore, correlated chains are more flexible than single chain models since they account for co-movement in the volatility regimes of different markets thus they should be preferred in turbulent periods (e.g. financial crisis). Finally, the combination of the expected utility framework with the regime-switching structure allows us to define a robust minimum variance hedging strategy. Changes in the optimal hedging strategies before and after the financial crisis are evidenced when the proposed robust hedging strategy is applied to crude oil spot and futures markets. It is therefore recommended that many and different models should be used in place of a single one in energy risk hedging since they could perform differently in various phases of the market.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

The recent volatile nature of energy commodity (e.g. crude oil, heating oil and gasoline) prices has caused a lot of concern among policy makers, investors, traders and companies operating in energy markets because of the important role it plays in modern economy and the potential losses that can be generated by price fluctuations. In line with this, a number of recent studies on how to hedge energy price risk have emerged in the literature. Hedging is an investment position taken to mitigate the adverse effect arising from changes in the price of a companion investment. A crucial issue is the determination of the optimal hedge ratio, i.e. the number of derivative contracts to buy (or sell) for each unit of the underlying asset on which the investor bears risk (see [Chen et al. \(2003\)](#) for a review). In this paper we focus on crude oil market and contribute to recent studies (see, for example, [Alizadeh et al. \(2008\)](#); [Chang et al. \(2010\)](#);

[Lee \(2010\)](#); [Wang and Wu \(2012\)](#); [Pan et al. \(2014\)](#)) dealing with hedging crude oil risk. We show how designing effective hedging strategies on oil spot and futures markets can help in mitigating against the adverse effect of volatility and volatility transmission across commodity markets. Since volatility regimes and correlation changes have been documented in other commodity markets, such as copper, gold and silver ([Bhar and Hammoudeh, 2011](#)), corn, and wheat ([Creti et al., 2013](#)), our hedging strategy can also be applied successfully to these markets.

In this paper, we propose a new hedging model based on minimizing the risk of a hedged portfolio. The result of this minimization exercise is the Minimum Variance (MV) hedge ratio defined as the ratio of the covariance between the underlying spot and futures returns to the variance of the futures return (see [Johnson \(1960\)](#)). To apply this optimum hedge ratio in practice, [Ederington \(1979\)](#) suggests regressing the underlying spot returns against the futures returns, and using the estimate of the slope as an MV hedge ratio. This approach has been widely criticized on the grounds that some of the well known stylized facts about asset returns are ignored. Accordingly, to improve the hedging performance, time-varying hedge ratios have been proposed in the literature and its estimation have been developed along

* Corresponding author.

E-mail addresses: billio@unive.it (M. Billio), r.casarin@unive.it (R. Casarin), ayokunlet@yahoo.com (A. Osuntuyi).

two major lines. The first approach involves the estimation of the conditional second order moments of the underlying spot and futures returns. The Generalized Autoregressive Conditional Heteroscedasticity (GARCH) models have been proposed for this by Haigh and Holt (2002), Chang et al. (2010), Cifarelli and Paladino (2015), Abul-Bashera and Sadorsky (2016) and Lai and Lien (2017), among others. The later approach treats the hedge ratio as a time-varying regression coefficient, and focuses on the estimation of such a parameter (e.g. see Alizadeh and Nomikos (2004), Lee et al. (2006), and Chang et al. (2010)). Note that this hedging strategy works by re-balancing the hedged portfolio on a period-by-period basis. As this may involve huge transaction costs, it may not be worthwhile to use this particular instrument for hedging. It has also been well documented in the empirical literature that the class of GARCH models exhibits high persistence of the conditional variance, i.e. the process is close to being nearly integrated. In view of this, some authors allow the optimal hedge ratio to be state-dependent. Alizadeh et al. (2008), Lee and Yoder (2007a), Lee and Yoder (2007b) Dark (2015), Philip and Shi (2016) and Su (2017), among others, propose various multivariate Markov-switching (MS) GARCH (MS-GARCH) models. More precisely, due to the path dependence problem of MS-GARCH models, these authors implement the multivariate extension of Gray (1996) model with different characterizations of the time-varying covariance matrix. While Gray's model is attractive, its analytical intractability is a drawback since it cannot be derived using any standard analytical approximation technique.

Our hedging model builds on the Billio et al. (2016) approach to MS-GARCH modeling and inference by extending it to a bivariate GARCH model with multiple and possibly dependent MS processes (multichain MS). More specifically, we assume a system of simultaneous equations modeling both return dynamics on the hedged portfolio and futures in lieu of the traditional spot-futures framework. Each component of this system is characterized by a path dependent MS-GARCH process. This specification firstly makes it possible to take care of the estimation risk of the parameters in the hedging model, which has not gained much attention in literature and secondly allows for a better understanding of the cross dynamics of futures and hedging portfolios and their effect on the optimal hedging decisions. Our modeling framework is close to Alizadeh and Nomikos (2004), but we differ in two ways. The first difference lies in the characterization of the time-varying variance process. While Alizadeh and Nomikos (2004) consider a time-varying variance defined by an exponential function of the lagged 4-week moving average of the difference between the logarithm of the underlying and the logarithm of the futures, we consider a MS-GARCH model. The second difference relates to the properties of the underlying hidden process governing the observable processes. Alizadeh and Nomikos (2004) either assume that the conditional variances of futures returns are regime independent or that the hidden process characterizing the dynamics of the hedged portfolio is independent of the one influencing the futures returns process. We account for these limitations in our econometric framework. Still regarding the MS-GARCH framework,

Sheu and Lee (2014) and Lai et al. (2017) argue that the dependence of both the derivative and the spot on the same hidden state process might be inappropriate. Thus, the authors propose the use of a multichain Markov regime switching GARCH (MSCG) model. In this paper, we also extend the work of Sheu and Lee (2014) by allowing for simultaneous dependence between the Markov chains of the MSCG model.

Another aim of the paper is to develop a robust hedging approach within the MS-GARCH framework. In practice, the parameters in the optimal hedge ratio are unknown, thus optimal hedge ratios are estimated by replacing the unknown parameters by their corresponding estimates. This approach is referred to in the literature as the “plug-in” or Parameter Certainty Equivalent (PCE) principle. Generally speaking, decision makers are left to provide, using an estimation technique of their choice, estimates of the model parameters, and to substitute them directly in the theoretical model. One of the problems with this approach is that it completely ignores estimation risk. Depending on the econometric specification considered for estimating the optimal hedge ratio, large differences are observed in the estimated MV hedge ratios on the same commodity. This observation further suggests that it may be very costly to ignore estimation risk. Moreover, possible relevant non-sample information (such as insider information or subjective prior) could be available to the hedger but discarded in the decision making process.

We thus recast the MV hedging model as an expected utility model and deal with the estimation risk problem within this framework. It may be argued that a rational decision maker would choose an action that maximizes its expected utility over the unknown parameter space. Early studies on this problem have been pursued by Raiffa and Schlaifer (1961) and DeGroot (2005), among others. A review of the application of this theory to portfolio choice, prior to 1978, is provided in Bawa et al. (1979). A more recent application can be found in Kan and Zhou (2007). As appealing as the expected utility theory sounds, it is laden with a number of computational issues. In many empirical analysis, analytical solutions to either the optimization exercise and/or the integration problem are often not achievable. Accordingly, alternative solutions, such as approximation or simulations, are called for. Müller et al. (2004), among others, propose simulation-based approaches to the expected utility optimization problem. In this paper, we propose a robust hedging ratio that accounts not only for parameter uncertainty, but also for different states of the market. We follow a Bayesian decision rule (see, for example, Lence and Hayes (1994a) and Lence and Hayes (1994b)) to account for parameter uncertainty in the definition of optimal hedging strategies.

The structure of the paper is as follows. In the next section, we present the conventional MV hedge ratio as well our revised approach and the Bayesian hedging strategy. In Section 3, we discuss the empirical application of our proposed model to West Texas Intermediate (WTI) crude oil spot and futures prices and compare the result with the conventional OLS method proposed by Ederington (1979). Section 4 concludes the paper.

2. Bayesian optimal hedging

Let $(Y, \mathcal{Y}, P_\theta)$ be a probability observation space, with $\{P_\theta\}_{\theta \in \Theta}$ a parametric family of probability distributions and θ a parameter in the measurable parameter space $(\Theta, \mathcal{F}^\Theta)$. Let $\mathbf{y}_t = (RS_t, RF_t)' \in Y \subset \mathbb{R}^2$, $t = 1, \dots, T$, be an observable process, where RS_t, RF_t , respectively, correspond to returns on the underlying and returns on the derivative (e.g., option, futures) at time t . Let us define the information set available at time t , as the σ -algebra $\mathcal{F}_t = \sigma(\{\mathbf{y}_s\}_{s \leq t})$ generated by \mathbf{y}_t , $t = 1, \dots, T$ and denote with $\mathbf{y}_{s:t} = (\mathbf{y}_s, \dots, \mathbf{y}_t)$ a collection of observable variables.

Considering on the basic paradigm of expected utility theory and following the standard hedging literature on commodities (e.g., see Haigh and Holt (2002) and references therein), the optimal hedge ratio at time t , h_t , is the solution of the following optimization problem

$$\arg \max_{h \in H} E \left(U \left(\mathcal{F}_{t-1}^\Theta \right) \right) = \arg \max_{h \in H} \int_Y U(h, \mathbf{y}_t) p(\mathbf{y}_t | \mathbf{y}_{1:t-1}, \theta) d\mathbf{y}_t, \quad (1)$$

where, $E(\cdot|\mathcal{F})$ is the conditional expectation operator, conditioning on a σ -algebra \mathcal{F} , $\mathcal{F}_t^\theta = \sigma(\mathcal{F}_t \vee \mathcal{F}^\theta)$ the information set generated by the collection of past values of observable process and parameter prior information, $U(\cdot)$ is the utility function, $r(h, \mathbf{y})$ is a function of decision variable, h , and a vector of random variables \mathbf{y} , H is the feasible set of hedge ratios, $p(\mathbf{y}_t|\mathbf{y}_{1:t-1}, \theta)$ is the joint probability density function (pdf) corresponding to \mathbf{y}_t conditional on the past values $\mathbf{y}_{1:t-1}$ and the parameter θ . The minimum variance (MV) hedge ratio proposed by Johnson (1960) fits into this setting by assuming that: (i) the utility function is quadratic, and (ii) the function $r(h, \mathbf{y})$ is the returns on the hedged portfolio ($RS_t - hRF_t$). Under these assumptions, the solution of the problem (1) is

$$h_t = \frac{\text{Cov}(RS_t, RF_t | \mathcal{F}_{t-1}^\theta)}{V(RF_t | \mathcal{F}_{t-1}^\theta)}. \quad (2)$$

An implicit, and important assumption in Eq. (1) is that $p(\mathbf{y}_t|\mathbf{y}_{1:t-1}, \theta)$ is known with certainty. Unfortunately, in practice we are faced with incomplete knowledge of the parameter value, θ (parameter uncertainty).

If the hedger does not know the values of θ , the optimal hedge ratio cannot be evaluated since it is a function of θ . The classic solution to this problem follows the “plug-in” principle (i.e. a point estimate $\hat{\theta} \in \mathcal{F}_{t-1}$ is substituted for the unknown parameter vector θ). Upon appropriate substitution, Eq. (1) becomes

$$\arg \max_{h \in H} E(U|\mathcal{F}_{t-1}) = \arg \max_{h \in H} \int_Y U(r(h, \mathbf{y}_t)) p(\mathbf{y}_t|\mathbf{y}_{1:t-1}, \hat{\theta}) d\mathbf{y}_t. \quad (3)$$

In the technique described above, we act as if the parameters are known, thus suggesting the name parameter certainty equivalent (PCE), as this technique is sometimes referred to in the literature (Lence and Hayes (1994a)). The uncertainty about the parameters in Eq. (1) are completely ignored in this approach and this clearly calls for care when applying this method. Based on this, we adopt the Bayes' decision criterion (see Lence and Hayes (1994a)) by integrating out the unknown parameters in the product of $E(U|\mathcal{F}_{t-1}^\theta)$ and the posterior distribution of θ , i.e.

$$\begin{aligned} \arg \max_{h \in H} E(E(U|\mathcal{F}_{t-1}^\theta) | \mathcal{F}_{t-1}) &= \\ &= \arg \max_{h \in H} \int_{\Theta} \left(\int_Y U(r(h, \mathbf{y}_t)) p(\mathbf{y}_t|\mathbf{y}_{1:t-1}, \theta) d\mathbf{y}_t \right) p(\theta|\mathbf{y}_{1:t-1}) d\theta \\ &= \arg \max_{h \in H} \int_Y U(r(h, \mathbf{y})) p(\mathbf{y}_t|\mathbf{y}_{1:t-1}) d\mathbf{y}_t, \end{aligned} \quad (4)$$

where $p(\mathbf{y}_t|\mathbf{y}_{1:t-1})$ is the marginal posterior predictive distribution. Unlike Eq. (1), Eq. (4) does not involve any unknown parameter, but requires some information about the parameters. The information can come from past values of the observation process or from other prior information included in \mathcal{F}^θ and in the prior distribution of the parameters. In this case, the MV hedge ratio is

$$h_t^{\text{BAY}} = \frac{E(\text{Cov}(RS_t, RF_t | \mathcal{F}_{t-1}^\theta) | \mathcal{F}_{t-1})}{E(V(RF_t | \mathcal{F}_{t-1}^\theta) | \mathcal{F}_{t-1})}. \quad (5)$$

As highlighted in Bawa et al. (1979), applying Bayes' criterion (4) in place of the PCE approach has at least three benefits. First, Bayes' criterion is supported by the basic axioms postulated by von Neumann-Morgenstern, whereas the PCE has no such axiomatic foundation. Second, all relevant (sample or non-sample) information about θ are taken into consideration through the posterior distribution in Bayes' method. In contrast, sample information contained in the point estimates $\hat{\theta}$ are only needed to implement the PCE. Lastly, optimal average risk decision is arrived at by using Bayes' criterion. This framework can be further enriched by accounting for ambiguity. See Guidolin and Rinaldi (2013) for a review.

In many situations, obtaining an analytical solution to the Bayesian optimal hedge ratio problem in Eq. (4) can be a daunting task. This is because, in some cases, neither the maximum nor the integrals in Eq. (4) can be computed analytically, thus demanding alternative approaches such as simulation based methods (see Müller (1999)). For example, the integrand may be too complex to integrate or the number of parameters to integrate over might be too large to evaluate analytically. In such a scenario, it is possible to approximate the optimization problem in Eq. (4) by using draws from the posterior distribution of θ given \mathcal{F}_{t-1} , which is a natural output of the MCMC approximation of the θ posterior distribution (see Amzal et al. (2006) and Müller et al. (2004)). It is worth noting that our approach is general and can be applied to several alternative specifications of the utility function existing in the literature for deriving the optimal hedge ratio (see, for example, Lence and Hayes (1994b) and Haigh and Holt (2002)), other than the quadratic function. In the following, we shall limit our attention to the MV hedge ratio as it is the most commonly used optimal hedging.

2.1. Econometric model specification

A popular econometric model used for calculating the optimal hedge ratio is the linear model proposed by Ederington (1979). In this model a linear relationship is assumed between the underlying spot and futures returns

$$RS_t = \mu + \nu RF_t + \epsilon_t, \quad \epsilon_t \stackrel{iid}{\sim} (0, \sigma^2), \quad (6)$$

where μ, ν and σ are the regression parameters. The ordinary least square (OLS) estimate of the coefficient of RF_t , ν , is then the MV hedging ratio. The assumption of constant variance and covariance in Eq. (6) implies time-invariant hedge ratio and thus makes this approach easy to

implement. However, as highlighted by Myers (1991), this method fails to properly account for all relevant conditioning information available to hedgers when making their decision. Also, this method fails to account for some of the well known stylized facts, such as conditional heteroscedasticity and volatility clustering, commonly observed in financial data. In view of this and to allow for changes in the market conditions to affect the hedge ratios, Eq. (6) is extended to an M state Markov switching model with a time-varying volatility process also characterized by regime switching.

Let us define two measurable spaces (X, \mathcal{X}) and (Z, \mathcal{Z}) and unobserved processes $s_t \in (X, \mathcal{X})$, and $z_t \in (Z, \mathcal{Z})$, $t = 1, \dots, T$, which represent, respectively, the hedging regime of the portfolio and the volatility state of the futures market at time t . Let \mathcal{F}_t^X and \mathcal{F}_t^Z be the sigma algebras generated respectively by s_u , and z_u , $u \leq t$. The following model defines the relationship between the hedged portfolio and the futures market volatility:

$$\begin{aligned} RS_t &= \mu(s_t) + \nu(s_t, z_t)RF_t + \sigma_t\eta_t, & \eta_t &\stackrel{iid}{\sim} \mathcal{N}(0, 1), \\ \sigma_t^2 &= \gamma(s_t) + \alpha(s_t)\epsilon_{t-1}^2 + \beta(s_t)\sigma_{t-1}^2, \\ RF_t &= a(z_t) + \tau_t\zeta_t, & \zeta_t &\stackrel{iid}{\sim} \mathcal{N}(0, 1), \\ \tau_t^2 &= \kappa(z_t) + \omega(z_t)\xi_{t-1}^2 + \psi(z_t)\tau_{t-1}^2, \end{aligned} \quad (7)$$

where, $\epsilon_t = \sigma_t\eta_t$, $\xi_t = \tau_t\zeta_t$, $\mu(s_t), \nu(s_t, z_t), \gamma(s_t) > 0, \alpha(s_t) \geq 0, \beta(s_t) \geq 0, a(z_t), \kappa(z_t) > 0, \omega(z_t) \geq 0, \psi(z_t) \geq 0$. As regards the hedging and volatility states, (s_t, z_t) , we assume that they take values in the set $\{1, \dots, M\}^2$, $t = 1, \dots, T$ and follow a first order Markov chain with transition probabilities

$$\pi_{ij,kl} = p(s_t = i, z_t = j | s_{t-1} = k, z_{t-1} = l), \sum_{i=1}^M \sum_{j=1}^M \pi_{ij,kl} = 1 \quad \forall k, l = 1, 2, \dots, M. \quad (8)$$

The parameter shift functions $\mu(s_t), a(z_t), \gamma(s_t), \alpha(s_t), \beta(s_t), \kappa(z_t), \omega(z_t)$ and $\psi(z_t)$ describe the dependence of parameters on the realized regimes s_t and z_t i.e.

$$\mu(s_t) = \sum_{i=1}^M \mu_i \mathbb{I}_{s_t=i}, a(z_t) = \sum_{j=1}^M a_j \mathbb{I}_{z_t=j}, \quad \text{with} \quad \mathbb{I}_{s_t=i} = \begin{cases} 1, & \text{if } s_t = i \\ 0, & \text{otherwise,} \end{cases}$$

The parameter shift function $\nu(s_t, z_t)$ plays a crucial role in our model, since it allows to separate the contribution of the spot and futures market volatilities to the hedging strategy. We assume:

$$\nu(s_t, z_t) = \sum_{i,j=1}^M \nu_{ij} \mathbb{I}_{s_t=i} \mathbb{I}_{z_t=j}.$$

In order to simplify the exposition, we define $s_{u:t} = (s_u, \dots, s_t)$, $z_{u:t} = (z_u, \dots, z_t)$, $(s, z)_{u:t} = \{(s_r, z_r)\}_{r=u:t}$, $RS_{u:t} = (RS_u, \dots, RS_t)$, $RF_{u:t} = (RF_u, \dots, RF_t)$ whenever $u < t$ $\theta_\pi = (\{\pi_{ij,kl}\}_{i,j,k,l=1,\dots,M})$, $\theta_u^{RS} = (\mu_1, \dots, \mu_M, \nu_{11}, \dots, \nu_{MM})$, $\theta_a^{RF} = (a_1, \dots, a_M)$, $\theta_\sigma = (\gamma_1, \dots, \gamma_M, \alpha_1, \dots, \alpha_M, \beta_1, \dots, \beta_M)$, $\theta_\tau = (\kappa_1, \dots, \kappa_M, \omega_1, \dots, \omega_M, \psi_1, \dots, \psi_M)$ and $\theta = (\theta_\pi, \theta_u^{RS}, \theta_a^{RF}, \theta_\sigma, \theta_\tau)$.

We summarize the theoretical implication of this extension on the optimal hedge ratio in the following proposition.

Proposition 1. Suppose θ is known and assume that the observations are generated by the process described in Eq. (7). Then the conditional minimum variance hedge ratio at time t , is the solution to

$$h_t = \arg \min_{h \in H} V(RS_t - hRF_t | \mathcal{F}_{t-1}^\Theta) \quad (9)$$

which is given by

$$h_t = \underbrace{\frac{\text{Cov}(\mu(s_t), a(z_t) | \mathcal{F}_{t-1}^\Theta)}{V(RF_t | \mathcal{F}_{t-1}^\Theta)}}_{\text{level-shift hedging}} + \underbrace{\sum_{i,j=1}^M \nu_{ij} w_{ij}}_{\text{volatility hedging}}, \quad (10)$$

where

$$w_{ij} = \frac{(\sum_{(s,z)_{1:t-1}} (a_j^2 + \tau_t^2(j) - a_j E[a(z_t) | \mathcal{F}_{t-1}^\Theta]) \pi_{ij, \dots} p((s, z)_{1:t-1} | \mathcal{F}_{t-1}, \theta))}{\sum_{i,j=1}^M (\sum_{(s,z)_{1:t-1}} (a_j^2 + \tau_t^2(j) - a_j E[a(z_t) | \mathcal{F}_{t-1}^\Theta]) \pi_{ij, \dots} p((s, z)_{1:t-1} | \mathcal{F}_{t-1}, \theta))},$$

$$E(a(z_t) | \mathcal{F}_{t-1}^\Theta) = \sum_{(s,z)_{1:t-1}} \sum_{i,j=1}^M a_j \pi_{ij} p((s,z)_{1:t-1} | \mathcal{F}_{t-1}, \theta),$$

$$\text{Cov}(\mu(s_t), a(z_t) | \mathcal{F}_{t-1}^\Theta) = \sum_{(s,z)_{1:t-1}} \sum_{i,j=1}^M (\mu_i a_j - \mu_i E[a(z_t) | \mathcal{F}_{t-1}^\Theta]) \pi_{ij} p((s,z)_{1:t-1} | \mathcal{F}_{t-1}, \theta),$$

$$V(RF_t | \mathcal{F}_{t-1}^\Theta) = \sum_{i,j=1}^M \left(\sum_{(s,z)_{1:t-1}} (a_j^2 + \tau_t^2(j) - a_j E[a(z_t) | \mathcal{F}_{t-1}^\Theta]) \pi_{ij} p((s,z)_{1:t-1} | \mathcal{F}_{t-1}, \theta) \right),$$

$$\pi_{ij} = p(s_t = i, z_t = j | s_{t-1}, z_{t-1}, \theta),$$

and $\tau_t^2(k) = \kappa_k + \omega_k \xi_{t-1}^2 + \phi_k \tau_{t-1}^2$ for $k = 1, \dots, M$ and $t = 1, \dots, T$.

Proof. See Appendix A. \square

Proposition 1 states that the optimal hedge ratio at any point in time can be determined by two components. The first one, which we termed “level-shift hedging”, is given by the conditional covariance between the intercepts ($a(z_t)$ and $\mu(s_t)$) scaled by the conditional variance of RF_t . If $\nu(s_t, z_t) = 0$ and the spot and futures returns go on average in the same direction within the same regime then the hedge ratio increases. The second component, is customarily called “volatility hedging”, is a weighted average of the hedge ratios conditioning on the different states ($\nu_{ij}, i, j = 1, \dots, M$). The weights are driven by the volatility of the returns on the derivative. This suggests that the dynamics of the variance process on the derivative plays an important role in estimating the MV hedge ratio. The role of the derivative's volatility in the hedging strategy is clear when the spot return level is regime independent. See [Remark 2](#).

Remark 2. If $a(z_t)$ is constant, then the optimal hedge ratio in Eq. (10) reduces to the volatility hedging component

$$h_t = \sum_{i,j=1}^M \nu_{ij} \left(\frac{\left(\sum_{(s,z)_{1:t-1}} \tau_t^2(j) p(s_t = i, z_t = j | s_{t-1}, z_{t-1}, \theta) p((s,z)_{1:t-1} | \mathcal{F}_{t-1}, \theta) \right)}{\sum_{i,j=1}^M \left(\sum_{(s,z)_{1:t-1}} \tau_t^2(j) p(s_t = i, z_t = j | s_{t-1}, z_{t-1}, \theta) p((s,z)_{1:t-1} | \mathcal{F}_{t-1}, \theta) \right)} \right). \quad (11)$$

The effect of the dependence between hedging regimes s_t and futures variance states z_t on the hedge ratio is discussed in the following remarks. We consider the two cases of maximal dependence, i.e. $s_t = z_t$ ([Remark 3](#)) and maximal independence, both lagged and simultaneous independence ([Remark 4](#)).

Remark 3. If the dynamics of both the hedged portfolio and the derivative are governed by the same unobserved state process, s_t , then the optimal hedge ratio at time t is given by

$$h_t = \frac{\text{Cov}(\mu(s_t), a(s_t) | \mathcal{F}_{t-1}^\Theta)}{V(RF_t | \mathcal{F}_{t-1}^\Theta)} + \sum_{j=1}^M \nu_j w_j, \quad (12)$$

where

$$w_j = \frac{\left(\sum_{s_{1:t-1}} (a_j^2 + \tau_t^2(j) - a_j E[a(s_t) | \mathcal{F}_{t-1}^\Theta]) p(s_t = j | s_{t-1}, \theta) p(s_{1:t-1} | \mathcal{F}_{t-1}, \theta) \right)}{\sum_{j=1}^M \left(\sum_{s_{1:t-1}} (a_j^2 + \tau_t^2(j) - a_j E[a(s_t) | \mathcal{F}_{t-1}^\Theta]) p(s_t = j | s_{t-1}, \theta) p(s_{1:t-1} | \mathcal{F}_{t-1}, \theta) \right)},$$

$$E(a(s_t) | \mathcal{F}_{t-1}^\Theta) = \sum_{s_{1:t-1}} \sum_{j=1}^M a_j p(s_t = j | s_{t-1}, \theta) p(s_{1:t-1} | \mathcal{F}_{t-1}, \theta),$$

$$\text{Cov}(\mu(s_t), a(s_t) | \mathcal{F}_{t-1}^\Theta) = \sum_{s_{1:t-1}} \sum_{j=1}^M (\mu_j a_j - \mu_j E[a(s_t) | \mathcal{F}_{t-1}^\Theta]) p(s_t = j | s_{t-1}, \theta) p(s_{1:t-1} | \mathcal{F}_{t-1}, \theta),$$

$$V(RF_t | \mathcal{F}_{t-1}^\Theta) = \sum_{j=1}^M \left(\sum_{s_{1:t-1}} (a_j^2 + \tau_t^2(j) - a_j E[a(s_t) | \mathcal{F}_{t-1}^\Theta]) p(s_t = j | s_{t-1}, \theta) p(s_{1:t-1} | \mathcal{F}_{t-1}, \theta) \right),$$

and $\tau_t^2(k) = \kappa_k + \omega_k \xi_{t-1}^2 + \phi_k \tau_{t-1}^2$ for $k = 1, \dots, M$ and $t = 1, \dots, T$.

Remark 4. If s_t and z_t are independent given (s_{t-1}, z_{t-1}) , and $z_{t-1}(s_{t-1})$ does not cause $s_t(z_t)$ one step ahead (see Billio and Di Sanzo (2015)) given $s_{t-1}(z_{t-1})$, then, under the further assumption $\nu(s_t, z_t) = \nu(s_t)$, the result by Alizadeh and Nomikos (2004) is obtained, that is

$$\begin{aligned} h_t &= \sum_{j=1}^M \nu_j \left(\frac{\left(\sum_{s_{1:t-1}} p(s_t = j | s_{t-1}, \theta) p(s_{1:t-1} | \mathcal{F}_{t-1}, \theta) \right)}{\left(\sum_{j=1}^M \left(\sum_{s_{1:t-1}} p(s_t = j | s_{t-1}, \theta) p(s_{1:t-1} | \mathcal{F}_{t-1}, \theta) \right) \right)} \right) \\ &= \sum_{j=1}^M \nu_j p(s_t = j | \mathcal{F}_{t-1}, \theta). \end{aligned} \quad (13)$$

We expect a more flexible and efficient hedging strategy using the above outlined framework over constant hedge ratio since the model allows for shifts in the mean and volatility of both RS_t and RF_t and recognizes the relationship between them. As noted in Section 2, the model parameters in Eq. (10) are not known in practice. In this respect, a natural approach to solving this problem will be to apply the PCE principle. Alternatively, following the Bayesian paradigm outlined above we have the following proposition.

Proposition 5. Assume that the observations are generated by the process described in Eq. (7). Then under certain regularity conditions the Bayesian conditional minimum hedge ratio at time t is the solution to

$$h_t^{BAY} = \arg \min_{h \in H} \left\{ E \left(V \left(RS_t - h RF_t \mid \mathcal{F}_{t-1}^\theta \right) \mid \mathcal{F}_{t-1} \right) \right\} \quad (14)$$

which is given by

$$h_t^{BAY} = \underbrace{\frac{\int_{\Theta} \left[\text{Cov} \left(\mu(s_t), a(z_t) \mid \mathcal{F}_{t-1}^\theta \right) \right] p(\theta | \mathbf{y}_{1:t-1}) d\theta}{\int_{\Theta} \left[V \left(RF_t \mid \mathcal{F}_{t-1}^\theta \right) \right] p(\theta | \mathbf{y}_{1:t-1}) d\theta}}_{\text{integrated level-shift hedging}} + \underbrace{\sum_{ij=1}^M \int_{\Theta} \nu_{ij} w_{ij}(\theta | \mathbf{y}_{1:t-1}) d\theta}_{\text{integrated volatility hedging}}, \quad (15)$$

where,

$$w_{ij}(\theta | \mathbf{y}_{1:t-1}) = \frac{\sum_{(s,z)_{1:t-1}} \left(a_j^2 + \tau_t^2(j) - a_j E \left[a(z_t) \mid \mathcal{F}_{t-1}^\theta \right] \right) \pi_{ij...} p((s,z)_{1:t-1}, \theta | \mathbf{y}_{1:t-1})}{\sum_{ij=1}^M \left(\int_{\Theta} \sum_{(s,z)_{1:t-1}} \left(a_j^2 + \tau_t^2(j) - a_j E \left[a(z_t) \mid \mathcal{F}_{t-1}^\theta \right] \right) \pi_{ij...} p((s,z)_{1:t-1}, \theta | \mathbf{y}_{1:t-1}) d\theta \right)},$$

$$\pi_{ij...} = p(s_t = i, z_t = j | s_{t-1}, z_{t-1}, \theta),$$

$\tau_t^2(k) = \kappa_k + \omega_k \xi_{t-1}^2 + \phi_k \tau_{t-1}^2$ for $k = 1, \dots, M$, $t = \bar{t}, \dots, T$ and \bar{t} is the minimum number of observations needed for the posterior distribution of θ to be proper.

Proof. See Appendix B.

Similar to Proposition 1, Proposition 5 states that the Bayesian optimal hedge ratio at any point in time can be determined by two components. The first component measures the expected covariance between the intercepts divided by the expected variance of the returns on futures after incorporating all available information about the unknown parameters through their joint posterior distribution. Conditional on past observations, the second component is the expected hedge ratio subject to a modified joint posterior distribution of the unknown parameters.

2.2. Computational issues

An important ingredient needed in the computation of the optimal hedge ratio in Eq. (15) is the posterior distribution of the augmented parameter vector $p((s,z)_{1:t-1}, \theta | \mathbf{y}_{1:t-1})$, $t = \bar{t}, \dots, T$. These quantities cannot be identified with any known distribution. This limitation makes the evaluation of Eq. (15) non-trivial. We shall address this problem by using simulations.

The computation of the MV hedge ratio will be broken down into two main stages. The first part consists of approximating the posterior distribution of the unknown parameters vector given past observations, while the second part involves evaluating the hedge ratio.

Following Billio et al. (2016), we propose an efficient simulation technique for Bayesian approximation of the posterior probability, $p((s,z)_{1:t-1}, \theta | \mathbf{y}_{1:t-1})$. The proposed approach is based on MCMC Gibbs algorithm which allows us to circumvent the path dependence problem inherent in MS-GARCH models and efficiently sample the state trajectories. The samples generated by this MCMC algorithm are used in the second stage for approximating the moments in Eq. (15).

We assume fairly informative prior for θ_π and independent uniform prior for θ_u^{RS} , θ_a^{RF} , θ_σ and θ_τ and denote with $f(\theta)$ the joint prior density. To avoid label switching we assume the identifiability restriction: $\gamma_1 < \gamma_2 < \dots < \gamma_M$, $\kappa_1 < \kappa_2 < \dots < \kappa_M$. In order to generate samples from the posterior density of the augmented parameter vectors:

$$f(\theta, (s, z)_{1:t} | RS_{1:t}, RF_{1:t}) \propto f(RS_{1:t} | (s, z)_{1:t}, \theta, RF_{1:t}) f(RF_{1:t} | (s, z)_{1:t}, \theta) p((s, z)_{1:t} | \theta) f(\theta) \quad (16)$$

our Gibbs sampler iterates over the following steps:

1. $p((s, z)_{1:t} | \theta, RS_{1:t}, RF_{1:t})$,
2. $f(\theta_\pi | \theta_u^{RS}, \theta_a^{RF}, \theta_\sigma, \theta_\tau, (s, z)_{1:t}, RS_{1:t}, RF_{1:t}) = f(\theta_\pi | (s, z)_{1:t})$,
3. $f(\theta_u^{RS}, \theta_a^{RF}, \theta_\sigma, \theta_\tau | \theta_\pi, (s, z)_{1:t}, RS_{1:t}, RF_{1:t}) = f(\theta_u^{RS}, \theta_a^{RF}, \theta_\sigma, \theta_\tau | (s, z)_{1:t}, RS_{1:t}, RF_{1:t})$.

The full joint distribution of the state variables, $s_{1:t}$ and $z_{1:t}$, given the parameter values and return series

$$p((s, z)_{1:t} | \theta, RS_{1:t}, RF_{1:t}) \propto f(RS_{1:t} | RF_{1:t}, \theta, (s, z)_{1:t}) f(RF_{1:t} | \theta, (s, z)_{1:t}) \quad (17)$$

is a non-standard distribution. In view of this, we consider a Metropolis Hastings (MH) strategy for generating proposals for the state variables. We construct the proposal distribution by first considering an approximation of the regime switching GARCH model and then derive the joint distribution of the state variables. See Billio et al. (2016) for alternative approximations. Samples of the state trajectory are then drawn by Forward Filter Backward Sampling scheme. Details of the proposal construction and the MH algorithm are given in Appendix C.

In the second stage, G MCMC samples from $p((s, z)_{1:t-1} | \theta, \mathcal{F}_{t-1}^\Theta)$ are used to approximate the hedge ratio:

$$\hat{h}_t^{BAY} = \frac{\sum_{g=1}^G [\text{Cov}(\mu^{(g)}(s_t), a^{(g)}(z_t) | \mathcal{F}_{t-1}^\Theta)]}{\sum_{g=1}^G [\text{Var}^{(g)}(RF_t | \mathcal{F}_{t-1})]} + \frac{1}{G} \sum_{i,j=1}^M \sum_{g=1}^G \nu_{ij}^{(g)} w_{ij}(\theta^{(g)} | \mathbf{y}_{1:t-1})$$

$$w_{ij}(\theta^{(g)} | \mathbf{y}_{1:t-1}) = \frac{\left(\left(\tau_t^{(g)} \right)^2 (j) + \left(a_j^{(g)} \right)^2 - a_j^{(g)} E \left[a^{(g)}(z_t) | \mathcal{F}_{t-1}^\Theta \right] \right) p^{(g)}(s_t = i, z_t = j | s_{t-1}^{(g)}, z_{t-1}^{(g)}, \theta^{(g)})}{\frac{1}{G} \sum_{j,k=1}^M \sum_{g=1}^G \left(\left(\tau_t^{(g)} \right)^2 (k) + \left(a_k^{(g)} \right)^2 - a_k^{(g)} E \left[a^{(g)}(z_t) | \mathcal{F}_{t-1}^\Theta \right] \right) p^{(g)}(s_t = j, z_t = k | s_{t-1}^{(g)}, z_{t-1}^{(g)}, \theta^{(g)})}$$

It is worth noting that the decision problem characterized by Proposition 5 may be classified as a sequential estimation problem. This is because, in contrast to a fixed decision problem, as new observation \mathbf{y}_{t+1} arrives, the hedger updates the posterior distribution, $f(\theta | \mathbf{y}_{1:t})$, about the unknown parameters and by induction revises the hedge ratio. In our computational procedure, at each date t an MCMC algorithm is employed for drawing samples from the posterior probability distribution of the unknown parameters which are then used in computing the moments in the Bayesian hedge ratio. A drawback of our Bayesian estimation approach is the potential computational burden involved running the MCMC algorithm on the posterior probability distribution at each date. However, it can be argued that, the procedure remains feasible in practice since the computation of hedging ratio only requires about an hour on daily basis. Alternative procedures such as the sequential MCMC or sequential Monte Carlo may be used to reduce the computing time when a timely updating of the hedge ratio is required.

3. An application to energy markets

The goal of this section is twofold. First, we aim to apply our model to provide empirical evidence of the effects of the recent financial crisis on the crude oil markets. Second we want to assess the efficiency of the proposed hedging models and to compare them.

3.1. Hedging on the crude oil market

We consider daily closing energy prices for West Texas Intermediate (WTI) crude oil futures for the period September 14, 2001 to July 31, 2013 (2967 observations). Both spot and futures daily settlement prices were obtained from the US Energy information Agency (<http://www.eia.doe.gov>). The daily returns are computed using the first difference of the natural logarithm of the daily settlements. Fig. 1 displays the sample path of the crude oil squared returns on spot and futures. We observe volatility clustering, which calls for the use of MS-GARCH models.

Before proceeding with the evaluation of the hedge ratio, we consider a full sample estimation of the parameters of the MS-GARCH model under different assumptions on the hidden state process (i.e. independent, dependent and same state variable). Fitting a full MS-GARCH model to empirical data can lead to parameter estimates close to the boundary of the parameter space and often results in slow convergence of the

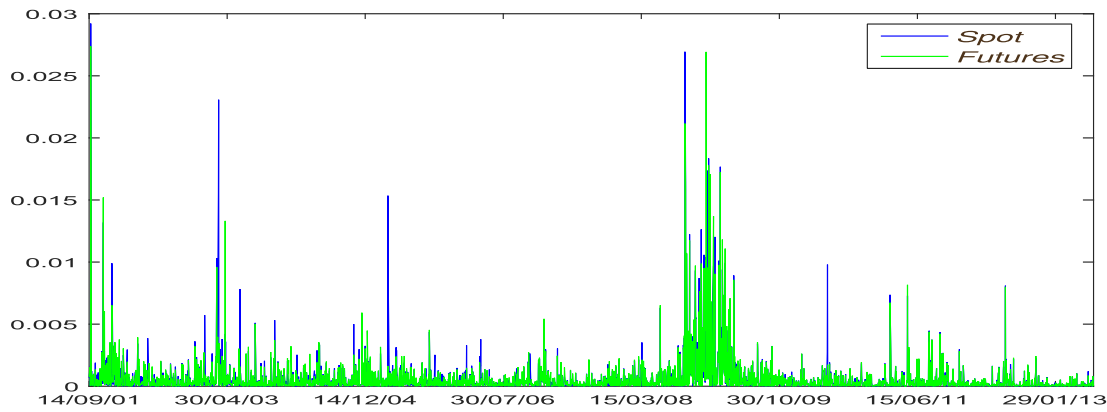


Fig. 1. Graphs for daily squared returns on WTI crude oil spot and futures from 14 September 2001 to 31 July 2013.

MCMC chain. Based on this, a common practice in the estimation of MS-GARCH models with empirical data is to impose some restriction on the parameters. See Billio et al. (2016), Bauwens et al. (2010) and Augustyniak (2014) for illustrations. In this paper, we assume that $a(z_t) = 0$ and $\mu(s_t) = 0$. This analysis will allow us to investigate some volatility features of the spot and futures oil markets. Moreover, this specification of the bivariate spot and futures MS-GARCH model will allow us to draw some conclusions on the optimal hedging strategy to use in these markets.

We perform 10,000 Gibbs iterations after convergence according to the Geweke's diagnostic (Geweke, 1992). To reduce serial correlation of the draws, we consider every 10th draw after convergence of the Gibbs iteration to obtain the results presented below. Tables 1 to 4 show the estimation results for two-state models using the full sample of observation described above. The choice of a two-regime process is motivated by its intuitive interpretation (i.e. low and high volatility state) and the fact that commodity markets dynamics (see Alizadeh et al. (2008)) can be captured in an efficient way by the model. It is however important to note that in our model we specify two regimes for each return (i.e. spot and futures) process and obtain a four-regime model which is firstly an improvement over existing studies and secondly a model flexible enough to capture the non-linearities in the data. From the estimated parameters in Table 1, regime 1 may be labeled as the low volatility state. With the exception of the constrained multichain MS-GARCH model, we observe from Table 1 that the volatility persistence of the hedged portfolio measured by the sum of GARCH parameter, β , and the ARCH parameter, α , is higher in regime 1 than in regime 2. In other words, large

Table 1
Parameter estimate of the MS-GARCH model and standard deviation in parenthesis.

| | MC-f-MSGARCH | MC-c-MSGARCH | SC-MSGARCH |
|---|--------------------|--------------------|---------------------|
| <i>Mean Equation</i> | | | |
| ν_{11} | 0.994(0.0011) | 0.993(0.0014) | 0.991(0.0013) |
| ν_{12} | 0.629(0.0097) | | |
| ν_{21} | 0.947(0.0011) | 0.875(0.0128) | 0.829(0.0189) |
| ν_{22} | 0.055(0.0097) | | |
| <i>Volatility equation for the hedged portfolio</i> | | | |
| γ_1 | 1.23e-06(4.93e-08) | 1.62e-06(7.85e-08) | 1.64e-06(2.15e-07) |
| γ_2 | 8.33e-05(6.06e-06) | 1.14e-04(9.67e-06) | 1.65e-04(1.79e-05) |
| α_1 | 0.560(0.0369) | 0.363(0.0310) | 0.868(0.0501) |
| α_2 | 0.586(0.0554) | 0.632(0.0708) | 0.091(0.0503) |
| β_1 | 0.037(0.0022) | 0.005(0.0032) | 0.022(0.0086) |
| β_2 | 0.292(0.0525) | 0.325(0.0675) | 0.442(0.0873) |
| <i>Volatility equation for future returns</i> | | | |
| κ_1 | 9.76e-06(3.25e-06) | 1.11e-06(8.99e-07) | 7.14e-06(3.315e-06) |
| κ_2 | 9.70e-05(4.37e-05) | 5.48e-05(1.34e-05) | 4.73e-05(1.813e-05) |
| ω_1 | 0.073(0.0104) | 0.026(0.0062) | 0.062(0.0122) |
| ω_2 | 0.093(0.0144) | 0.122(0.0220) | 0.084(0.0226) |
| ψ_1 | 0.908(0.0123) | 0.965(0.0067) | 0.918(0.0176) |
| ψ_2 | 0.794(0.0097) | 0.789(0.0388) | 0.872(0.0370) |

Notes: SC-MSGARCH stands for single chain MS-GARCH; MC-c-MSGARCH stands for constrained multichain MS-GARCH model; and MC-f-MSGARCH stands for unconstrained multichain MS-GARCH.

Table 2
Transition matrix for the unconstrained multichain MS-GARCH model i.e. MC-f-MSGARCH model.

| | $s_{t-1} = 1, z_{t-1} = 1$ | $s_{t-1} = 1, z_{t-1} = 2$ | $s_{t-1} = 2, z_{t-1} = 1$ | $s_{t-1} = 2, z_{t-1} = 2$ |
|--------------------|----------------------------|----------------------------|----------------------------|----------------------------|
| $s_t = 1, z_t = 1$ | 0.9124 | 0.6383 | 0.2672 | 0.2176 |
| $s_t = 1, z_t = 2$ | 0.0026 | 0.0866 | 0.0168 | 0.2635 |
| $s_t = 2, z_t = 1$ | 0.0766 | 0.0534 | 0.6682 | 0.3781 |
| $s_t = 2, z_t = 2$ | 0.0084 | 0.2217 | 0.0478 | 0.1408 |

Note: s_t and z_t , respectively, denote the hedging regime of the portfolio and the volatility state of the futures market at time t .

Table 3

Transition matrices $P(s_t = i | s_{t-1} = j)$ and $P(z_t = i | z_{t-1} = j)$, $i, j = 1, 2$, for the MC-c-MSGARCH model.

| (a) Hedged portfolio | | | (b) Futures | | |
|----------------------|---------------|---------------|-------------|---------------|---------------|
| | $s_{t-1} = 1$ | $s_{t-1} = 2$ | | $z_{t-1} = 1$ | $z_{t-1} = 2$ |
| $s_t = 1$ | 0.894 | 0.383 | $z_t = 1$ | 0.974 | 0.059 |
| $s_t = 2$ | 0.106 | 0.617 | $z_t = 2$ | 0.026 | 0.941 |

Table 4

Transition matrix $P(s_t = i | s_{t-1} = j)$, $i, j = 1, 2$, for the SC-MSGARCH model.

| | $s_{t-1} = 1$ | $s_{t-1} = 2$ |
|-----------|---------------|---------------|
| $s_t = 1$ | 0.930 | 0.424 |
| $s_t = 2$ | 0.070 | 0.576 |

persistence tends to be associated with a low volatility regime. This result may largely be a reflection of the dependence assumption between the chains driving the two series as observed in the unconstrained multichain MS-GARCH model, and the single chain MS-GARCH model case.

Transition probabilities in Table 2 show that the probability that the hedged portfolio and futures return simultaneously remain in the high regime is very low. Whereas, from Table 4 we observe that the single chain MS-GARCH model gives a relatively high probability for the two variables being in the high state simultaneously. The implication of this observation is that, when possible misalignments between the states of the chains driving the two dynamics are not taken into account, our results may be a reflection of an under- or over-estimation of the volatility. Nevertheless, a robust deduction from all the MS-GARCH specifications under consideration is that when both returns are in the low regime at time $t - 1$, it is likely that this scenario will be maintained in the next period. Also, Table 2 suggests that when the returns on the hedged portfolio are in a different state with respect to the returns on the futures at time $t - 1$, then the most probable scenario at time t will be the alignment of the futures to the same scenario of the hedged portfolio.

Given the current volatility states across markets, the unconstrained multichain MS-GARCH framework provides an avenue to studying the influence of a change in the state of the futures on the probability of the hedged portfolio remaining in the same regime, and vice versa, i.e.

$$p(s_t = i | s_{t-1} = h, z_{t-1} = l) \quad \text{and} \quad p(z_t = j | s_{t-1} = h, z_{t-1} = l) \quad \forall \quad i, j, h, l = 1, 2$$

In Table 5, we report these probabilities. The influence of z_{t-1} on the changes in regime for the hedged portfolio is evident; in fact, the probability of a hedged portfolio staying in regime 1, when the futures is in regime 1 in the previous month, is 0.92, but decreases to 0.72 when the futures is in regime 2. In a similar way, the futures remains in regime 1 with a 98% chance when the hedged portfolio is in the same regime, but switches to regime 2 with a probability equal to 31% when the hedged portfolio is in regime 2.

The correlations between the returns on the spot and the futures can be obtained by evaluating

$$\rho_t = \frac{1}{\sqrt{\frac{\sigma_{t,t}^2}{\sigma_{t,t}^2} + 1}}.$$

Differently from the single chain MS-GARCH model with two correlations, the multichain MS-GARCH models have four correlation regimes at each point in time. To get an idea of the relative importance of the correlations in each MS-GARCH specifications, we replace the time varying variance with their respective regime unconditional variances. In the unconstrained (constrained) multichain case, when both the spot and futures are in the high volatility regime, the correlation is equal to 0.997 (0.989), and when both of them are in the low volatility regime, the correlation is equal to 0.016 (0.389). When the spot return is in the high volatility regime and the futures return is in the low volatility regime, the correlation is equal to 0.995 (0.998), and when the spot return is in the low volatility regime and the futures return is in the high volatility regime, the correlation is equal to 0.634 (0.185). In brief, we find that the correlation of spot and futures return series tends to be higher when the spot is in the low volatility regime. The estimated correlations for the first and second regime of the single chain MS-GARCH model are equal, respectively, to 0.979 and 0.822. These values are somewhere between the highest and the lowest correlation estimated from the multichain MS-GARCH model. Overall, more model flexibility may be achieved with the unconstrained MS-GARCH model, since it is the one with the widest correlation range, among the models considered in the analysis.

Table 5

Conditional probabilities, $p(s_t = i | s_{t-1} = h, z_{t-1} = l)$ and $p(z_t = j | s_{t-1} = h, z_{t-1} = l)$, $i, j, h, l = 1, 2$, for the unconstrained multichain MS-GARCH model i.e. MC-f-MSGARCH model.

| | $s_{t-1} = 1, z_{t-1} = 1$ | $s_{t-1} = 1, z_{t-1} = 2$ | $s_{t-1} = 2, z_{t-1} = 1$ | $s_{t-1} = 2, z_{t-1} = 2$ |
|-----------|----------------------------|----------------------------|----------------------------|----------------------------|
| $s_t = 1$ | 0.9150 | 0.7249 | 0.2840 | 0.4811 |
| $s_t = 2$ | 0.0850 | 0.2751 | 0.7160 | 0.5189 |
| $z_t = 1$ | 0.9890 | 0.6917 | 0.9354 | 0.5957 |
| $z_t = 2$ | 0.0110 | 0.3083 | 0.0646 | 0.4043 |

Note: s_t and z_t , respectively, denote the hedging regime of the portfolio and the volatility state of the futures market at time t .

Table 6

Lower bound for the pairwise estimate of the log-Bayes factor $\log(BF(j', j))$ in favor of M_j over $M_{j'}$ for models SC-MSGARCH ($j = 1$), MC-c-MSGARCH ($j = 2$) and MC-f-MSGARCH ($j = 3$).

| Model | SC-MSGARCH ($j = 1$) | MC-c-MSGARCH ($j = 1$) | MS-f-MSGARCH ($j = 3$) |
|---------------------------|---------------------------|-----------------------------|-----------------------------|
| SC-MSGARCH ($j' = 1$) | 1 | 1.65×10^4 | 1.50×10^3 |
| MC-c-MSGARCH ($j' = 2$) | -1.65×10^4 | 1 | -1.50×10^4 |
| MC-f-MSGARCH ($j' = 3$) | -1.50×10^3 | 1.50×10^4 | 1 |

Note: The results in this Table is computed using lower bound of the log-Bayes factor given in Eq. (21) i.e. $\log(\widehat{BF}(j', j)) \geq \frac{1}{C} \sum_{i=1}^C \log(f(y_{1:T}^{(i)}, (s, z)_{1:T}^{(i)}, M_j)) - \frac{1}{C} \sum_{i=1}^C \log(f(y_{1:T}^{(i)}, (s, z)_{1:T}^{(i)}, M_{j'}))$, for $j, j' = 1, 2, 3$. SC-MSGARCH stands for single chain MS-GARCH; MC-c-MSGARCH stands for constrained Multichain MS-GARCH model; and MC-f-MSGARCH stands for unconstrained Multichain MS-GARCH.

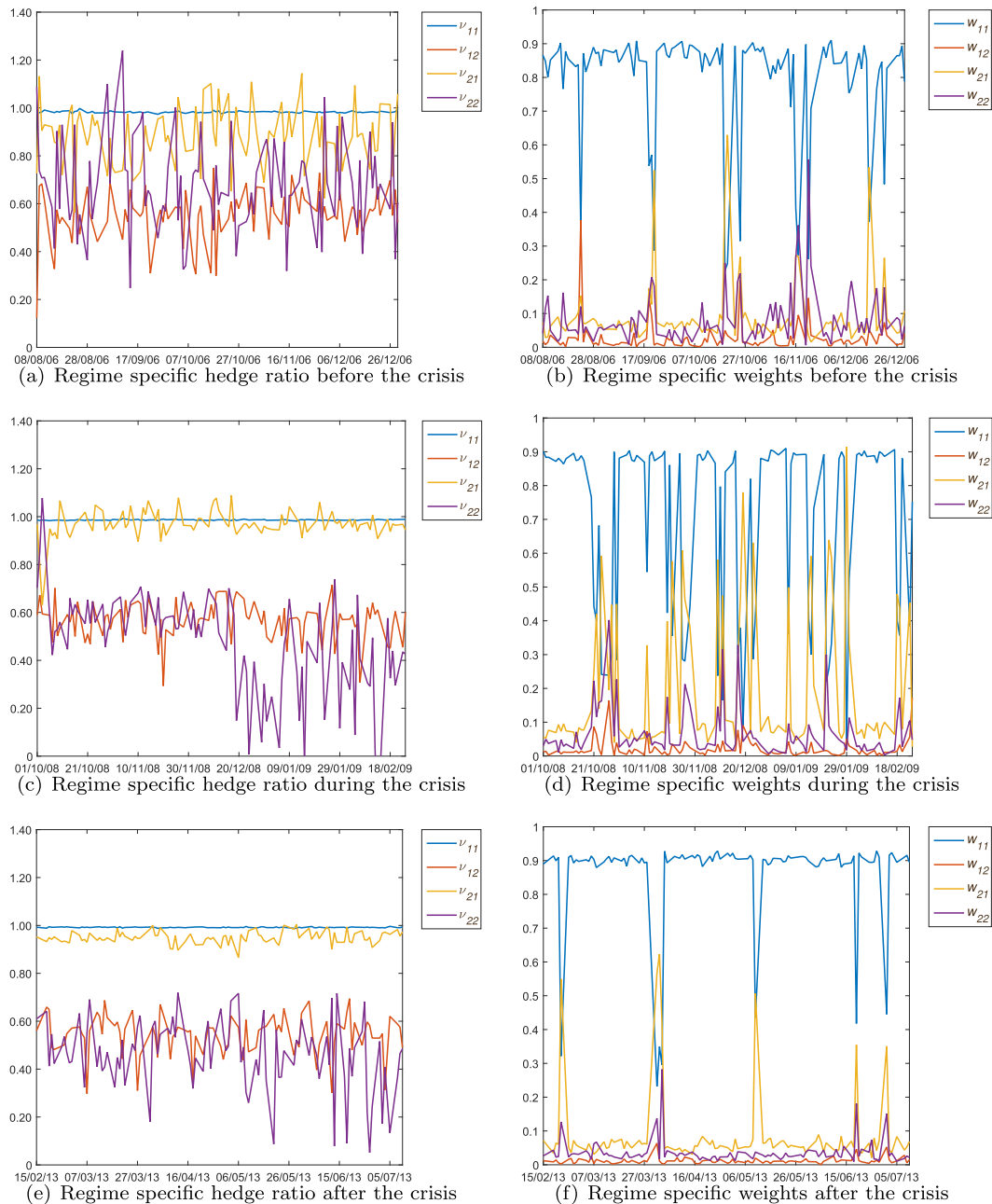


Fig. 2. Regime specific hedge ratio ($\nu_{ij}, i, j = 1, 2$) and corresponding weights ($w_{ij}, i, j = 1, 2$) for the unconstrained multichain MS-GARCH model (MC-f-MSGARCH). First row pre-financial crisis (08/08/2006 to 03/01/2007); second row during financial crisis (01/10/2008 to 25/03/2009); and third row post-financial crisis (15/02/2013 to 31/07/2013).

Lastly, a pairwise comparison of all the competing models, $M_1 = \text{SC-MSGARCH}$, $M_2 = \text{MC-c-MSGARCH}$, $M_3 = \text{MC-f-MSGARCH}$, is carried out by evaluating the log-Bayes factor:

$$\log(BF(j', j)) = \log \left(\frac{f(\mathbf{y}_{1:T}|M_j)}{f(\mathbf{y}_{1:T}|M_{j'})} \right), \quad (18)$$

between model M_j and $M_{j'}$ for all pairs of models based on the full sample. However, it is analytically intractable to evaluate the marginal joint probability density function,

$$f(\mathbf{y}_{1:T}|M_j) = \int_{\theta} \sum_{(s,z)_{1:T}} f(\mathbf{y}_{1:T}, \theta, (s,z)_{1:T}|M_j) d\theta \quad j = 1, 2, 3,$$



Fig. 3. Regime specific conditional variance ($\tau_i, i = 1, 2$) and corresponding predicted probabilities ($p_{ij}, i, j = 1, 2$) for the unconstrained multichain MS-GARCH model (MC-f-MSGARCH). First row pre-financial crisis (08/08/2006 to 03/01/2007); second row during financial crisis (01/10/2008 to 25/03/2009); third row post-financial crisis (15/02/2013 to 31/07/2013).

of the observations given the model. Hence, we turn to numerical procedures. There are several ways of numerically computing the log-Bayes factor in Eq. (18). See Kass and Raftery (1995). In this paper, we apply the procedure due to Newton and Raftery (1994).

$$BF(j', j) = E \left[\frac{1}{f(\mathbf{y}_{1:T} | \theta, (s, z)_{1:T}, M_{j'})} \right] E \left[\frac{1}{f(\mathbf{y}_{1:T} | \theta, (s, z)_{1:T}, M_j)} \right] \quad (19)$$

where the expectation is evaluated over the posterior distribution of $(\theta, (s, z)_{1:T})$. Using the samples obtained from the MCMC procedure described for the posterior distribution, $f(\theta, (s, z)_{1:T} | \mathbf{y}_{1:T}, M_j)$, the pairwise Bayes factor given in Eq. (19) can be estimated by

$$\widehat{BF}(j', j) = \frac{1}{G} \sum_{i=1}^G \frac{1}{f(\mathbf{y}_{1:T} | \theta^{(i)}, (s, z)_{1:T}^{(i)}, M_{j'})} \bigg/ \frac{1}{G} \sum_{i=1}^G \frac{1}{f(\mathbf{y}_{1:T} | \theta^{(i)}, (s, z)_{1:T}^{(i)}, M_j)}, \quad (20)$$

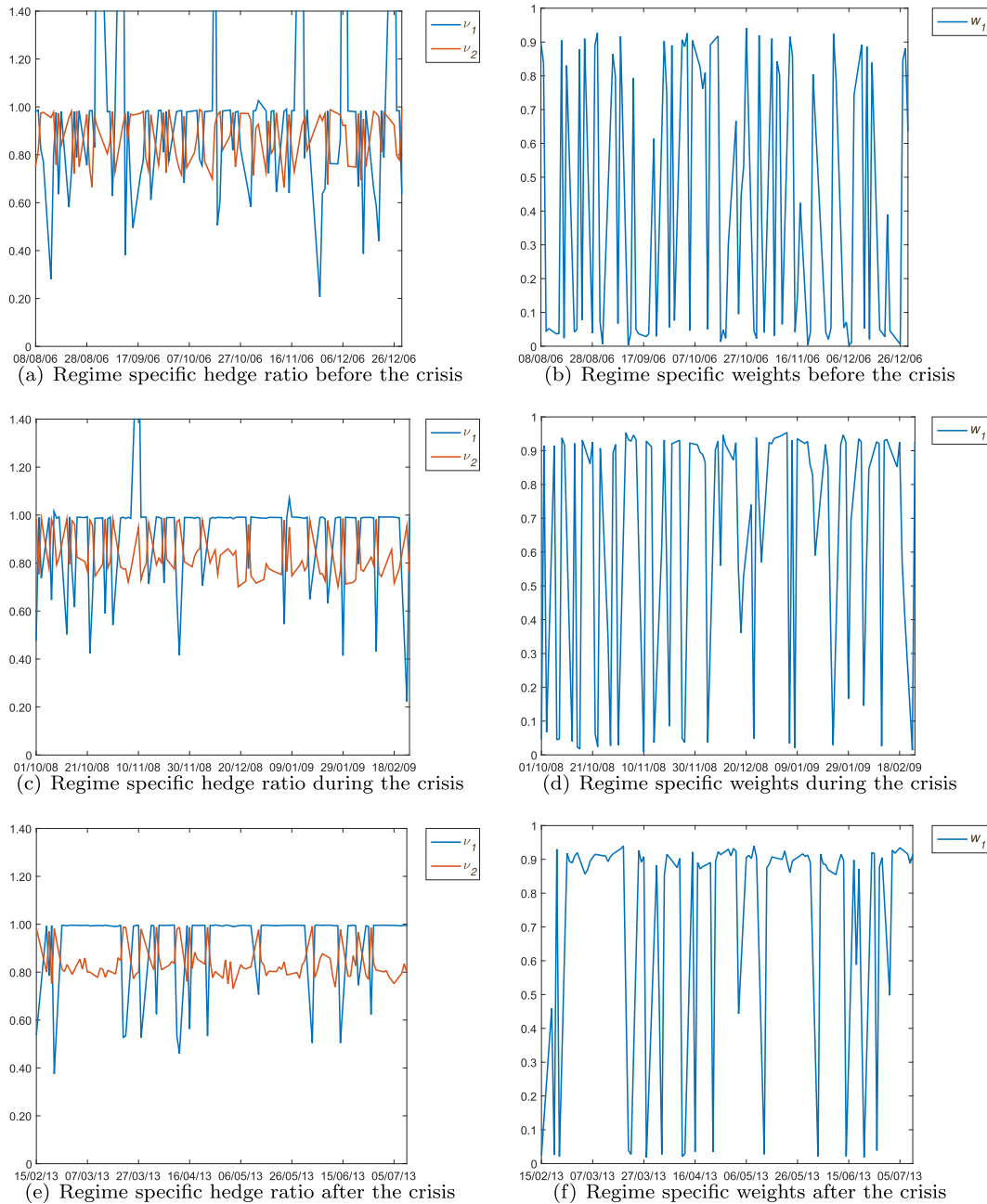


Fig. 4. Regime specific hedge ratio ($\nu_i, i = 1, 2$) and corresponding weights (w_1) for the single chain MS-GARCH model (SC-MSGARCH). First row pre-financial crisis (08/08/2006 to 03/01/2007); second row during financial crisis (01/10/2008 to 25/03/2009); and third row post-financial crisis (15/02/2013 to 31/07/2013).

and an estimate of the log-Bayes factor (BF) in favor of M_j over $M_{j'}$ obtained by

$$\begin{aligned} \log(\widehat{BF}(j', j)) &= \log \left(\frac{1}{G} \sum_{i=1}^G \frac{1}{f(\mathbf{y}_{1:T} | \theta^{(i)}, (s, z)_{1:T}^{(i)}, M_j)} \right) / \frac{1}{G} \sum_{i=1}^G \frac{1}{f(\mathbf{y}_{1:T} | \theta^{(i)}, (s, z)_{1:T}^{(i)}, M_{j'})} \\ &\geq \frac{1}{G} \sum_{i=1}^G \log(f(\mathbf{y}_{1:T} | \theta^{(i)}, (s, z)_{1:T}^{(i)}, M_j)) - \frac{1}{G} \sum_{i=1}^G \log(f(\mathbf{y}_{1:T} | \theta^{(i)}, (s, z)_{1:T}^{(i)}, M_{j'})) \end{aligned} \quad (21)$$

Table 6 reports the pairwise evaluation of the log-Bayes factor for all the models under consideration. We follow the Kass and Raftery (1995) scale for $\log(BF)$. From Table 6 we conclude that there is a very strong evidence in favor of both multichain MSGARCH models against the

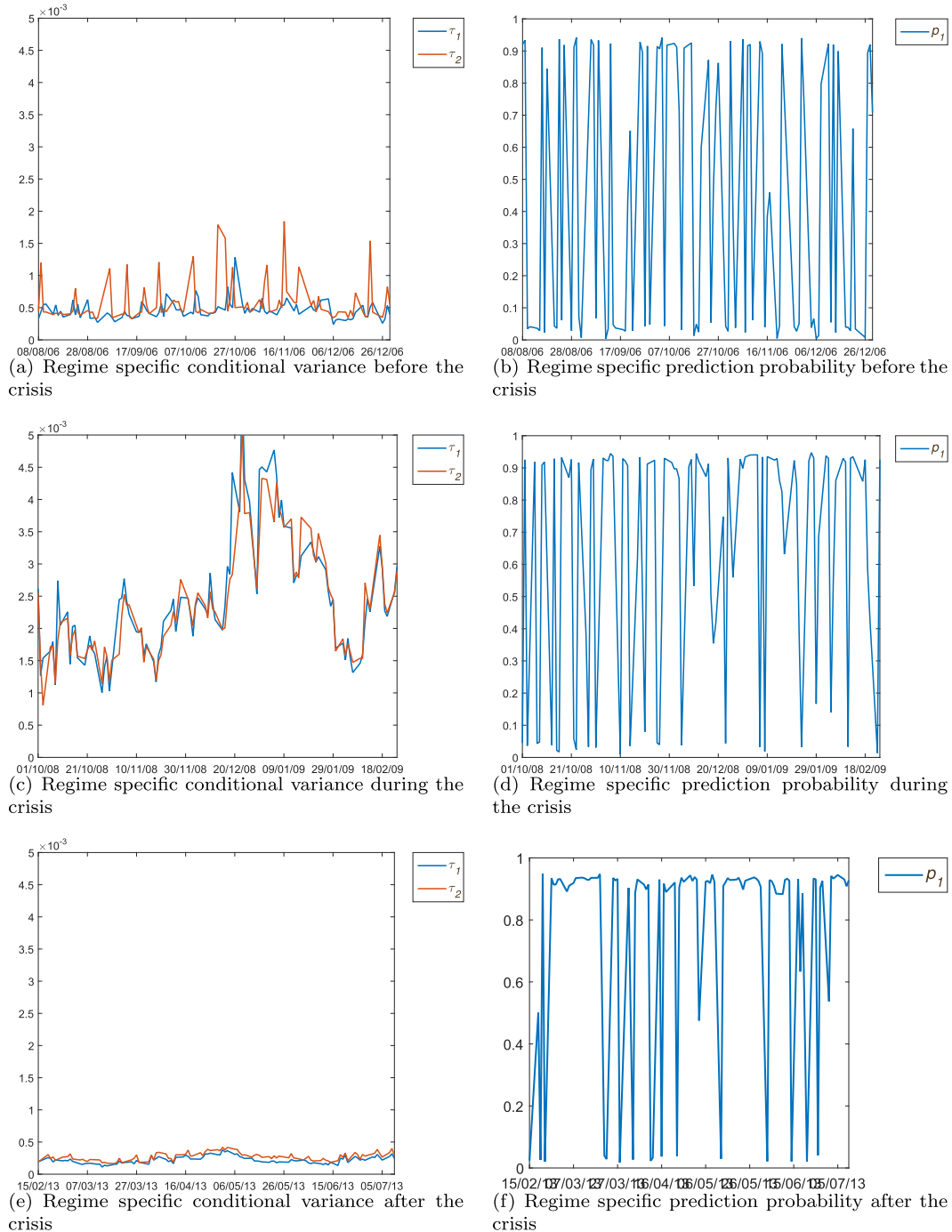


Fig. 5. Regime specific conditional variance ($\tau_i, i = 1, 2$) and the corresponding prediction probabilities (p_1) for the single chain MS-GARCH model (SC-MSGARCH). First row pre-financial crisis (08/08/2006 to 03/01/2007); second row during crisis (01/10/2008 to 25/03/2009); and third row post-financial crisis (15/02/2013 to 31/07/2013).

single-chain MSGARCH model. Moreover, the independent multichain MS-GARCH model is favored to the correlated multichain MSGARCH model.

Based on the previous discussion, it can be deduced that multichain MS-GARCH models have an important role to play in the optimal hedge ratio theory.

3.2. Hedge ratio

To check whether our proposed model is of practical interest, we conduct a sequential estimation exercise to investigate its performance. Moreover, the sequential analysis allows us to provide evidence of time-changes in the volatility transmission mechanisms, and in the correlation between the two markets. We exploit the hedge ratio interpretation of one of the parameters used in our model to study the effects on the optimal hedging strategy due to possible time-variations in the volatility and correlation structures of these markets. For each hedging model, an out-of-sample analysis of its hedging performance with a daily re-balancing is carried out. Accordingly, on a daily basis, an estimate of the MV hedge ratio is obtained, and the futures position to be taken at the end of that day until the following day is also determined. The sample is then extended by one day, the hedge ratios re-estimated, and the hedge rebalanced and held until the end of the next day. For each MS-GARCH specification we consider the sequential estimation of the hedge ratio for three sub-periods i.e. 08/08/2006 to 03/01/2007, 01/10/2008

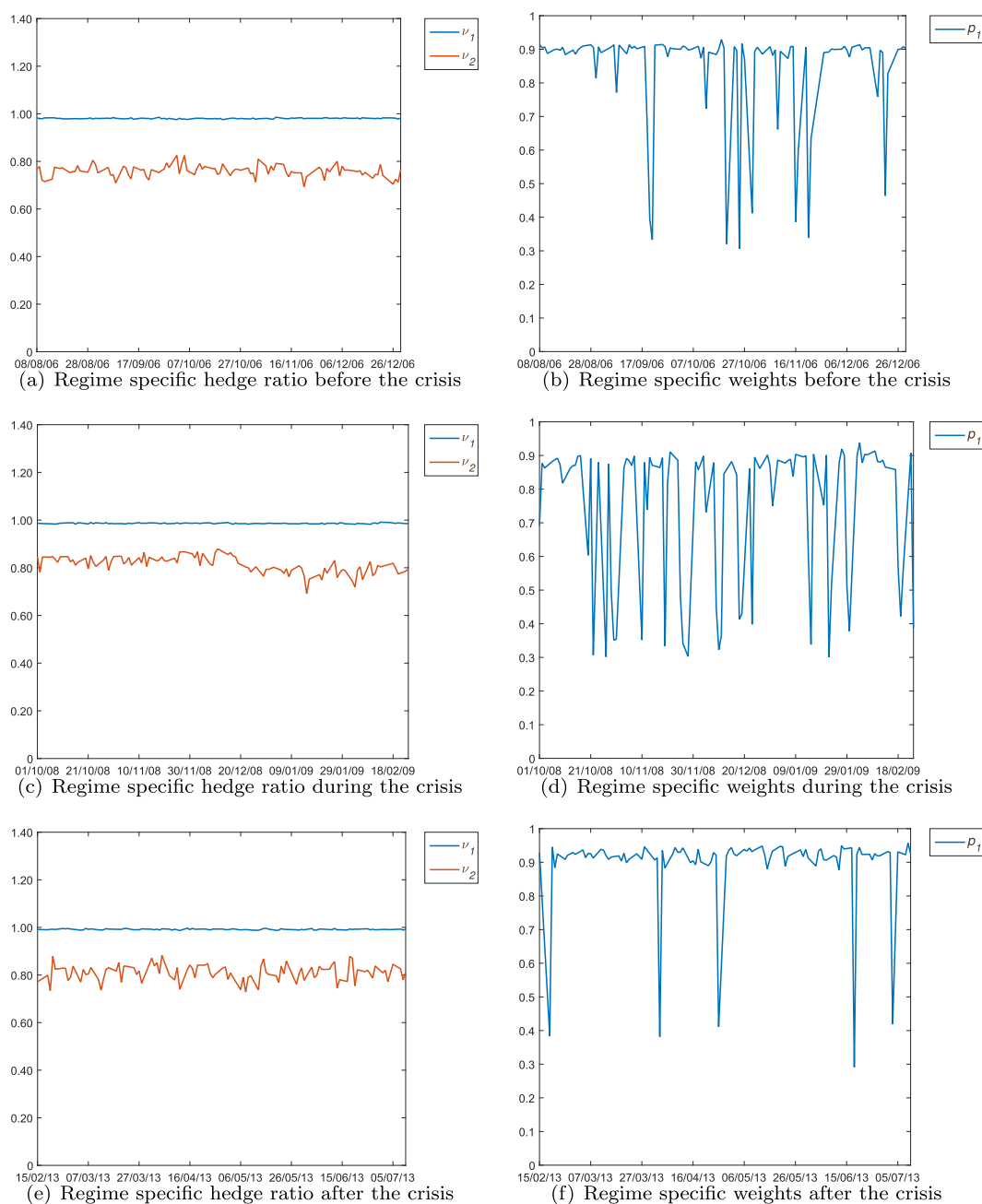


Fig. 6. Regime specific hedge ratio ($\nu_i, i = 1, 2$) and corresponding prediction probabilities (p_1) for the constrained multichain MS-GARCH model (MC-c-MSGARCH). First row pre-financial crisis (08/08/2006 to 03/01/2007); second row during crisis (01/10/2008 to 25/03/2009); and third row post-financial crisis (15/02/2013 to 31/07/2013).

to 25/03/2009 and 15/02/2013 to 31/07/2013. These periods correspond, respectively, to the period before, during and after the 2008/2009 global financial crisis. We fix the 2008–2009 global financial crisis time line following the results of previous papers (e.g. see Phillips (2011)) and considered it as exogenously given in the model comparison.

Figs. 2–6 show how the different regime-specific hedge ratios, conditional variances and prediction probabilities have evolved over the three sub-sample periods. For the unconstrained MS-GARCH model, we observe a small range of values for the regime specific hedge ratios (see Fig. 2 (a)), and high variability in the conditional variance of the futures (see Fig. 3 (a)), prior to the 2008/2009 global financial crisis. However, after the 2008/2009 global financial crisis, we observe a clear separation of the hedge ratios (Fig. 2 (e)) into two groups determining the change in the hidden process on the futures returns. Also, the variability of the conditional variance of the futures is very low (Fig. 2 (e)) relative to our observation prior to the global financial crisis. One possible argument is that investors are more careful and learned from their experience during the financial crisis. The application of the single chain MS-GARCH model tends to suffer from an under- or over-estimation problem arising from the use of a single chain (see Fig. 4 (a)–Fig. 4 (e)). In the case of the constrained MS-GARCH model, there is no significant difference in the evolution of the hedge ratio before and after the global financial crisis. A direct comparison between the hedge ratios formed by the unconstrained and constrained MS-GARCH suggests that the unconstrained model is more flexible as it produces a wider range of values for the hedge ratio. Overall, the unconstrained MS-GARCH model seems to perform best among the set of models under comparison. Also, the high probability of staying in the low volatility regime implies low transaction costs because the investor only needs to re-balance his/her portfolio occasionally.

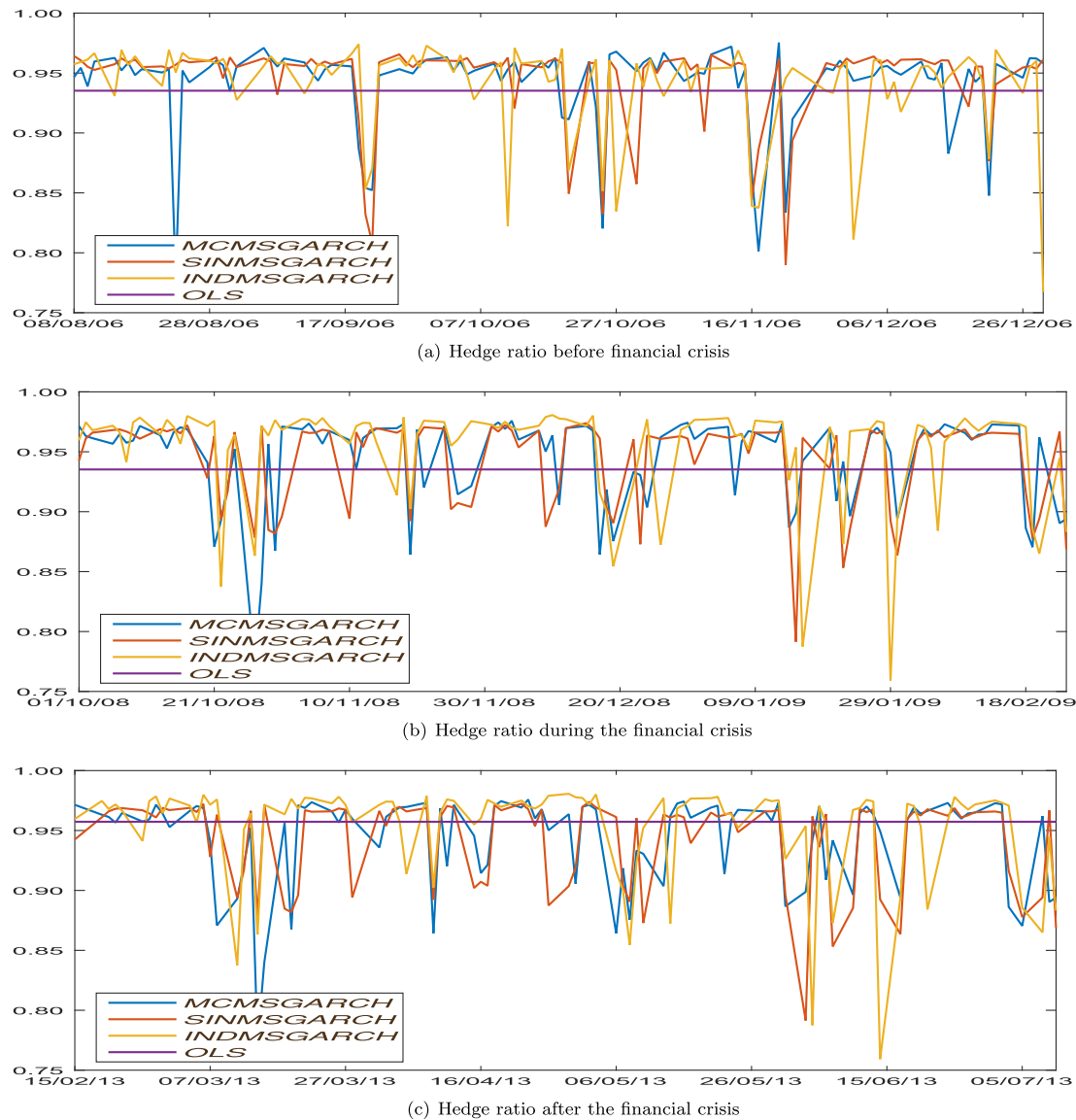


Fig. 7. Comparison of average hedge ratio for the unconstrained multichain MSGARCH (MC-f-MSGARCH), constrained MSGARCH (MC-c-MSGARCH) and single chain MSGARCH (SC-MSGARCH) models. First row pre-financial crisis (08/08/2006 to 03/01/2007); second row during crisis (01/10/2008 to 25/03/2009); and third row post-financial crisis (15/02/2013 to 31/07/2013).

In the above exercise, it is assumed that the prevailing state of the world is known. However the current state of the world cannot be correctly identified by the hedger. In this situation, the mean hedge ratio implemented is:

$$h_t = E_{\theta} E [v(s_t, z_t) | \mathcal{F}_{t-1}, \Theta] = E_{\theta} \left[\sum_{m, m'=1}^M v_m p(s_t = m, z_t = m' | \mathcal{F}_{t-1}, \Theta) \right], \quad (22)$$

and has been approximated using,

$$h_t = \frac{1}{G} \sum_{i=1}^G \sum_{m, n=1}^M v_{m,n}^{(i)} p^{(i)}(s_t = m, z_t = n | \mathcal{F}_{t-1}, \Theta), \quad (23)$$

where G is the number of Gibbs samples.

In Fig. 7, we report the estimation results for each model and compare them with the OLS hedge ratio over the three subsamples. The MS-GARCH hedge ratios display similar time-varying characteristics. However, we occasionally observe that the time-varying hedge ratios fall below the OLS hedge ratio. Also, the hedge ratios are observed to shift closer to 1 after the global financial crisis. This confirms our earlier intuition that hedgers seem to have become more careful in their investment decisions after the global financial crisis.

3.3. Hedging effectiveness

Following the estimation of the hedge ratios, we formally assess the performance of these hedges by first constructing the portfolio implied by the computed hedge ratios and then calculate the variance of the returns of these portfolios over each subsample period on a daily basis. In mathematical forms, we evaluate

$$V(RS_t - h_t^* RF_t) \quad (24)$$

where h_t^* are the estimated hedge ratios. The relative variance improvement of the MS-GARCH model against the OLS model is calculated as follows

$$\frac{V(\text{OLS}) - V(\text{MS-GARCH})}{V(\text{OLS})} \times 100, \quad (25)$$

where $V(\text{OLS})$ and $V(\text{MS-GARCH})$ are respectively the variance of the returns on the hedged portfolio (Eq. (24)) estimated using hedge ratios obtained from the OLS and MS-GARCH models. A positive value of Eq.(25) is an evidence that the MS-GARCH hedge ratio performs better than the OLS hedge ratio. Four different measures of the hedge ratio h_t^* in Eq.(25) are considered. The first is the average hedge ratio given by Eq.(23), the second is the average hedge ratio at time t given the most probable state at time $t - 1$, the third measure assumes the most probable hedge ratio at time t given the most state of the market at time t , and the fourth is hedge ratio obtained via the plug-in principle.

From Table 7, it appears that Markov-switching models provide more efficient hedge ratios relative to the OLS estimate, both before and during the 2008/2009 global financial crisis. The OLS hedge ratio, on the other hand, seems to perform better than MS-GARCH models after the financial crisis. This observation may be due to the low conditional variance of the markets after the 2008/2009 global financial crisis. Among the MS-GARCH specifications under consideration, the constrained multichain MS-GARCH model provides the most consistent measure of hedging effectiveness across the four different measures of hedge ratios used in the evaluation of Eq.(25), while, the unconstrained multichain MS-GARCH model provides the least effective hedging across the four different measures of hedge ratios considered. With the exception of a few cases, Bayes hedge ratio performs at least as good as the hedge ratio obtained by following the plug-in principle. Furthermore, prior to the financial crisis, the hedging effectiveness obtained using the most probable hedge ratio suggests that the OLS hedge ratio performs better than the unconstrained multichain model. This is in contrast with our observation when the average hedge ratio is applied. This result suggests that the unconstrained multichain model is flexible enough to detect events that are not evident when average hedge ratio is applied. This result is in line with the observation of Sephton (1998) who finds that the Regime Switching strategy outperforms both OLS and GARCH strategies

Table 7
Hedging effectiveness of MS-GARCH against constant hedge ratio.

| | Variance improvement | | | | | | | | | | | |
|--------------|--------------------------|--------|-------|---|--------|-------|-----------------------------------|--------|-------|-----------------------------------|--------|-------|
| | Bayes hedge ratio | | | | | | | | | Plug-in hedge ratio | | |
| | $h_t^* = E[v(s_t, z_t)]$ | | | $h_t^* = E[v_t \hat{s}_{t-1}, \hat{z}_{t-1}]$ | | | $h_t^* = v(\hat{s}_t, \hat{z}_t)$ | | | $h_t^* = E[v(s_t, z_t) \theta]$ | | |
| | Before | During | After | Before | During | After | Before | During | After | Before | During | After |
| SC-MSGARCH | 6.9 | 7.8 | −3.8 | 0.8 | 9.3 | 4.5 | 2.2 | 16.4 | 4.6 | 7.1 | 7.8 | −4.0 |
| MC-c-MSGARCH | 6.3 | 5.9 | −6.3 | 6.9 | 5.9 | −6.3 | 11.7 | 12.1 | −14.9 | 6.3 | 5.9 | −6.3 |
| MC-f-MSGARCH | 3.9 | 4.7 | −4.8 | 1.9 | 4.6 | 0.2 | −5.9 | 1.0 | −3.3 | 4.2 | 4.6 | −4.9 |

Notes: • $(\hat{s}_t, \hat{z}_t) = \arg \max_{s_t, z_t} p(s_t, z_t | \mathcal{F}_{t-1})$. • Variance Improvement of MS-GARCH models measures the incremental variance reduction of the MSGARCH models versus the OLS model. This is estimated using the formula: $(V(\text{OLS}) - V(\text{MSGARCH})) / V(\text{OLS}) \times 100$. • Before, during and after respectively signify the period before, during and after the 2008/2009 global financial crisis. • For each block in the Table, variance improvement is evaluated under different measures of the hedge ratio e.g. the hedge ratio used in columns 11 to 13 is obtained via the plug-in principle. The hedge ratio used in columns 2 to 4, columns 5 to 7, columns 8 to 10 and columns 11 to 13 are respectively the average hedge ratio given by Eq. (23), the average hedge ratio at time t given the most probable state at time $t - 1$, the most probable hedge ratio at time t given the most state of the market at time t and the hedge ratio obtained via the plug-in principle. • SC-MSGARCH stands for single chain MS-GARCH; MC-c-MSGARCH stands for constrained Multichain MS-GARCH model; and MC-f-MSGARCH stands for unconstrained Multichain MS-GARCH.

in the low variance state, but performs far worse than either strategy in the high variance state. This is an indication that multichain Markov-switching models have the potentials to compete favorably with other time-varying models.

It is important to emphasize that our measure of hedging effectiveness has been shown to be inadequate in evaluating minimum-variance hedge ratios other than OLS. See Lien (2005) and Lien (2009) for discussion. Based on this, alternative measures of effectiveness may provide better insights into the relative advantages of the multichain regime switching model. Following Lee and Yoder (2007a), we aim at providing a statistical comparison of the hedging effectiveness. In order to evaluate the significance of the different hedging strategies and since our models are Bayesians we follow Gelfand and Ghosh (1998) and Kadane and Lazar (2004) and we select the best model comparing observed data to predictors.

The goal of the Gelfand and Ghosh (1998) approach is to obtain good predictions for replicates of the observed data and at the same time, to be faithful to the observed values.

To attain this objective, in the spirit of Kadane and Lazar (2004) we consider the loss function

$$D_k(m) = \sum_{t=1}^T \min_{a_t} \mathbb{E}L(y_{1t}, y_{2t}, a_t) + kL(y_{10}, y_{20}, a_t) \quad (26)$$

for $k \geq 0$, where y_{10}, y_{20} are the observed data of the spot and future return rates and y_{1t}, y_{2t} are the replicates data of the spot and future return rates to be predicted with the same distribution of the observed data.

Under standard assumption of quadratic loss, $L(y_1, y_2, a) = (y_1 - ay_2)^2$, the general loss function given above can be written as

$$D_k(m) = \sum_{t=1}^T \mathbb{E} \left[(y_{t,1t} - a_t y_{t,2t})^2 + k(y_{t,10} - a_t y_{t,20})^2 \right] \quad (27)$$

with

$$a_t = \frac{\mu_{1t}\mu_{2t} + \sigma_{12t} + y_{t,10}y_{t,20}k}{\mu_{2t}^2 + \sigma_{2t}^2 + y_{t,20}^2 k}$$

where μ_{1t} and μ_{2t} are the posterior predictive means and σ_{1t}^2 and σ_{2t}^2 the posterior predictive variances of $y_{t,10}$ and $y_{t,20}$ given $y_{1:t-1,10}$ and $y_{1:t-1,20}$ and σ_{12t} the posterior predictive covariance between the two variables. Substituting for a_t in Eq. (27), we obtain

$$D_k(m) = \sum_{t=1}^T \mu_{1t}^2 + \sigma_{1t}^2 + ky_{t,10}^2 - \frac{(\mu_{1t}\mu_{2t} + \sigma_{12t} + ky_{t,10}y_{t,20})^2}{(\mu_{2t}^2 + \sigma_{2t}^2 + ky_{t,20}^2)}. \quad (28)$$

The weight k expresses how important it is to be close to $y_{1:T,10}$ and $y_{1:T,20}$ relative to $y_{1:T,1t}$ and $y_{1:T,2t}$. In Table 8 we consider a selection of MSGARCH models incorporating various interactions between the hedging regime of the portfolio and the volatility state of the future market at time t . Table 8 presents results for $D_k(m)$ for $k = 1, 3, 9$. From Table 8 the multichain MSGARCH models emerge as best choices for various values of k . The unconstrained multichain MSGARCH model on the other hand performed marginally better than the constrained multichain MSGARCH model. This result further suggests that multichains are required for effective performance of hedging models.

4. Conclusion

In this paper we propose a new Bayesian multichain MS-GARCH model with dependent chains. We apply the model to hedging in energy markets, thus extending the existing literature on MV hedging. The proposed parameterization of the multichain MS-GARCH model allows for a straightforward interpretation of the parameters of the models as level-shift and variance-covariance hedging components. Both the Bayesian model and the inference approach allow us to easily account for parameter uncertainty in the hedging decision. We apply this multichain MS-GARCH to estimate a state-dependent time-varying minimum variance hedge ratio and investigate the effect of relaxing the assumption of common switching

dynamic on the effectiveness of the hedging strategy. The practical implementation of our hedging model to crude oil spot and futures markets shows strong evidence in favor of the multichain MS-GARCH models when compared to other models, in terms of hedging effectiveness. Nevertheless, a sequential model comparison on the three sub-periods, i.e. before, during and after the 2008/2009 global financial crisis, provides evidence in favor of the best MS-GARCH models, before and during the financial crisis. In the period after the crisis, the reduction in the volatility level makes the MS-GARCH less appealing than the standard OLS approach. Based on this observation, we suggest that many and different models should be used in place of a single one in energy risk hedging since they could perform differently in various phases of the market.

First, the hedging strategy and measures of hedging effectiveness as considered in this paper ignore transaction cost. Generalizing our hedging framework by incorporating transaction cost may provide better insight into the practical usefulness of the proposed strategy. Also, the performance of our hedging strategy could be further enriched by accounting for model uncertainty. This may be achieved by embedding Bayesian Model Averaging (BMA) into our hedging strategy. Lastly, in order to create a balance between variance reduction and incremental transaction cost, alternative measures of hedging effectiveness such as a utility framework should be considered. See for example Alizadeh and Nomikos (2004) for an illustration.

Table 8

Model choice, m , for the MS-GARCH models.

| m | Model | $D_1(m)$ | $D_3(m)$ | $D_9(m)$ |
|-----|--------------|----------|----------|----------|
| 1 | SC-MSGARCH | 2942.3 | 3049.0 | 3166.1 |
| 2 | MC-c-MSGARCH | 1337.7 | 1754.1 | 2114.4 |
| 3 | MC-f-MSGARCH | 1223.7 | 1646.3 | 2065.0 |

Note: The loss function $D_k(m)$, $k = 1, 3, 9$ is evaluated using Eq. (28). SC-MSGARCH stands for single chain MS-GARCH; MC-c-MSGARCH stands for constrained Multichain MS-GARCH model; and MC-f-MSGARCH stands for unconstrained Multichain MS-GARCH.

Acknowledgments

We thank the seminar participants at the SIS 2013 Statistical Conference, Brescia, 2013. The authors' research is supported by funding from the European Union, Seventh Framework Programme FP7/2007–2013 under grant agreement SYRTO-SSH-2012-320270, by the Institut Europlace of Finance, "Systemic Risk grant", the Global Risk Institute in Financial Services, the Louis Bachelier Institute, "Systemic Risk Research Initiative", and by the Italian Ministry of Education, University and Research (MIUR) PRIN 2010–11 grant MISURA.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.eneco.2017.06.001>.

References

- Abul-Bashera, S., Sadosky, P., 2016. Hedging emerging market stock prices with oil, gold, VIX, and bonds: a comparison between DCC, ADCC and GO-GARCH. *Energy Econ.* 54, 235–247.
- Alizadeh, A.H., Nomikos, N.K., 2004. A Markov regime switching approach for hedging stock indices. *J. Futur. Mark.* 24, 649–674.
- Alizadeh, A.H., Nomikos, N.K., Pouliasis, P.K., 2008. A Markov regime switching approach for hedging energy commodities. *J. Bank. Financ.* 32, 1970–1983.
- Amzal, B., Bois, F.Y., Parent, E., Robert, C.P., 2006. Bayesian-optimal design via interacting particle systems. *J. Am. Stat. Assoc.* 101 (474), 773–785.
- Augustyniak, M., 2014. Maximum likelihood estimation of the Markov-switching GARCH model. *Comput. Stat. Data Anal.* 76, 61–75.
- Bauwens, L., Preminger, A., Rombouts, J., 2010. Theory and inference for a Markov switching GARCH model. *Econ. J.* 13, 218–244.
- Bawa, V., Brown, S., Klein, R., 1979. Estimation Risk and Optimal Portfolio Choice. *Studies in Bayesian Econometrics*, North-Holland Publishing Company.
- Bhar, R., Hammoudeh, S., 2011. Commodities and financial variables: analyzing relationships in a changing regime environment. *Int. Rev. Econ. Financ.* 469–484.
- Billio, M., Casarin, R., Osuntuyi, A., 2016. Efficient Gibbs sampling for Markov switching GARCH models. *Comput. Stat. Data Anal.* 100, 37–57.
- Billio, M., Di Sanzo, S., 2015. Granger-causality in Markov switching models. *J. Appl. Stat.* 42 (5), 956–966.
- Chang, C.Y., Lai, J.Y., Chuang, I.Y., 2010. Futures hedging effectiveness under the segmentation of bear/bull energy markets. *Energy Econ.* 32, 442–449.
- Chen, S., Lee, C., Shrestha, K., 2003. Futures hedge ratios: a review. *Q. Rev. Econ. Finance* 43, 433–465.
- Cifarelli, G., Paladino, G., 2015. A dynamic model of hedging and speculation in the commodity futures markets. *J. Financ. Mark.* 25, 1–15.
- Creti, A., Joëts, M., Mignon, V., 2013. On the links between stock and commodity markets' volatility. *Energy Econ.* 37, 16–28.
- Dark, J., 2015. Futures Hedging with Markov switching vector error correction FIE-GARCH and FIAPARCH. *J. Bank. Financ.* 61, S269–S285.
- DeGroot, M.H., 2005. Optimal statistical decisions. vol. 82. Wiley-Interscience.
- Ederington, L., 1979. The hedging performance of the new futures markets. *J. Financ.* 34 (1), 157–170.
- Gelfand, A.E., Ghosh, S.K., 1998. Model choice: a minimum posterior predictive loss approach. *Biometrika* 85 (1), 1–11.
- Geweke, J., 1992. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In: Bernardo, J.M., Berger, J.O., Dawid, A.P., Smith, A.F.M. (Eds.), *Bayesian Statistics 4*. Oxford University Press, Oxford, pp. 169–193.
- Gray, S.F., 1996. Modeling the conditional distribution of interest rates as a regime-switching process. *J. Financ. Econ.* 42, 27–62.
- Guidolin, M., Rinaldi, F., 2013. Ambiguity in asset pricing and portfolio choice: a review of the literature. *Theor. Decis. 74*, 183–217.
- Haigh, M.S., Holt, M.T., 2002. Crack spread hedging: accounting for time-varying volatility spillovers in the energy futures markets. *J. Appl. Economet.* 17, 269–289.
- Johnson, L.L., 1960. The theory of hedging and speculation in commodity futures. *Rev. Econ. Stud.* 27 (3), 139–151.
- Kadane, J.B., Lazar, N.A., 2004. Methods and criteria for model selection. *J. Am. Stat. Assoc.* 99 (465), 279–290.
- Kan, R., Zhou, G., 2007. Optimal portfolio choice with parameter uncertainty. *J. Financ. Quant. Anal.* 42 (3), 621.
- Kass, R.E., Raftery, A.E., 1995. Bayes factor. *J. Am. Stat. Assoc.* 90 (430), 773–795.
- Lai, Y., Lein, D., 2017. A bivariate high-frequency-based volatility model for optimal futures hedging. *J. Futur. Mark.*
- Lai, Y.S., Sheu, H.J., Lee, H.T., 2017. A multivariate Markov regime-switching high-frequency-based volatility model for optimal futures hedging. *J. Futur. Mark.*
- Lee, H., Yoder, J.K., 2007a. A bivariate Markov regime switching GARCH approach to estimate time varying minimum variance hedge ratio. *Appl. Econ.* 39 (10), 1253–1265.
- Lee, H., Yoder, J.K., 2007b. Optimal hedging with a regime switching time varying correlation GARCH model. *J. Futur. Mark.* 27 (5), 495–516.
- Lee, H.T., 2010. Regime switching correlation hedging. *J. Bank. Financ.* 34, 2728–2741.
- 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. *J. Futur. Mark.* 26 (2), 103–129.
- Lence, S.H., Hayes, D.J., 1994a. The empirical minimum-variance hedge. *Am. J. Agric. Econ.* 76 (1), 94–104.
- Lence, S.H., Hayes, D.J., 1994b. Parameter-based decision marketing under estimation risk: an application to futures trading. *J. Financ.* 49 (1), 345–357.
- Lien, D., 2005. A note on the superiority of the OLS hedge ratio. *J. Futur. Mark.* 25, 1121–1126.
- Lien, D., 2009. A note on the hedging effectiveness of GARCH models. *Int. Rev. Econ. Financ.* 18, 110–112.
- Müller, P., 1999. Simulation-based optimal design. *Bayesian Stat.* 6, 459–474.
- Müller, P., Sansó, B., Iorio, M.D., 2004. Optimal Bayesian design by inhomogeneous Markov chain simulation. *J. Am. Stat. Assoc.* 99 (467), 788–798.
- Myers, R., 1991. Estimating time-varying optimal hedge ratios on futures markets. *J. Futur. Mark.* 11, 39–53.
- Newton, M.A., Raftery, A.E., 1994. Approximate Bayesian inference by weighted likelihood bootstrap (with discussion). *J. R. Stat. Soc.* 56, 3–48.
- Pan, Z., Wang, Y., Yang, L., 2014. Hedging crude oil using refined product: a regime switching asymmetric DCC approach. *Energy Econ.* 46, 472–484.
- Philip, D., Shi, Y., 2016. Optimal hedging in carbon emission markets using Markov regime switching models. *J. Int. Financ. Mark. Inst. Money* 43, 1–15.
- Phillips, P.C.B., 2011. Dating the timeline of financial bubbles during the subprime crisis. *Quant. Econ.* 2, 455–491.
- Raiffa, H., Schlaifer, R., 1961. *Applied Statistical Decision Theory*. Harvard University Press.
- Sephton, P., 1998. GARCH and Markov hedging at the Winnipeg Commodity Exchange. *Can. J. Agric. Econ.* 46, 117–126.
- Sheu, H.-J., Lee, H.-T., 2014. Optimal futures hedging under multichain Markov regime switching. *J. Futur. Mark.* 32 (2), 172–202.
- Su, E., 2017. Stock index hedging using a trend and volatility regime-switching model involving hedging Cost. *Int. Rev. Econ. Financ.* 47, 233–254.
- Wang, Y., Wu, C., 2012. Energy prices and exchange rates of the U.S. dollar: further evidence from linear and nonlinear causality analysis. *Econ. Model.* 29 (6), 2289–2297.