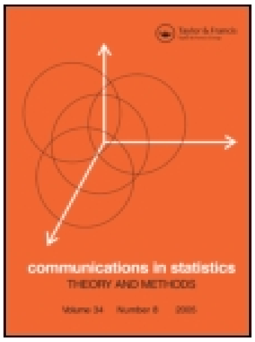

Communications in Statistics - Theory and Methods

ISSN: 0361-0926 (Print) 1532-415X (Online) Journal homepage: http://www.tandfonline.com/loi/lsta20

Markov switching component GARCH model: Stability and forecasting

N. Alemohammad, S. Rezakhah & S. H. Alizadeh

To cite this article: N. Alemohammad, S. Rezakhah & S. H. Alizadeh (2016) Markov switching component GARCH model: Stability and forecasting, Communications in Statistics - Theory and Methods, 45:15, 4332-4348, DOI: 10.1080/03610926.2013.841934

To link to this article: http://dx.doi.org/10.1080/03610926.2013.841934

COMMUNICATIONS IN STATISTICS—THEORY AND METHODS
2016, VOL. 45, NO. 15, 4332–4348
http://dx.doi.org/10.1080/03610926.2013.841934

Taylor & Francis
Taylor & Francis Group

Markov switching component GARCH model: Stability and forecasting

N. Alemohammad[a], S. Rezakhah[a], and S. H. Alizadeh[b]

[a]Faculty of Mathematics and Computer Science, Amirkabir University of Technology, Tehran, Iran; [b]Department of Computer Engineering and Information Technology, Qazvin Branch, Islamic Azad University, Qazvin, Iran

ABSTRACT

This paper introduces an extension of the Markov switching GARCH model where the volatility in each state is a convex combination of two different GARCH components with time varying weights. This model has the dynamic behavior to capture the variants of shocks. The asymptotic behavior of the second moment is investigated and an appropriate upper bound for it is evaluated. Using the Bayesian method via Gibbs sampling algorithm, a dynamic method for the estimation of the parameters is proposed. Finally, we illustrate the efficiency of the model by simulation and also by considering two different set of empirical financial data. We show that this model provides much better forecasts of the volatility than the Markov switching GARCH model.

ARTICLE HISTORY
Received 14 October 2012
Accepted 29 August 2013

KEYWORDS
Bayesian inference; Component GARCH models; Forecasting; GARCH models; Griddy Gibbs sampling; Markov process; Stability.

MATHEMATICS SUBJECT CLASSIFICATION
60J10; 62M10; 62F15

1. Introduction

In the past three decades, there has been a growing interest in using non linear time series models in finance and economics. For financial time series, the ARCH and GARCH models, introduced by Engle (1982) and Bollerslev (1986), are surely the most popular classes of volatility models. Although these models have been applied extensively in the modeling of financial time series, the dynamic structure of volatility cannot be captured passably by such models. For more flexible volatility modeling, models with time varying parameters are introduced. One class of such models is that of smooth transition GARCH models presented by Gonzalez-Rivera (1998), lubrano (2001) (see also Hagerud (1997) and Medeiros and Veiga (2009)). The component GARCH models, introduced first by Ding and Granger (1996), provide another generalization of GARCH models. The structure of the component GARCH model (Ding and Granger (1996)) is established by introducing a linear combination of two GARCH components to contribute to the overall variance at time t, where shocks to one component die out rapidly, whereas the other component is rather persistent. Such models have been applied in modeling financial time series (e.g., Maheu (2005) and Engle and Rangel (2008)). A generalization of the component GARCH model of Ding and Granger is the weighted GARCH model that is proposed by Bauwens and Storti (2009). They considered the weights of GARCH components to be functions of lagged values of the conditional standard deviation or squared past observations.

CONTACT S. Rezakhah rezakhah@aut.ac.ir Faculty of Mathematics and Computer Science, Amirkabir University of Technology, Tehran, Iran.
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/lsta.

Markov switching also provides a nice generalization for (G)ARCH models. These models are obtained by merging the (G)ARCH model with a Markov process, where each state of the Markov model allows a different (G)ARCH behavior. Such models have been introduced by Cai (1994) and Hamilton and Susmel (1994). This feature extends the dynamic formulation of the model and potentially enables improving forecasts of the volatility (Abramson and Cohen, 2007). Gray (1996), Klaassen (2002), and Haas et al. (2004) proposed different variants of Markov-switching GARCH models. See also further studies, Abramson and Cohen (2007), Alexander and Lazar (2008), and Bauwens et al. (2010).

In this paper, we consider a Markov switching model where the volatility of each state is a time varying convex combination of two GARCH models which are in effect of the last observed volatility. This model has the potential to switch between regimes with various level of volatilities and also is able to consider variant volatilities in each level. As an explanation in the state of high volatility we could have a dynamic time varying combination of very high volatilities and high volatilities, and in the state of low volatility we could have also such combination of low volatilities and moderate ones, as often observed in financial time series.

The dynamic structure of this model has the potential to capture sudden changes and moreover to account gradual changes by the smooth transition of time varying weights in each state. As using all past observations for forecasting could increase the complexity of the model, we reduce the volume of calculations by proposing a dynamic programming algorithm. We derive necessary and sufficient conditions for stability and obtain an upper bound for the limit of the second moment by using the method of Abramson and Cohen (2007) and Medeiros and Veiga (2009). For the estimation of the parameters, we use the Bayesian inference via the Gibbs sampling. We compare the performance of our model with the Markov switching GARCH model. The Markov switching component GARCH model can forecast the conditional variance much better than MS-GARCH model.

The paper is organized as follows: in Sec. 2 we introduce the Markov switching component GARCH model. Section 3 investigates the statistical properties of the model. Section 4 is devoted to the estimation of the parameters of the model. Section 5 is dedicated to the analyzing of the efficiency of the proposed model through simulation and the comparison of the forecast errors with the MS-GARCH model. The performance of our model has been compared with MS-GARCH for empirical financial indices and discussed in Sec. 6. Section 7 concludes.

2. Markov switching component GARCH model

The Markov switching component GARCH model, MS-CGARCH, for time series $\{y_t\}$ is defined as

$$y_t = \varepsilon_t \sqrt{H_{t,Z_t}}, \tag{2.1}$$

where $\{\varepsilon_t\}$ are iid standard normal variables, $\{Z_t\}$ is an irreducible and aperiodic Markov chain on finite state space $E = \{1, 2, \ldots, K\}$ with transition probability matrix $P = ||p_{ij}||_{K\times K}$, where $p_{ij} = p(Z_t = j|Z_{t-1} = i)$, $i, j \in \{1, \ldots, K\}$, and stationary probability measure $\pi = (\pi_1, \ldots, \pi_K)'$. Also given that $Z_t = j$, $H_{t,j}$ (the conditional variance in regime j) is driven by

$$H_{t,j} = w_{t,j} h_{1,t,j} + (1 - w_{t,j}) h_{2,t,j}, \tag{2.2}$$

where

$$h_{1,t,j} = a_{0j} + a_{1j} y_{t-1}^2 + a_{2j} H_{t-1,j}, \tag{2.3}$$

$$h_{2,t,j} = b_{0j} + b_{1j}y_{t-1}^2 + b_{2j}H_{t-1,j}, \tag{2.4}$$

and each of the weights ($w_{t,j}$) is a function of the past observation as

$$w_{t,j} = \frac{1 - \exp(-\gamma_j |y_{t-1}|)}{1 + \exp(-\gamma_j |y_{t-1}|)} \quad \gamma_j > 0, \tag{2.5}$$

which is bounded, $0 < w_{t,j} < 1$. The parameter γ_j is called the slope parameter, that explains the speed of transition from one component to the other one: the higher γ_j, the faster the transition. $H_{t-1,j}$ in (2.2) is the conditional variance of state j at time $t-1$, that is a combination of conditional variances of both components at the state j. Since $\gamma_j > 0$, when the absolute value of y_{t-1} increases, the impact of $h_{1,t,j}$ increases and consequently the effect of $h_{2,t,j}$ decreases and vice versa.

The MS-CGARCH structure (2.1)–(2.5) represents that the total volatility in each state ($H_{t,j}$) is in effect of the volatilities of two components ($h_{1,t,j}$ and $h_{2,t,j}$), and the volatility of these components are in effect of the total volatility at previous step ($H_{t-1,j}$). This implies that as in Markov switching GARCH the total volatility of each state ($H_{t,j}$) to be updated by its previous value. It is imposed that the volatility of the components in each state have effects on the next volatility of the other component via the total volatility as it is expected in many macroeconomic factors and stock returns. For more elaboration see Booth and Ciner (2001), Baumeister et al. (2008), and Hyde and Bredin (2005) .

Another good feature of our model is that it overcomes the problem of path dependency[1] (that appears in some MS-GARCH specifications).

If $w_{t,j}$ becomes a constant value (for example, when γ_j tending to zero or infinity), the MS-CGARCH model will be a MS-GARCH model. In the case of single regime, if $a_{2.} = b_{2.}$, our model is the generalization of the smooth transition GARCH model that is introduced by lubrano (2001).

It is assumed that $\{\varepsilon_t\}$ and $\{Z_t\}$ are independent. Sufficient conditions to guarantee strictly positive conditional variance are a_{0j}, b_{0j} to be positive and a_{1j}, a_{2j}, b_{1j}, b_{2j} being non negative.

Let $\mathcal{I}_t$ be the observation set up to time t. Following the method of Alizadeh and Rezakhah (2013), the conditional density function of y_t given past observations is obtained as follows:

$$\begin{aligned} f(y_t|\mathcal{I}_{t-1}) &= \sum_{j=1}^{K} f(y_t, Z_t = j|\mathcal{I}_{t-1}) \\ &= \sum_{j=1}^{K} p(Z_t = j|\mathcal{I}_{t-1}) f(y_t|\mathcal{I}_{t-1}, Z_t = j) \\ &= \sum_{j=1}^{K} \alpha_j^{(t)} \phi\left(\frac{y_t}{\sqrt{H_{t,j}}}\right) \end{aligned} \tag{2.6}$$

in which $\alpha_j^{(t)} = p(Z_t = j|\mathcal{I}_{t-1})$ (that is obtained in next section), and $\phi(.)$ is the probability density function of the standard normal distribution.

[1] Path dependency happens when the volatility of each regime at time t depends on the entire sequence of past regimes because of the recursive property of GARCH processes.

3. Statistical properties of the model

In this section, the statistical properties of the MS-CGARCH model are investigated and the conditional variance of the process is obtained. We show that the model, under some conditions on coefficients and transition probabilities, is asymptotically stable in the second moment. An appropriate upper bound for the limiting value of the second moment is obtained.

3.1. *Forecasting*

The forecasting volatility (conditional variance) of MS-CGARCH model is given by

$$Var(Y_t|\mathcal{I}_{t-1}) = \sum_{j=1}^{K} \alpha_j^{(t)} H_{t,j} = \sum_{j=1}^{K} \alpha_j^{(t)} (w_{t,j} h_{1,t,j} + (1 - w_{t,j}) h_{2,t,j}). \tag{3.7}$$

This relation shows that the conditional variance of this model is affected by the changes in states, the volatility of components, and the weight functions in each state.

At each time t, $\alpha_j^{(t)}$ (in Eqs. (2.6) and (3.7)) can be obtained from a dynamic programming method based on forward recursion algorithm, proposed in remark (3.1).

Remark 3.1. The value of $\alpha_j^{(t)}$ is obtained recursively by

$$\alpha_j^{(t)} = \frac{\sum_{m=1}^{K} f(y_{t-1}|Z_{t-1} = m, \mathcal{I}_{t-2}) p(Z_{t-1} = m|\mathcal{I}_{t-2}) p_{m,j}}{\sum_{m=1}^{K} f(y_{t-1}|Z_{t-1} = m, \mathcal{I}_{t-2}) p(Z_{t-1} = m|\mathcal{I}_{t-2})}. \tag{3.8}$$

Proof. As the hidden variables $\{Z_t\}_{t\geq 1}$ have Markov structure in the MS-CGARCH model, so

$$\begin{aligned}
\alpha_j^{(t)} &= p(Z_t = j|\mathcal{I}_{t-1}) = \sum_{m=1}^{K} P(Z_t = j, Z_{t-1} = m|\mathcal{I}_{t-1}) \\
&= \sum_{m=1}^{K} p(Z_t = j|Z_{t-1} = m, \mathcal{I}_{t-1}) p(Z_{t-1} = m|\mathcal{I}_{t-1}) \\
&= \sum_{m=1}^{K} p(Z_t = j|Z_{t-1} = m) p(Z_{t-1} = m|\mathcal{I}_{t-1}) \\
&= \frac{\sum_{m=1}^{K} f(\mathcal{I}_{t-1}, Z_{t-1} = m) p_{m,j}}{\sum_{m=1}^{K} f(\mathcal{I}_{t-1}, Z_{t-1} = m)} \\
&= \frac{\sum_{m=1}^{K} f(y_{t-1}|Z_{t-1} = m, \mathcal{I}_{t-2}) p(Z_{t-1} = m|\mathcal{I}_{t-2}) p_{m,j}}{\sum_{m=1}^{K} f(y_{t-1}|Z_{t-1} = m, \mathcal{I}_{t-2}) p(Z_{t-1} = m|\mathcal{I}_{t-2})},
\end{aligned} \tag{3.9}$$

where

$$f(y_{t-1}|Z_{t-1} = m, \mathcal{I}_{t-2}) = \phi\left(\frac{y_{t-1}}{\sqrt{H_{t-1,m}}}\right).$$

□

3.2. *Stability*

In this subsection, we investigate the stability of the second moment of MS-CGARCH model. Indeed we are looking for an upper bound for the second moment of our model. The second moment of the model can be calculated as

$$\begin{aligned} E(y_t^2) &= E(H_{t,Z_t}) = E_{Z_t}[E_{t-1}(H_{t,Z_t}|z_t)] \\ &= \sum_{z_t=1}^{K} \pi_{z_t} E_{t-1}(H_{t,Z_t}|z_t). \end{aligned} \tag{3.10}$$

$E_t(\cdot)$ denotes the expectation with respect to the information up to time t. We use $E(\cdot|z_t)$ and $p(\cdot|z_t)$ to denote $E(\cdot|Z_t = z_t)$ and $P(\cdot|Z_t = z_t)$, respectively, where z_t is the realization of the state at time t. We investigate the conditional variance under the chain state, m, as follows:

$$\begin{aligned} E_{t-1}(H_{t,m}|z_t) &= E_{t-1}[w_{t,m}(a_{0m} + a_{1m}y_{t-1}^2 + a_{2m}H_{t-1,m})|z_t] \\ &\quad + E_{t-1}[(1 - w_{t,m})(b_{0m} + b_{1m}y_{t-1}^2 + b_{2m}H_{t-1,m})|z_t] \\ &= b_{0m} + \underbrace{b_{1m}E_{t-1}[y_{t-1}^2|z_t]}_{I} + \underbrace{(a_{0m} - b_{0m})E_{t-1}[w_{t,m}|z_t]}_{II} \\ &\quad + \underbrace{b_{2m}E_{t-1}(H_{t-1,m}|z_t)}_{III} \\ &\quad + \underbrace{(a_{1m} - b_{1m})E_{t-1}[w_{t,m}y_{t-1}^2|z_t]}_{IV} + \underbrace{(a_{2m} - b_{2m})E_{t-1}(w_{t,m}H_{t-1,m}|z_t)}_{V} \end{aligned} \tag{3.11}$$

The term (I) in (3.11) can be evaluated as

$$\begin{aligned} E_{t-1}[y_{t-1}^2|z_t] &= \sum_{z_{t-1}=1}^{K} \int_{S_{\mathcal{I}_{t-1}}} y_{t-1}^2 p(\mathcal{I}_{t-1}|z_t, z_{t-1}) p(z_{t-1}|z_t) d\mathcal{I}_{t-1} \\ &= \sum_{z_{t-1}=1}^{K} p(z_{t-1}|z_t) E_{t-1}[y_{t-1}^2|z_{t-1}, z_t], \end{aligned} \tag{3.12}$$

where $S_{\mathcal{I}_{t-1}}$ is the support of $\mathcal{I}_{t-1} = (y_1, \ldots, y_{t-1})$. Since the expected value of y_{t-1}^2 given z_{t-1} is independent of any future state, so

$$E_{t-1}[y_{t-1}^2|z_{t-1}, z_t] = E_{t-1}[y_{t-1}^2|z_{t-1}]. \tag{3.13}$$

Also using the tower property of the conditional expectation, $E[E(Y|X, Z)|X] = E(Y|X)$ (see Grimmett and Stirzaker (2001, p. 69)), we have

$$\begin{aligned} E_{t-1}[y_{t-1}^2|z_{t-1}] &= E_{t-2}\big[E_{t-1}\big(y_{t-1}^2\big|\mathcal{I}_{t-2}, z_{t-1}\big)|z_{t-1}\big] \\ &= E_{t-2}[H_{t-1,Z_{t-1}}|z_{t-1}]. \end{aligned} \tag{3.14}$$

The calculation of $E_{t-1}[w_{t,m}|z_t]$, $E_{t-1}[w_{t,m}y_{t-1}^2|z_t]$, and $E_{t-1}(w_{t,m}H_{t-1,m}|z_t)$ is a problem that cannot be easily done, for this reason we will try to find an upper bound for them.

Upper bound for II in (3.11): As $0 < w_{t,m} < 1$, so an upper bound for the relation II in (3.11) is obtained by

$$(a_{0m} - b_{0m})E_{t-1}[w_{t,m}|z_t] \le |a_{0m} - b_{0m}| < \infty. \tag{3.15}$$

The term (III) in (3.11) can be evaluated as

$$\begin{aligned} b_{2m}E_{t-1}(H_{t-1,m}|z_t) &= b_{2m}\int_{S_{\mathcal{I}_{t-1}}} H_{t-1,m}p(\mathcal{I}_{t-1}|z_t)d\mathcal{I}_{t-1} \\ &= b_{2,m} \sum_{z_{t-1}=1}^{K} p(z_{t-1}|z_t)E_{t-2}(H_{t-1,m}|z_{t-1}). \end{aligned} \tag{3.16}$$

Upper bound for IV in (3.11): Let $0 < M < \infty$ be a constant, so

$$E_{t-1}[w_{t,z_t} y_{t-1}^2 | z_t] = E_{t-1}[w_{t,z_t} y_{t-1}^2 I_{|y_{t-1}|<M} | z_t] \\ + E_{t-1}[w_{t,z_t} y_{t-1}^2 I_{|y_{t-1}|\geq M} | z_t]$$

in which

$$I_{x<a} = \begin{cases} 1 & \text{if} x < a \\ 0 & \text{otherwise.} \end{cases}$$

As by (2.5), $0 < w_{t,z_t} < 1$ and so

$$E_{t-1}[w_{t,z_t} y_{t-1}^2 | z_t] \leq M^2 + E_{t-1}[w_{t,z_t} y_{t-1}^2 I_{|y_{t-1}|\geq M} | z_t],$$

also

$$E_{t-1}[w_{t,z_t} y_{t-1}^2 I_{|y_{t-1}|\geq M} | z_t] = \int_{S_{\mathcal{I}_{t-2}}, y_{t-1} \leq -M} y_{t-1}^2 [w_{t,z_t}] p(\mathcal{I}_{t-1} | z_t) d\mathcal{I}_{t-1} \\ + \int_{S_{\mathcal{I}_{t-2}}, y_{t-1} \geq M} y_{t-1}^2 [w_{t,z_t}] p(\mathcal{I}_{t-1} | z_t) d\mathcal{I}_{t-1},$$

by (2.5),

$$\lim_{y_{t-1} \to +\infty} w_{t,z_t} = 1 \quad \text{and} \quad \lim_{y_{t-1} \to -\infty} w_{t,z_t} = 1, \tag{3.17}$$

so for $\epsilon > 0$, there exist $M > 0$ such that if $y_{t-1} \geq M$, then $|w_{t,z_t} - 1| \leq \epsilon$ and if $y_{t-1} \leq -M$, then $|w_{t,z_t} - 1| \leq \epsilon$. Hence

$$E_{t-1}[w_{t,z_t} y_{t-1}^2 I_{|y_{t-1}|\geq M} | z_t] \leq (\epsilon + 1) \int_{S_{\mathcal{I}_{t-2}}, y_{t-1} \leq -M} y_{t-1}^2 p(\mathcal{I}_{t-1} | z_t) d\mathcal{I}_{t-1} \\ + (\epsilon + 1) \int_{S_{\mathcal{I}_{t-2}}, y_{t-1} \geq M} y_{t-1}^2 p(\mathcal{I}_{t-1} | z_t) d\mathcal{I}_{t-1}.$$

Since the distribution of the $\{\varepsilon_t\}$ is symmetric, then

$$(\epsilon + 1) \int_{S_{\mathcal{I}_{t-2}}, y_{t-1} \leq -M} y_{t-1}^2 p(\mathcal{I}_{t-1} | z_t) d\mathcal{I}_{t-1} \leq (\epsilon + 1) \int_{S_{\mathcal{I}_{t-2}}, -\infty < y_{t-1} < 0} y_{t-1}^2 p(\mathcal{I}_{t-1} | z_t) d\mathcal{I}_{t-1} \\ = (\epsilon + 1) \frac{E_{t-1}[y_{t-1}^2 | z_t]}{2}$$

and

$$(\epsilon + 1) \int_{S_{\mathcal{I}_{t-2}}, y_{t-1} \geq M} y_{t-1}^2 p(\mathcal{I}_{t-1} | z_t) d\mathcal{I}_{t-1} \leq (\epsilon + 1) \int_{S_{\mathcal{I}_{t-2}}, 0 < y_{t-1} < \infty} y_{t-1}^2 p(\mathcal{I}_{t-1} | z_t) d\mathcal{I}_{t-1} \\ = (\epsilon + 1) \frac{E_{t-1}[y_{t-1}^2 | z_t]}{2}.$$

Therefore,

$$(a_{1m} - b_{1m}) E_{t-1}[w_{t,z_t} y_{t-1}^2 | z_t] \leq |a_{1m} - b_{1m}| (M^2 + (\epsilon + 1) E_{t-1}[y_{t-1}^2 | z_t]).$$

Upper bound for V in (3.11): Since $0 < w_{t,m} < 1$, so

$$(a_{2m} - b_{2m}) E_{t-1}(w_{t,m} H_{t-1,m} | z_t) \leq |a_{2m} - b_{2m}| E_{t-1}(H_{t-1,m} | z_t). \tag{3.18}$$

By replacing the obtained upper bounds and relations (3.12)–(3.14) in (3.11), the upper bound for $E_{t-1}(H_{t,Z_t}|z_t)$ is acquired as

$$\begin{aligned} E_{t-1}(H_{t,m}|z_t) &\leq a_{0m} + |a_{1m} - b_{1m}|M^2 \\ &+ \sum_{z_{t-1}=1}^{K} [b_{1m} + |a_{1m} - b_{1m}|(\epsilon + 1)] p(z_{t-1}|z_t) E_{t-2}[H_{t-1,Z_{t-1}}|z_{t-1}] \\ &+ \sum_{z_{t-1}=1}^{K} a_{2m} p(z_{t-1}|z_t) E_{t-2}[H_{t-1,m}|z_{t-1}], \end{aligned} \tag{3.19}$$

in which by Bayes' rule

$$p(z_{t-1}|z_t) = \frac{\pi_{z_{t-1}}}{\pi_{z_t}} \{P_{z_{t-1}z_t}\},$$

where P is the transition probability matrix. Let

$$\Omega = [a_{01} + |a_{11} - b_{11}|M^2, \ldots, a_{0K} + |a_{1K} - b_{1K}|M^2]', \tag{3.20}$$

be a vector with K component, C denotes a K^2-by-K^2 block matrix as

$$C = \begin{pmatrix} C_{11} & C_{21} & \cdots & C_{K1} \\ C_{12} & C_{22} & \cdots & C_{K2} \\ \vdots & & & \vdots \\ C_{1K} & C_{2K} & \cdots & C_{KK} \end{pmatrix} \tag{3.21}$$

with each block given by

$$C_{jk} = p(Z_{t-1} = j|Z_t = k)(ue_j' + \upsilon), \qquad j, k = 1, \ldots, K, \tag{3.22}$$

where $u = [b_{11} + (\epsilon + 1)|a_{11} - b_{11}|, \ldots, b_{1K} + (\epsilon + 1)|a_{1K} - b_{1K}|]'$, e_j is a K-by-1 vector of all zeros, except its jth element, which is one, and υ is a diagonal K-by-K matrix with elements $[a_{21}, \ldots, a_{2K}]$ on its diagonal.

Let $A_t(j, k) = E_{t-1}[H_{t,j}|Z_t = k]$, $A_t = [A_t(1, 1), A_t(2, 1), \ldots, A_t(K, 1), A_t(1, 2), \ldots, A_t(K, K)]$ be a K^2-by-1 vector and consider $\dot{\Omega} = (\Omega', \ldots, \Omega')'$ be a vector that is made of K vector Ω.

Hence by (3.20)–(3.22) we have the following recursive inequality vector form for A_t, as

$$A_t \leq \dot{\Omega} + CA_{t-1}, \quad t \geq 0 \tag{3.23}$$

with some initial conditions A_{-1}.

Let $\Pi = [\pi_1 e_1', \ldots, \pi_K e_K']$ and consider $\rho(A)$ denotes the spectral radius of a matrix A, then we have the following theorem for the stationarity condition of the MS-CGARCH model.

Theorem 3.1. *Let $\{Y_t\}_{t=0}^{\infty}$ follows the MS-CGARCH model, defined by (2.1)–(2.5), the process is asymptotically stable in variance and $\lim_{t\to\infty} E(Y_t^2) \leq \Pi'(I - C)^{-1}\dot{\Omega}$, if $\rho(C) < 1$.*

Proof. The recursive inequality (3.23) implies that

$$A_t \leq \dot{\Omega} \sum_{i=0}^{t-1} C^i + C^t A_0 := B_t. \tag{3.24}$$

Following the matrix convergence theorem Lancaster and Tismenetsky (1985), the necessary and sufficient condition for the convergence of B_t when $t \to \infty$ is that $\rho(C) < 1$. Under this condition, C^t converges to zero as t goes to infinity and $\sum_{i=0}^{t-1} C^i$ converges to $(I - C)^{-1}$ provided that matrix $(I - C)$ is invertible. So if $\rho(C) < 1$,

$$\lim_{t\to\infty} A_t \leq (I - C)^{-1}\dot{\Omega}.$$

By (3.10) the upper bound for the asymptotic behavior of unconditional variance is given by

$$lim_{t\to\infty} E\left(y_t^2\right) \leq \Pi'(I - C)^{-1}\dot{\Omega}.$$

□

4. Estimation

In this section, we describe the estimation of the parameters of the MS-CGARCH model. We consider Bayesian MCMC method using Gibbs algorithm by following methods of sampling of a hidden Markov process (Chib (1996) and Kaufman and Fruhwirth-Schnatter (2002)), MS-GARCH model, and weighted GARCH model (Bauwens and Storti (2009) and Bauwens et al. (2010)) for estimation of the parameters.

Let $Y_t = (y_1, \ldots, y_t)$ and $Z_t = (z_1, \ldots, z_t)$. For the case of two states, the transition probabilities are $\eta = (\eta_{11}, \eta_{12}, \eta_{21}, \eta_{22})$ and the parameters of the model are $\theta = (\theta_1, \theta_2)$, where $\theta_k = (a_{0k}, b_{0k}, a_{1k}, b_{1k}, a_{2k}, b_{2k}, \gamma_k)$ for $k = 1, 2$.

The purpose of Bayesian inference is to simulate from the distributions of the parameters and the state variables given the observations. As $Z = (z_1, \ldots, z_T)$ and $Y = (y_1, \ldots, y_T)$, the posterior density of our model is

$$p(\theta, \eta, Z|Y) \propto p(\theta, \eta)p(Z|\theta, \eta)f(Y|\theta, \eta, Z), \tag{4.25}$$

in which $p(\theta, \eta)$ is the prior of the parameters. The conditional probability mass function of Z given the (θ, η) is independent of θ, so

$$\begin{aligned} p(Z|\theta, \eta) &= p(Z|\eta_{11}, \eta_{22}) \\ &= \prod_{t=1}^{T} p(z_{t+1}|z_t, \eta_{11}, \eta_{22}) \\ &= p_{11}^{n_{11}}(1 - p_{11})^{n_{12}} p_{22}^{n_{22}}(1 - p_{22})^{n_{21}}, \end{aligned} \tag{4.26}$$

where $n_{ij} = \#\{z_t = j|z_{t-1} = i\}$. The conditional density function of Y given the realization of Z and the parameters is factorized in the following way:

$$f(Y|\eta, \theta, Z) = \prod_{t=1}^{T} f(y_t|\theta, z_t = k, Y_{t-1}), \quad k = 1, 2, \tag{4.27}$$

where the one step ahead of the predictive densities are

$$f(y_t|\theta, z_t = k, Y_{t-1}) = \frac{1}{\sqrt{2\pi H_{t,k}}} \exp\left(-\frac{y_t^2}{H_{t,k}}\right). \tag{4.28}$$

Since the posterior density (4.25) is not standard we cannot sample it in a straightforward manner. Gibbs sampling of Gelfand and Smith (1990) is a repetitive algorithm to sample consecutively from the posterior distribution. Under regularity conditions, the simulated distribution converges to the posterior distribution (see, e.g., Robert and Casella (2004)). The blocks of parameters are θ, η, and the realizations of Z.

A brief description of the Gibbs algorithm: let us use the superscript (r) on Z, θ, and η to denote the estimators of Z, η, and θ at the rth iteration of the algorithm. Each iteration of the algorithm consists of three steps:

(i) Drawing an estimator random sample of the state variable $Z^{(r)}$ given $\eta^{(r-1)}$ and $\theta^{(r-1)}$.
(ii) Drawing a random sample of the transition probabilities $\eta^{(r)}$ given $Z^{(r)}$.
(iii) Drawing a random sample of the $\theta^{(r)}$ given $Z^{(r)}$ and $\eta^{(r)}$.

These steps are repeated until the convergency is obtained. In what follows sampling of each block is explained.

4.1. *Sampling z_t*

The purpose of this step is to obtain the sample of $p(z_t|\eta, \theta, Y_t)$ that is performed by Chib (1996) (see also Kaufman and Fruhwirth-Schnatter (2002)). Suppose $p(z_1|\eta, \theta, Y_0,)$ be the stationary distribution of the chain,

$$p(z_t|\eta, \theta, Y_t) \propto f(y_t|\theta, z_t = k, Y_{t-1})p(z_t|\eta, \theta, Y_{t-1}), \tag{4.29}$$

where the predictive density $f(y_t|\theta, z_t = k, Y_{t-1})$ is calculated by the relation (4.28) and by the law of total probability $p(z_t|\eta, \theta, Y_{t-1})$ is given by

$$p(z_t|\eta, \theta, Y_{t-1}) = \sum_{z_{t-1}=1}^{K} p(z_{t-1}|\eta, \theta, Y_{t-1})\eta_{z_{t-1}z_t}. \tag{4.30}$$

Given the filter probabilities $(p(z_t|\eta, \theta, Y_t))$, we run a backward algorithm, starting from $t = T$ that z_T is derived from $p(z_T|\eta, \theta, Y)$. For $t = T - 1, \dots, 0$ the sample is derived from $p(z_t|z_{t+1}, \dots, z_T, \theta, \eta, Y)$,which is obtained by

$$p(z_t|z_{t+1}, \dots, z_T, \theta, \eta, Y) \propto p(z_t|\eta, \theta, Y_t)\eta_{z_t,z_{t+1}}.$$

To derive z_t from $p(z_t|\cdot) = p_{z_t}$ is by sampling from the conditional probabilities, for example, $q_1 = p(Z_t = 1|Z_t \geq 1, .)$ which are given by

$$p(Z_t = 1|Z_t \geq 1, \dots) = \frac{p_1}{\sum_{l=1}^{2} p_l}.$$

After generating a number U from uniform (0,1), if $U \leq q_1$ then $z_t = 1$, otherwise $z_t = 2$.

4.2. *Sampling η*

This stage is devoted to sample $\eta = (\eta_{11}, \eta_{22})$ from the posterior probability $p(\eta|\theta, Y_t, Z_t)$ that is independent of Y_t, θ. We consider independent beta prior density for each of η_{11} and η_{22}. For example,

$$p(\eta_{11}|Z_t) \propto p(\eta_{11})p(Z_t|\eta_{11}) = \eta_{11}^{c_{11}+n_{11}-1}(1 - \eta_{11})^{c_{12}+n_{12}-1},$$

where c_{11} and c_{12} are the parameters of beta prior, n_{ij} is the number of transition from $z_{t-1} = i$ to $z_t = j$. In the same way the sample of η_{22} is obtained.

4.3. Sampling θ

The posterior density of θ given the prior $p(\theta)$ is given by

$$p(\theta|Y,Z,\eta) \propto p(\theta)\prod_{t=1}^{T} f(y_t|\theta, z_t = k, Y_{t-1}) = p(\theta)\prod_{t=1}^{T}\frac{1}{\sqrt{2\pi H_{t,k}}}\exp\left(-\frac{y_t^2}{H_{t,k}}\right), \quad (4.31)$$

which is independent of η. Since the conditional distribution of θ does not have a closed-form (because, for example, $p(a_{0k}|Y_t, Z_t, \theta_{-a_{0k}})$, in which $\theta_{-a_{0k}}$ is the parameter vector without a_{0k}, contains $H_{t,k}$, which is also a function of a_{0k}; therefore, it cannot be a normal density) using the Gibbs sampling in this situation may be complicated. The Griddy Gibbs algorithm, that introduced by Ritter and Tanner (1992), can be a solution of this problem. This method is very applicable in researches (for example, Bauwens and Lubrano (1998), Bauwens and Storti (2009), and Bauwens et al. (2010)).

Given samples at iteration r the Griddy Gibbs at iteration $r+1$ proceeds as follows:

1. Select a grid of points, such as $a_{0i}^1, a_{0i}^2, \ldots, a_{0i}^G$. Using (4.31), evaluate the conditional posterior density function $k(a_{0i|Z_t,Y_t,\theta_{-a_{0i}}})$ over the grid points to obtain the vector $G_k = (k_1, \ldots, k_G)$.
2. By a deterministic integration rule using the G points, compute $G_\Phi = (0, \Phi_2, \ldots, \Phi_G)$ with

$$\Phi_j = \int_{a_{0i}^1}^{a_{0i}^j} k(a_{01}|\theta_{-a_{0i}}^{(r)}, Z_t^{(r)}, Y_t)da_{0i}, \quad i = 2, \ldots, G. \quad (4.32)$$

3. Simulate $u \sim U(0, \Phi_G)$ and invert $\Phi(a_{0i}|\theta_{-a_{0i}}^{(r)}, Z_t^{(r)}, Y_t)$ by numerical interpolation to obtain a sample $a_{0i}^{(r+1)}$ from $a_{0i}|\theta_{-a_{0i}}^{(r)}, Z_t^{(r)}, Y_t$.
4. Repeat steps 1–3 for other parameters.

For the prior densities of all elements of θ, it can be can considered independent uniform densities over the finite intervals.

5. Simulation results

In this section we provide some simulation results of MS-CGARCH model defined by Eqs. (2.1)–(2.5) for two states. We simulate 300 sample from the following MS-CGARCH model:

$$y_t = \varepsilon_t\sqrt{H_{Z_t,t}}, \quad (5.33)$$

where $\{\varepsilon_t\}$ is an iid sequence of standard normal variables, $\{Z_t\}$ is a Markov chain on finite state space $E = \{1, 2\}$ with transition probability matrix

$$P = \begin{pmatrix} .85 & .15 \\ .05 & .95 \end{pmatrix}$$

and

$$\begin{aligned} H_{1,t} &= \frac{1-\exp(-2|y_{t-1}|)}{1+\exp(-2|y_{t-1}|)}(2.2 + .75y_{t-1}^2 + .15H_{1,t-1}) \\ &\quad + \left[1 - \frac{1-\exp(-2|y_{t-1}|)}{1+\exp(-2|y_{t-1}|)}\right](.7 + .3y_{t-1}^2 + .2H_{1,t-1}), \\ H_{2,t} &= \frac{1-\exp(-.5|y_{t-1}|)}{1+\exp(-.5|y_{t-1}|)}(.4 + .15y_{t-1}^2 + .1H_{2,t-1}) \end{aligned}$$

Table 1. Descriptive statistics for the simulated data (sample size = 300).

Mean	Std. dev.	Skewness	Maximum	Minimum	Kurtosis
0.034	0.860	0.289	3.109	−2.997	4.502

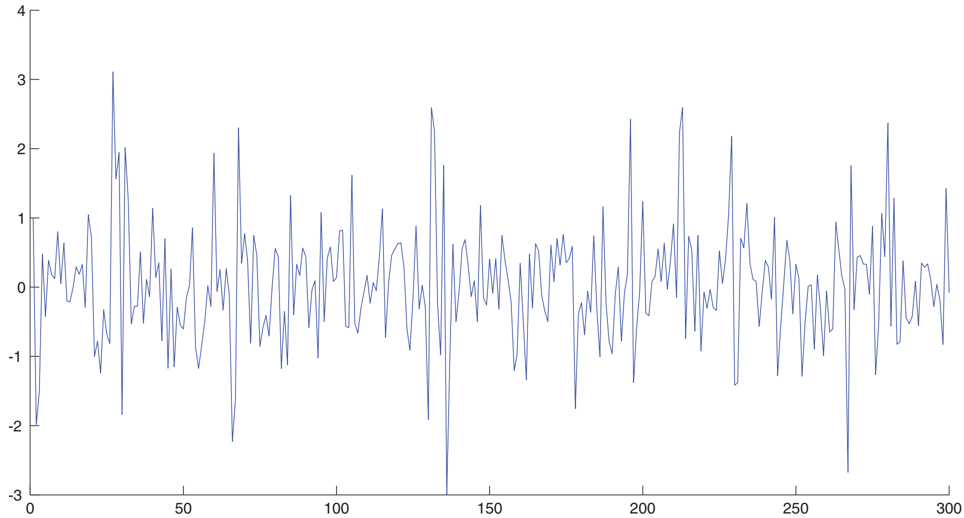

Figure 1. Simulated time series of MS-CGARCH model.

$$+\left[1-\frac{1-\exp(-.5|y_{t-1}|)}{1+\exp(-.5|y_{t-1}|)}\right](.2+.1y_{t-1}^2+.2H_{2,t-1}). \tag{5.34}$$

The first state implies a higher conditional variance than the second one and in each state, the first component has the higher volatility than the other component.

In Table 1, we report summery statistics for simulated data and Fig. 1 shows the plot of the simulated time series.

Using the Bayesian inference, we estimate the parameters of the MS-CGARCH model. The prior density of each parameter is assumed to be uniform restricted over a finite interval (except for η_{11} and η_{22}, since they are drawn from the beta distribution). Table 2 demonstrates

Table 2. Results of the Bayesian estimation of the simulated MS-CGARCH model.

	True values	Mean	Std. dev.
a_{01}	2.200	2.301	0.415
a_{11}	0.750	0.721	0.060
a_{21}	0.150	0.147	0.047
b_{01}	0.700	0.661	0.085
b_{11}	0.300	0.270	0.070
b_{21}	0.200	0.213	0.056
a_{02}	0.400	0.361	0.084
a_{12}	0.150	0.176	0.043
a_{22}	0.100	0.119	0.056
b_{02}	0.200	0.181	0.094
b_{12}	0.100	0.050	0.026
b_{22}	0.200	0.203	0.081
γ_1	2.000	2.01	0.603
γ_2	0.500	0.742	0.150
η_{11}	0.850	0.620	0.086
η_{22}	0.950	0.869	0.042

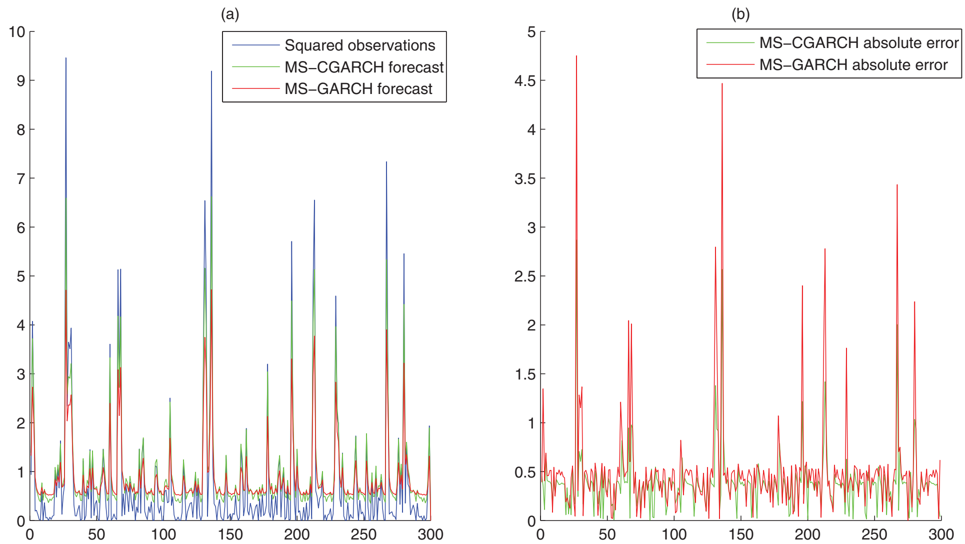

Figure 2. (a) Squared observations of the simulated time series (blue), forecast by MS-GARCH (red), and forecast by MS-CGARCH (green). (b) Absolute forecast error of squared simulated time series in the MS-GARCH (red) and in MS-CGARCH (green).

Table 3. Descriptive statistics of DJIA and $S\&P500$ stock market returns.

	Mean	Std. dev.	Skewness	Maximum	Minimum	Kurtosis
DJIA	0.05	1.01	−0.18	3.82	−3.67	4.76
$S\&P500$	−0.01	1.04	−0.54	2.87	−3.53	4.14

the performance of the estimation methods. The results of this table show that the standard deviation are small enough in most cases.

For clarifying the performance of MS-CGARCH model toward MS-GARCH model, we compare the forecasting volatility ($E(Y_t^2|\mathcal{F}_{t-1})$) of each model with the squared observations. Figure 2 shows that the forecasting volatility of MS-CGARCH is much better than MS-GARCH model and the absolute forecast error (the difference between the forecasting volatility and the squared observations) of our model is often smaller than the MS-GARCH model. The root of mean squared error of the MS-GARCH and MS-CGARCH, respectively, are 0.738 and .483 and the mean absolute error of them are 0.510 and .3804.

6. Empirical applications

We apply the daily stock market index of Dow Jones industrial average (DJIA) from October 7, 2009 to December 14, 2010 (300 observations) and $S\&P500$ from December 12, 2006 to February 22, 2008 (300 observations) for estimation. Figure 3 demonstrates the stock market index and the percentage returns[2] of both DJIA and $S\&P500$. It is evident that the stock market index of DJIA and $S\&P500$ have the divers of shocks. A summary of descriptive statistics of these returns are in Table 3.

In Tables 4 and 5, the posterior means and standard deviations from the estimation of MS-CGARCH and MS-GARCH models for DJIA and $S\&P500$ daily returns are reported.

[2] Percentage returns are defined as $r_t = 100 * \log(\frac{P_t}{P_{t-1}})$, where P_t is the index level at time t.

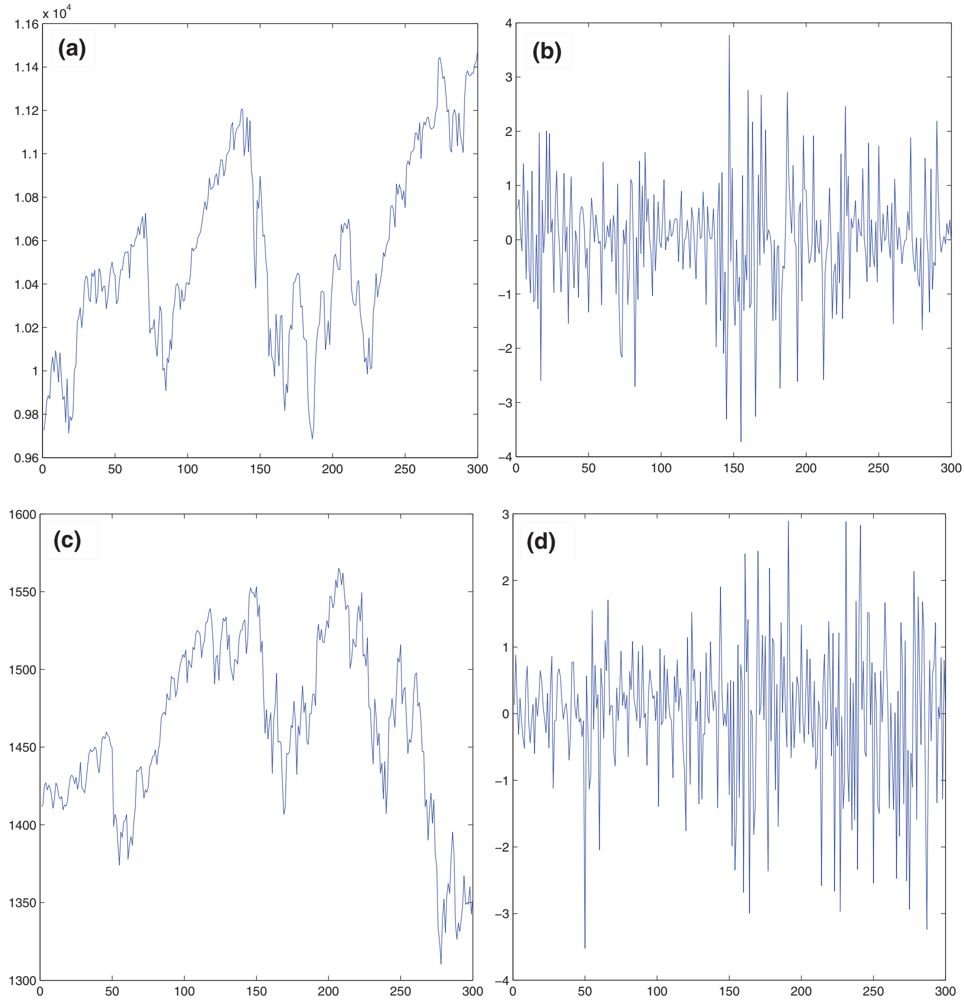

Figure 3. (a) DJIA stock market index, (b) percentage daily returns of DJIA, (c) *S*&*P*500 stock market index, and (d) percentage daily returns of *S*&*P*500.

Table 4. Posterior means and standard deviations (DJIA daily returns).

	MS-CGARCH		MS-GARCH	
	Mean	Std. dev.	Mean	Std. dev.
a_{01}	3.150	0.590	1.859	0.316
a_{11}	0.651	0.145	0.504	0.102
a_{21}	0.094	0.043	0.216	0.054
b_{01}	0.821	0.142	—	—
b_{11}	0.306	0.061	—	—
b_{21}	0.277	0.050	—	—
a_{02}	0.658	0.157	0.498	0.091
a_{12}	0.296	0.054	0.189	0.055
a_{22}	0.203	0.053	0.242	0.071
b_{02}	0.291	0.049	—	—
b_{12}	0.092	0.049	—	—
b_{22}	0.334	0.087	—	—
γ_1	1.554	0.307	—	—
γ_2	0.756	0.136	—	—
η_{11}	0.806	0.140	0.542	0.092
η_{22}	0.941	0.038	0.899	0.031

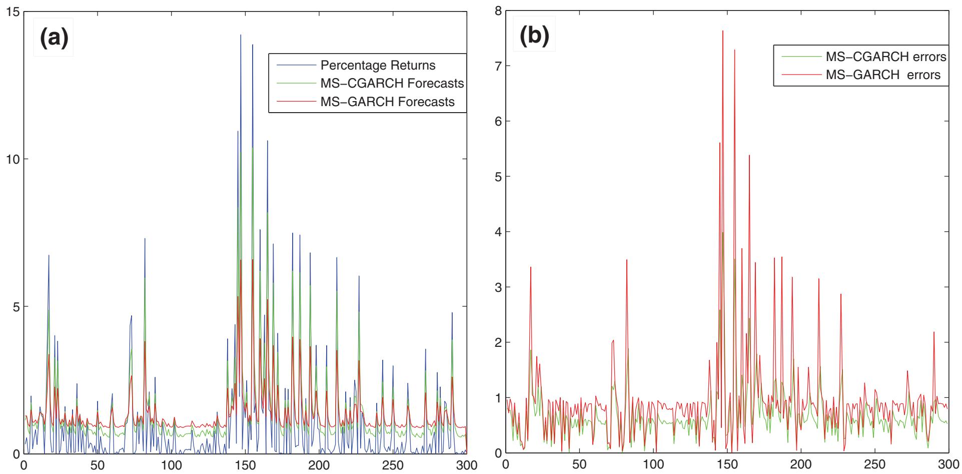

Figure 4. (a) Squared returns of DJIA (blue), forecast by MS-GARCH (red), and forecast by MS-CGARCH (green). (b) Absolute forecast error of squared returns (DJIA) in the MS-GARCH (red) and in the MS-CGARCH (green).

The results of estimating MS-GARCH in both cases (DJIA and $S\&P500$ daily returns) show that the first regime is the high volatility regime. In the high volatility state, the conditional variance is more sensitive to recent shocks ($a_{11} > a_{21}$) and less persistence ($a_{21} < a_{22}$) than the low volatility regime. Also the outcomes of estimating MS-CGARCH (Tables 4 and 5) show that the first regime is the high volatility state that in each state the first component is higher volatile than the second one. The values of γ_1 and γ_2 show the speed of transition (in each regime) from one component to the other one. This specification causes the MS-CGARCH to be more flexible than the MS-GARCH to capture the variants of shocks: very high, high, moderate, and low shocks. Indeed the MS-CGARCH model is able to model the gradual changes in high and low volatile states by the effect of their components in each state.

We compare the forecasting volatility of each model with the squared returns. Figures 4 and 5 show that the forecasting volatility of MS-CGARCH is much better than MS-GARCH model. Table 6 reports the measures of performance forecasting, the mean absolute error, and

Table 5. Posterior means and standard deviations ($S\&P500$ daily returns).

	MS-CGARCH		MS-GARCH	
	Mean	Std. dev.	Mean	Std. dev.
a_{01}	2.016	0.548	1.330	0.329
a_{11}	0.609	0.132	0.454	0.102
a_{21}	0.184	0.050	0.303	0.062
b_{01}	0.767	0.129	—	—
b_{11}	0.286	0.053	—	—
b_{21}	0.352	0.059	—	—
a_{02}	0.622	0.145	0.500	0.092
a_{12}	0.249	0.092	0.162	0.069
a_{22}	0.130	0.056	0.232	0.065
b_{02}	0.313	0.058	—	—
b_{12}	0.086	0.049	—	—
b_{22}	0.315	0.077	—	—
γ_1	1.856	0.505	—	—
γ_2	0.725	0.126	—	—
η_{11}	0.774	0.086	0.821	0.065
η_{22}	0.915	0.027	0.936	0.022

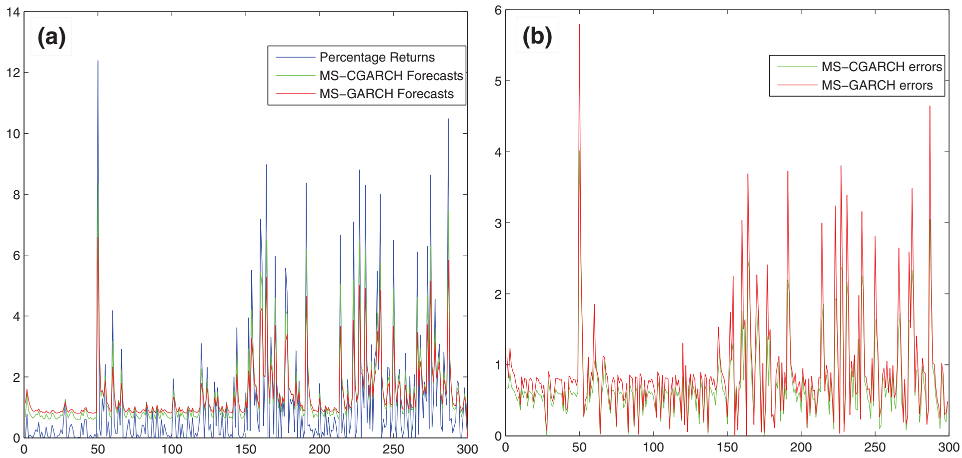

Figure 5. (a) Squared returns of *S&P*500 (blue), forecast by MS-GARCH (red), and forecast by MS-CGARCH (green). (b) Absolute forecast error of squared returns (*S&P*500) in the MS-GARCH (red) and in the MS-CGARCH (green).

Table 6. Measures of performance forecasting

	DJIA		*S&P*500	
	MS-GARCH	MS-CGARCH	MS-GARCH	MS-CGARCH
RMSE	1.281	0.834	1.169	0.902
MAE	0.940	0.687	0.904	0.723

root of mean squared error, for both MS-CGARCH and MS-GARCH models. Based on the results given in Table 6, the MS-CGARCH model has a much better forecast than MS-GARCH model.

7. Conclusion

In this paper a generalization of the MS-GARCH model is presented where the conditional variance in each state is a time varying convex combination of two GARCH components. Such model provides a better evaluation of volatility models which is in effect of time varying volatilities. By imposing one of the component with higher volatility than the other in each state, the MS-CGARCH provides a dynamic structure to model the effect of such time varying volatilities in each regime by reacting to the various species of shocks. Our model can provide more better forecast of volatility in compare to MS-GARCH, when the time series is in effect of wide range of volatilities. The Bayesian method based on Gibbs algorithm is applied which performs very well. We also presented a simple necessary and sufficient condition for the existence of an upper bound for the second moment.

This work has the potential to be applied in the context of financial time series. The empirical distribution of daily returns does not generally have a Gaussian distribution. They have fat tails densities (they are called leptokurtic). One of the extension of this work is considering the fat tail densities instead of Gaussian distribution, that can cause better modeling of the financial time series. Also we can generalize this model by allowing an ARMA structure for the conditional mean.

Acknowledgment

We would like to thank the anonymous reviewer for the valuable comments and suggestions which helped to improve the quality of this paper.

Funding

The second author research was in part supported by a grant from IPM (No. 91620035).

References

Abramson, A., Cohen, I. (2007). On the stationarity of Markov-Switching GARCH processes. *Econometric Theory.* 23:485–500.

Alexander, C., Lazar, E. (2008). Markov-Switching GARCH diffusion. *ICMA Center Discussion Paper in Finance 2008-01*, University of Reading.

Alizadeh, S.H., Rezakhah, S. (2013). Hidden Markov mixture autoregressive models:stability and moments. *Commun. Stat. Theory Methods.* 42(6): 1087–1104.

Baumeister, C., Durinck, E., Peerman, G. (2008). Liquidity, inflation and asset prices in a time-varying framework for the euro area. Available at: http://ideas.repec.org/p/nbb/reswpp/200810-17.html. Research Series 200810-17, National Bank of Belgium.

Bauwens, L., Lubrano, M. (1998). Bayesian inference on GARCH models using the Gibbs sampler. *Econom. J.* 1:23–46.

Bauwens, L., Storti, G. (2009). A component GARCH model with time varying weights. *Stud. Nonlinear Dyn. Econom.* 13(2): 1–33.

Bauwens, L., Preminger, A., Rombouts, V.K. (2010). Theory and inference for Markov switching GARCH model. *Econom. J.* 13:218–244.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *J. Econom.* 31:307–327.

Booth, G.C., Ciner, C. (2001). The relationship between nominal interest rates and inflation: International evidence. *J. Multinational Financial Manage.* 11(6): 269–280. (Intervention).

Cai, J. (1994). A Markov model of switching-regime ARCH. *J. Bus. Econ. Stat.* 12:309–316.

Chib, S. (1996). Calculating posterior distributions and model estimates in Markov mixture models. *J. Econom.* 75:79–97.

Ding, Z., Granger, C.W.J. (1996). Modeling volatility persistence of speculative returns: a new approach. *J. Econom.* 173:185–215.

Engle, R.F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica.* 50:987–1007.

Engle, R.F., Rangel, J. (2008). The Spline-GARCH model for low frequency volatility and its global macroeconomic causes. *Rev. Financial Stud.* 21:1187–1222.

Gelfand, A., Smith, A. (1990). Sampling based approaches to calculating marginal densities. *J. Am. Stat. Assos.* 85:398–409.

Gonzalez-Rivera, G. (1998). Smooth transition GARCH models. *Stud. Nonlinear Dyn. Econom.* 3:61–78.

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

Grimmett, G., Stirzaker, D. (2001). *Probability and Random Processes.* New York: Oxford University press.

Haas, M., Mittnik, S., Paolella, M.S. (2004). A new approach to Markov-switching GARCH models. *J. Financial Econom.* 2:493–530.

Hamilton, J.D., Susmel, R. (1994). Autoregressive conditional heteroskedasticity and changes in regime. *J. Econom.* 64:307–333.

Hagerud, G. (1997). A smooth transition ARCH model for asset returns, SSE/EFI, Working Paper Series in Economic and Finance, No. 162.

Hyde, S., Bredin, D. (2005). Regime Changes in the Relationship between Stock Returns and the Macroeconomy. http://dx.doi.org/10.2139/ssrn.686878.

Kaufman, S., Fruhwirth-Schnatter, S. (2002). Bayesian analysis of switching ARCH models. *J. Time Ser. Anal.* 23:425–458.

Klaassen, F. (2002). Improving GARCH volatility forecasts with regime-switching GARCH. *Empirical Econ.* 27:363–394.

Lancaster, P., Tismenetsky, M. (1985). *The Theory of Matrices,* 2nd ed. San Diego, California: Academic press.

Lubrano, M. (2001). Smooth transition GARCH models: a Bayesian approach. *Recherches Economiques de Louvain.* 67:257–287.

Maheu, J. (2005). Can GARCH models capture long range dependence?*Stud. Nonlinear Dyn. Econom.* 9(4): Article 1.

Medeiros, M.C., Veiga, A. (2009). Modeling multiple regimes in financial volatility with a flexible coefficient GARCH(1,1). *J. Econom. Theory.* 25:117–161.

Robert, C., Casella, G. (2004). *Monte Carlo Statistical Methods*. New York: Springer.