An asymmetric Markov switching GARCH model: estimation and forecasting

N. Alemohammad[1], S.Rezakhah[2] and S. Hosseinalizadeh [3] [2]

[1]**Department of Mathematics and Computer Science, Shahed University, Tehran, 18151/159 , Iran. Email: n.alemohammad@shahed.ac.ir.**

[2]**Faculty of Mathematics and Computer Science, Amirkabir University of Technology, Tehran, Iran. Email: rezakhah@aut.ac.ir.**

[3]**Department of Computer Engineering and Information Technology, Qazvin Branch, Islamic Azad University, Qazvin, Iran. Email: sasan.h.alizadeh@qiau.ac.ir.**

Abstract: In this paper an extension of Markov switching GARCH model is proposed to model leverage effect, conditional heteroscedasticity and nonlinearity attributes of financial time series. Greedy Gibbs Bayesian learning method is used to estimate the parameters of the model. Due to the complexity of the model a dynamic programming algorithm for forecasting is proposed. Finally we illustrate the model on S&P500 daily returns.

Keywords: GARCH models, Markov switching, Forecasting, Bayesian inference, Griddy Gibbs sampling.

Mathematics Subject Classification (2010): 60J10, 62M10, 62F15.

1 Introduction

In the past four decades, dynamic financial time series analysis based on nonlinear models has become a topic of interest for some researchers. For financial time series, the ARCH model and GARCH model are surely the most popular classes of volatility models. Merging (G)ARCH model

[2]**N. Alemohammad :** `n.alemohammad@shahed.ac.ir`

with a hidden Markov chain, where each state of the chain allows a different (G)ARCH behavior, was introduced by Hamilton. Gray (1996), Haas et al. (2012) proposed different variants of Markov-Switching GARCH (MS-GARCH) model and its applications. Also some other researchers such as Bauwens et al. (2010), Billio et al. (2014) have done extensive studies on the properties of this model.

One restitution of the GARCH model is the symmetry in the response of volatility to past shocks (because the volatility is a function of the lagged squared observations). In other words, the dynamics of conditional variance in (G)ARCH models change only with the size of squared observations. Financial markets become more volatile in response to negative shocks than to positive shocks . The asymmetric GARCH model (AGARCH) introduced by Engle (1990), Threshold GARCH (TGARCH) model of Zakoian (1994) and Skew normal mixture and Markov-switching GARCH Haas (2010) capture the asymmetric effects of negative and positive shocks.

In this paper, we propose a new structure for the Markov switching GARCH model where the volatility of each regime has different reactions to positive and negative shocks. This model is talented to switch between regimes with distinct levels of volatilities and also encompasses the leverage effect in each regime. By considering differen smooth transition functions for each state of Markov chain (to capture asymmetry property in each state), our model is able to have more accurate forecast than the MS-GARCH. Since 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. For the estimation of the parameters, the Bayesian inference is used.

The organization of this paper is as follows: in the next section the structure of our model is presented. Section 3 is devoted to the estimation of the parameters of the model. Section 4 is dedicated to the analysis of the efficiency of the proposed model via an empirical application and the comparison of forecast errors with the MS-GARCH model. Section 5 concludes.

2 Markov switching smooth transition GARCH model

The Markov switching smooth transition GARCH model, MS-STGARCH, 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 $E = \{1, 2, \cdots, K\}$ with transition probability matrix $P = ||p_{ij}||_{K\times K}$, where $p_{ij} = p(Z_t = j|Z_{t-1} = i),\ i, j \in \{1, \cdots, K\}$, and stationary probability measure $\pi = (\pi_1, \cdots, \pi_K)'$. Also given that $Z_t = j$, $H_{t,j}$ (the conditional variance in regime j) is driven by

$$H_{t,j} = a_{0,j} + a_{1,j}y_{t-1}^2(1 - w_{j,t-1}) + a_{2,j}y_{t-1}^2 w_{j,t-1} + b_j H_{t-1,j}, \tag{2.2}$$

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

$$w_{j,t-1} = \frac{1}{1 + \exp(-\gamma_j y_{t-1})} \quad \gamma_j > 0,\ j = 1, \cdots, K, \tag{2.3}$$

which are bounded, $0 < w_{j,t-1} < 1$, and monotonically increasing. The parameter γ_j is called the slope parameter, that explains the speed of transition from negative shocks to positive one: the higher γ_j, the faster the transition. Since $\gamma_j > 0$, the weight function $0 < w_{j,t-1} < 1$ goes to zero when $y_{t-1} \to -\infty$ and to one when $y_{t-1} \to +\infty$. As y_{t-1} increases from negative values to positive values the impact of y_{t-1}^2 proceeds proportionally from $a_{1,j}$ to $a_{2,j}$. So $a_{1,j}$ in each state will characterize negative shocks and $a_{2,j}$ positive ones. It refers to the fact that negative and positive shocks have different effects on volatility. When γ_j tending to zero, $w_{j,t-1}$ goes to 1/2 and the MS-STGARCH model tends to MS-GARCH model.

2.1 Forecasting

The conditional variance of MS-STGARCH model is given by

$$\begin{aligned} Var(Y_t|\mathcal{I}_{t-1}) = \sum_{i=1}^{K} \alpha_i^{(t)} H_{t,i} = \sum_{i=1}^{K} \alpha_i^{(t)} (a_{0,i} + a_{1,i}y_{t-1}^2(1 - w_{i,t-1}) \\ + a_{2,i}y_{t-1}^2 w_{i,t-1} + b_i H_{t-1,i}) \end{aligned} \tag{2.4}$$

as $H_{t,i}$ is the conditional variance of i-th 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.

Remark 2.1. *The value of* $\alpha_j^{(t)}$ *is obtained recursively by*

$$\alpha_j^{(t)} = \frac{\sum_{m=1}^{K} f(y_{t-1}|Z_{t-1} = m, \mathcal{I}_{t-2})p(Z_{t-1} = m|\mathcal{I}_{t-2})p_{m,j}}{\sum_{m=1}^{K} f(y_{t-1}|Z_{t-1} = m, \mathcal{I}_{t-2})p(Z_{t-1} = m|\mathcal{I}_{t-2})}. \tag{2.5}$$

3 Estimation

For the estimation of parameters, we apply the Bayesian inference, that is extensively used in literature see Bauwens et al. (2010).

Let Y_t be the vector $(y_1, \cdots, y_t)$. In what follows, for the case of two regimes, we are going to estimate the $Z_t = (z_1, \cdots, z_t)$ and the parameter vectors $\eta = (\eta_{11}, \eta_{22})$ and $\theta = (\theta_1, \theta_2)$, where $\theta_k = (a_{0k}, a_{1k}, a_{2k}, b_k, \gamma_k)$ for $k = 1, 2$ from the posterior density

$$p(\theta, \eta, Z|Y) \propto p(\theta, \eta)p(Z|\theta, \eta)f(Y|\theta, \eta, Z), \tag{3.1}$$

in which $Y = (y_1, \cdots, y_T)$, $Z = (z_1, \cdots, z_T)$ and $p(\theta, \eta)$ is the prior of the parameters.

3.1 Sampling z_t

This step is devoted to the sampling of $p(z_t|\eta, \theta, Y_t)$ that is performed by Chib. Suppose $p(z_1|\eta, \theta, Y_0,)$ be the stationary distribution of the chain,

$$p(z_t|\eta, \theta, Y_t) \propto f(y_t|\theta, z_t = k, Y_{t-1})p(z_t|\eta, \theta, Y_{t-1}), \tag{3.2}$$

that by the law of total probability $p(z_t|\eta, \theta, Y_{t-1})$ is given by:

$$p(z_t|\eta, \theta, Y_{t-1}) = \sum_{z_{t-1}=1}^{K} p(z_{t-1}|\eta, \theta, Y_{t-1})\eta_{z_{t-1}z_t}. \tag{3.3}$$

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 $t = T-1, \cdots, 0$ the sample is derived from $p(z_t|z_{t+1}, \cdots, z_T, \theta, \eta, Y)$,which is obtained by

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

To derive z_t from $p(z_t|.) = p_{z_t}$ is by sampling from the conditional probabilities $q_j = p(Z_t = j|Z_t \geq j, .)$ which are given by

$$p(Z_t = j|Z_t \geq j, .) = \frac{p_j}{\sum_{l=j}^{K} q_l}.$$

After generating a uniform (0,1) number U, if $U \leq q_j$ then $z_t = j$, otherwise increase j to $j+1$ and generate another uniform (0,1) and compare it by q_{j+1}.

3.2 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}.

3.3 Sampling θ

To sample from the $p(\theta|Y, Z, \eta)$ we use the Griddy Gibbs algorithm that introduced by Ritter and Tanner. This method is very applicable in researches (see **?**). Given samples at iteration r the Griddy Gibbs at iteration $r+1$ proceeds as follows:

Algorithm 1. *This algorithm has the following steps:*

- *Step 1 Select a grid of points, such as* $a_{0i}^1, a_{0i}^2, \cdots, a_{0i}^G$. *Using* (**??**), *evaluate the conditional posterior density function* $k(a_{0i|Z_t,Y_t,\theta_{-a_{0i}}})$ *over the grid points to obtain the vector* $G_k = (k_1, \cdots, k_G)$.

- *Step 2 By a deterministic integration rule using the G points, compute* $G_\Phi = (0, \Phi_2, \cdots, \Phi_G)$ *with*

$$\Phi_j = \int_{a_{0i}^1}^{a_{0i}^j} k(a_{01}|\theta_{-a_{0i}}^{(r)}, Z_t^{(r)}, Y_t) da_{0i}, \quad i = 2, \cdots, G. \tag{3.4}$$

- *Step 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* $a_{0i}|\theta_{-a_{0i}}^{(r)}, Z_t^{(r)}, Y_t$.

- *Step 4 Repeat steps 1-3 for other parameters.*

In this section we apply the daily stock market index of $S\&P500$ from 04/08/2010 to 24/04/2012 (450 observations) for estimation. Figure 1 demonstrates the stock market index, the percentage returns [1] and the kernel density of the data set. Using the Bayesian inference, we estimate the parameters of the MS-STGARCH (2.1)-(2.3) and MS-GARCH models (that each model has two regimes). 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). Parameters γ_1

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

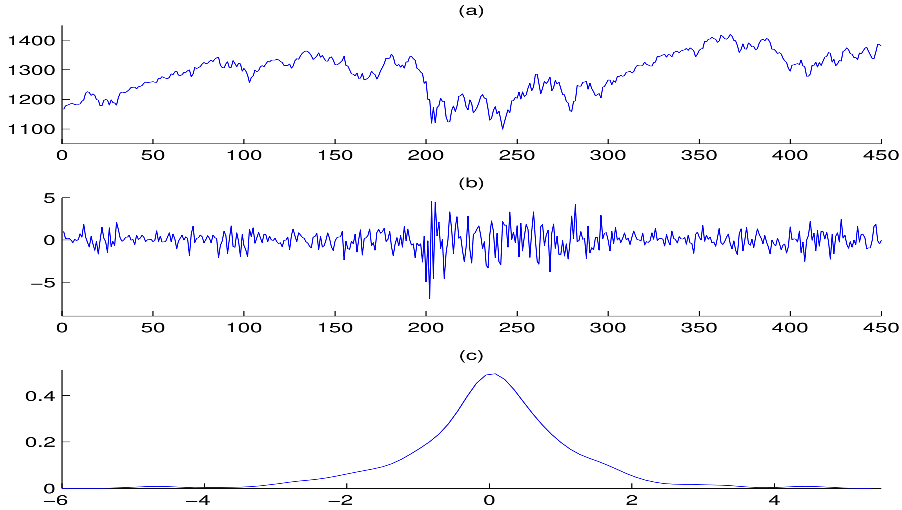

Figure 1: (a): $S\&P500$ from 04/08/2010 to 24/04/2012 stock market index, (b): Percentage daily returns , (c): kernel density of the returns.

and γ_2 in the MS-GARCH model are zero. The estimation results for both models are reported in Table 2 and show that the standard deviations are small enough in most cases. The outcomes of estimating display that the MS-STGARCH has different reactions to positive and negative shocks in both regimes while the MS-GARCH has the similar responses to positive and negative shocks. Since in the MS-STGARCH, $a_{11} > a_{21}$ and according to structure of $w_{j,t-1}$ (2.3) we deduce that the negative shocks have more effect on volatility than the positive shocks in the two regimes. Figure 2 pictures the logistic transition functions for both regimes in MS-STGARCH.

For clarifying the efficiency of the MS-STGARCH model toward the MS-GARCH model, we compare the forecasting volatility ($E(Y_t^2|\mathcal{F}_{t-1})$) of each model with the squared returns .Also for evaluation of forecasting accuracy of our model toward the MS-GARCH model we compute some measures of performance forecasting [2] that are reported in Table 3. The results of this table display that all performance measures of our model are smaller than that of MS-GARCH model.

[2]If e_t is the forecast error (the difference between the actual value and the forecasted value), some of the performance measures are as follows:

- Mean absolute error (MAE)$=\frac{1}{n}\sum_{t=1}^{n}|e_t|$,
- Mean forecast error (MFE)$=\frac{1}{n}\sum_{t=1}^{n}e_t$,
- Mean squared error (MSE)$=\frac{1}{n}\sum_{t=1}^{n}e_t^2$.

Table 1: Posterior means and standard deviations ($S\&P500$ daily returns).

	MS-STGARCH		MS-GARCH	
	Mean	Std.dev.	Mean	Std.dev
a_{01}	0.200	0.007	0.199	0.007
a_{11}	0.288	0.008	0.288	0.007
a_{21}	0.103	0.004	0.103	0.005
b_1	0.290	0.006	0.300	.007
γ_1	2.157	0.395	0.000	0.000
a_{02}	0.585	0.054	0.584	0.055
a_{12}	0.673	0.051	0.664	0.046
a_{22}	0.3463	0.039	0.362	0.048
b_2	0.310	0.041	0.311	0.042
γ_2	1.049	0.265	0.000	0.000
η_{11}	0.948	0.029	0.942	0.044
η_{22}	0.9674	0.030	0.968	0.026

Table 2: Forecast performance measures

Model	Mean Square Error	Mean Absolute Error	Mean Forecast Error
MS-GARCH	3.211	0.924	-0.178
MS-STGARCH	1.930	0.811	-0.254

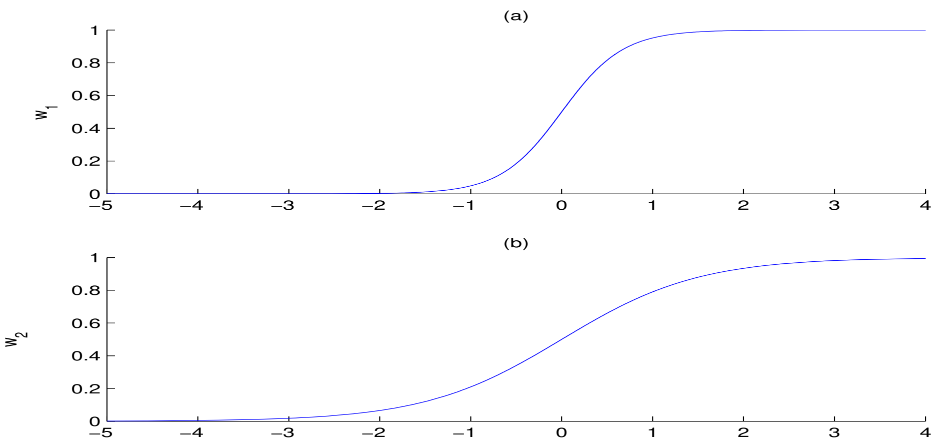

Figure 2: Logistic smooth transition functions for the MS-STGARCH model, (a): first regime (w_1), (b): second regime (w_2).

Conclusion

In this paper an extension of the MS-GARCH model is introduced where the volatility in each state captures the leverage effect property of financial time series. For the estimation of parameters the Bayesian inference is used by applying the Griddy Gibbs algorithm. Through an empirical example, we show that our model can provide better forecast of volatility than the MS-GARCH.

This model has the potential to be applied for modeling and forecasting conditional volatility of financial time series . For the sake of simplicity it was assumed that the process conditional mean is zero, this assumption could be relaxed by refining the structure of model to allow ARMA structure for conditional mean.

References

Bauwens, L. Preminger, A. Rombouts, V.K. (2010). Theory and inferencefor Markov switching GARCH model, *Econometrics journal*, **13**, 218–244.

Billio, L., Casarin, R. Osuntuyi, A. (2014). Efficient Gibbs sampling for Markov switching GARCH models , *Computational Statistics and Data Analysis*, In Press.

Engle, R. F.(1990). Discusion: stock volatility and the crash of '87, *Review of Financial Studies*, **3**, 103-106.

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

Haas, M. (2010). Skew normal mixture and Markov-switching GARCH processes, *Studies in Nonlinear Dynamic & Econometrics*, **14**, 1–54.

Haas, M. Paolella, M.S.(2012). Mixture and Regime-Switching GARCH Models, *Handbook of Volatility Models and Their Applications, First Edition*, John Wiley & Sons, Inc.

Zakoian, J. M. (1994). Threshold heteroskedastic models, *Journal of Economic Dynamic and Control*, **27**, 577–597.