DYNAMIC ECONOMETRIC MODELS
Vol. 8 – Nicolaus Copernicus University – Toruń – 2008

Monika Kośko
The University of Computer Science and Economics in Olsztyn
Michał Pietrzak[*]
Nicolaus Copernicus University

Modeling Financial Time Series Volatility with Markov Switching Models

1. Introduction

An analysis of financial time series volatility is an important issue in making many economic decisions. The volatility of high frequency financial series changes over time and the periods of the high volatility are clustering. Many authors use GARCH models, introduced by Bollerslev (1986), to capture these dependences. GARCH models describe the conditional variance clustering effect but their forecasts are often overstated (Anderson and Bollerslev, 1998). An application of the Markov switching specification to GARCH models can outperform forecasts of the standard GARCH structure. The first Markov switching model was used by Hamilton (1989) in the analysis of the business cycle. The ARCH model with Markov switching (SWARCH) was the first specification in this class of models (Hamilton, Susmel, 1994). Next the SWARCH structure was extended to GARCH parameters, giving the MS-GARCH model. The Markov switching GARCH model was characterized by Davidson (1994), Klassen (2002) and Gray (1996), and each of them defined an equation of the conditional variance in a different way. The conditional variance equation is then exploited in the estimation of MS-GARCH model parameters.

The main purpose of this article is to check whether, a better quality volatility predictions can be obtained from MS-AR-GARCH than from AR-GARCH models. At first the estimation of those types of models has been carried out for

[*] In the case of Michał Pietrzak the publication was funded by the Nicolaus Copernicus University within the confines of the 404 E grant.

the investigated series. Next, having well-estimated models, one-day predictions have been pointed out for the future thirty sessions. The following step of the research was calculating the ex post prediction error, in the form of RMSE, in order to compare the prediction properties of the both analyzed types of models.

2. MS-AR-GARCH Model

First of all, AR-GARCH and MS-AR-GARCH[1] models differ between themselves in the conception of a volatility clustering explanation. The volatility in AR-GARCH models is described by a consideration of previous volatility levels (t-1). Therefore, this kind of a specification characterizes the volatility clustering effect quite well, when periods of a low variance follow long periods of a high variance. In case of MS-AR-GARCH modeling the volatility clustering is discussed as a result of staying in a one state for some time and then violent switching to another state. Because of the r states occurrence, thereare r equations of the conditional variance, so it is called a mixture of distributions. This property allows to characterize a volatility clustering in financial time series as well as AR-GARCH models. However the Markov switching structure should provide a better forecast abilities in comparison to AR-GARCH model.

The Markov switching MS-AR-GARCH model is given by:

$$y_t \mid Y_{t-1}, s_t, \theta^{(i)} \sim \begin{cases} y_t^1 & dla \quad s_t = 1 \\ \vdots & \\ y_t^r & dla \quad s_t = r \end{cases}, \tag{1}$$

$$y_t^i = \mu_t^i + \varepsilon_t^i, \tag{2}$$

$$\varepsilon_t^i = u_t \sqrt{h_t^i}\,,\ u_t \sim IID(0,1), \tag{3}$$

where

y_t -the empirical value of the process in the t-th moment,

Y_{t-1} - process information from the past to the (t-1)-th moment,

$\theta^{(i)} = \left[p_{ii}, \beta_0^{(i)}, \beta_1^{(i)}, \gamma_1^{(i)}, df^{(i)}\right]$- the estimated parameters vector in i state,

$s_t = i$ -the state of the process in the t-th moment, $i \in \{1,2,...,r\}$.

The conditional mean can be described by the autoregressive AR(p) process in the t-th moment and for i-th state. AR(p) process can be written as:

$$\mu_t^i = \alpha_0 (s_i = i) + \alpha_1 (s_t = i) y_{t-1} + ... + \alpha_p (s_t = i) y_{t-p}, \tag{4}$$

[1] The GARCH and the MS-GARCH models were expanded, respectively to AR-ARCH and MS-AR-GARCH specifications with autoregressive process AR(p), witch enable to describe an autocorrelation in time series.

and the conditional variance in the t-th moment for i-th state is given by:

$$h_t^i = \beta_0 (s_t = i) + \sum_{l=1}^{p} \beta_l (s_t = i)\varepsilon_{t-l}^2 + \sum_{k}^{q} \gamma_k (s_t = i) h_{t-k} \,. \tag{5}$$

3. MS-AR-GARCH Model Estimation

An estimation of the MS-AR-GARCH model based on Equation (1) is a computationally quite difficult. When you need to calculate the conditional variance h_t^i, for two states and GARCH(1,1) process, you have to consider two equations of the conditional variance h_{t-1}^i depending on $s_{t-1} = i$. Then for the each of variance equations h_{t-1}^i you have to take into account two equations of conditional variance h_{t-2}^i for the sake of the two states $s_{t-2} = i$. This pattern is done until the $t = 1$ moment. So it is seen that the number of needed states increases with a number of time series observations. Therefore the estimation of this kind of equation becomes unworkable. There are three approaches to solve this problem in the literature.

The first approach was introduced by Davidson (2004), where the GARCH structure is an autoregressive process with infinite number of lags ARCH(∞):

$$h_t^i = \frac{\beta_0 (s_t = i)}{1 - \beta_1 (s_t = i) - \ldots - \beta_q (s_t = i)} + \sum_{k=1}^{\infty} \delta_k (s_t = i)\varepsilon_{t-1}^2 \,. \tag{6}$$

An advantage of this solution is that the conditional variance h_i^t depends only on state s_t-th.

The second approach is a proposition of Gray (1996), where lag of conditional variance is the expected value of conditional variances for each state. Hence in the case of GARCH model the conditional variance is given by:

$$h_t^i = \beta_0 (s_t = i) + \beta_1 (s_t = i)\varepsilon_{t-1}^2 + E(h_{t-1}) \,, \tag{7}$$

$$E(h_{t-1}) = P(s_{t-1} = 1 | t-2) h_{t-1}^1 + P(s_{t-1} = 2 | t-2) h_{t-1}^2 \,, \tag{8}$$

where $P(s_t = i | t-1)$ is a probability that the process in *t*-th moment is in *i*-th state and the information about the process until (*t*-1)-th moment is known.

The third approach was proposed by Klassen (2002), where the values of the conditional variance h_{t-1}^1, h_{t-1}^2 are needed. These values in Klassen's approach are calculated from equation:

$$h_{t-1}^i = \beta_0 (s_t = i) + \beta_1 (s_t = i)\varepsilon_{t-2}^2 + E(h_{t-2}) \,, \tag{9}$$

where an expected value $E(h_{t-2})$ is calculated by an analogy from Equation (8).

In this paper the Davidson solution was used to define an equation of conditional variance. The parameters of the research model were obtained by the maximization of the log-likelihood function,which can be written as[2]:

$$l=\sum_{t=1}^{T}\ln\left[P(s_t=i|t-1)\cdot f(y_t \mid Y_{t-1},s_t=1)+(1-P(s_t=i|t-1))\cdot f(y_t \mid Y_{t-1},s_t=2)\right], \quad (10)$$

$$P(s_t=i|t-1)=P(s_{t-1}=1|t-1)p_{1i}+P(s_{t-1}=2|t-1)p_{2i}, \quad (11)$$

$$P(s_{t-1}=i|t-1)=\frac{f(y_{t-1} \mid Y_{t-2},s_{t-1}=i)\cdot P(s_{t-1}=i|t-2)}{f(y_{t-1} \mid Y_{t-2},s_{t-1}=1)\cdot P(s_{t-1}=1|t-2)+f(y_{t-1} \mid Y_{t-2},s_{t-1}=2)\cdot P(s_{t-1}=2|t-2)}, \quad (12)$$

where

$f(y_t \mid Y_{t-1}, s_t = i)$ - the density function of a distribution,

$P(s_t = i|t)$- probability of the process in *t*-th moment being in *i*-th state and information about the process until *t*-th moment is known,

p_{ij} - the conditional probability of the process switching from the *i*-th state to the *j*-th state (the transition probability) .

The return rate forecasts and the conditional variance forecasts for *k* periods were calculated from equations given by:

$$y_{T+k}^{P}=\sum_{i=1}^{2}P(s_{T+k}=i|T)\cdot \mu_{T+k}^{(i)}, \quad (13)$$

$$h_{T+k}^{P}=\sum_{i=1}^{2}P(s_{T+k}=i|T)\cdot h_{T+k}^{(i)}, \quad (14)$$

where the conditional probabilities $P(s_{T+k} = i|T)$ are received recursively from Equation (11).

4. The Financial Time Series Results

The empirical analysis refer to the daily return rates of the companies quoted on the Warsaw Stock Exchange, which create WIG20 index[3]. The values

[2] The log-likelihood function in (10) was constructed for two states but it can be done for any number of states.

[3] To estimate the AR-GARCH and the MS-AR-GARCH models, 12 from 20 companies of the WIG20 index were chosen. It has been the consequence of the assumption, that at least 1000 observations of a time series should be taken into consideration. The analyzed series come from the period from November 17, 2000 (the date of implementing WARSET system) to March 30, 2007. The logarithmic rates of return have been multiplied by 100.

of distribution's characteristics and some tests have been presented in Table 1. These results verify a time series properties for the selected stock market returns (Agora S.A., Telekomunikacja Polska S.A., KGHM Polska Miedź S.A., PKN Orlen S.A.[4]). All series have increased kurtosis (more than 3) and the hypothesis of the normal distribution in accordance with Jarque-Bera's test results is rejected. The ARCH effect occurs in every time series, what is shown by the Ljung-Box test for the squares of return rates. According to Ljung-Box test in the case of Agora S.A. and KGHM S.A., the autocorrelation phenomenon appears.

Table 1. Distribution characteristics of rates of return for the chosen series

Distribution's characteristics	AGORA	TP	KGHM	PKN Orlen
Standard daviation	2.1748	2.1457	2.54483	1.8936
Assymetry	0.0739	0.1762	-0.2252	0.1448
Kurtosis	4.8796	4.0895	4.925	4.189
Jarque-Bera test	236.83 [0.000]	87.36 [0.000]	260.57 [0.000]	99.83 [0.000]
Ljung-Box test (autocorelation) - $Q(20)$	33.24 [0.032]	19.69 [0.477]	37.11 [0.011]	20.77 [0.410]
Ljung-Box test (effekt ARCH) - $Q(20)$	191.30 [0.000]	330.54 [0.000]	140.86 [0.000]	53.62 [0.000]

Source: Calculations in TSM programme, p-values have been presented in brackets.

The results[5] of the estimation for the AR(p)-GARCH(p,q) models are presented in Table 2. The estimated parameters are statistically significant for all time series. While analysing the residuals it is worth to pay an attention on the increased kurtosis in residual processes and on the fact, that according to Jarque-Bera's test for null hipothesis, normality of residual distribution is rejected. This is the result of the conditional t-Student distribution of residuals, which with low level degree of freedom have higher curtosis in comparison to the normal distribution. The ARCH effect and the autocorrelation phenomenon have been successfully eliminated in the case of the all series. The Schwarz information criterion and RMSE values are presented in the last two rows of Table 2. The RMSE values were calculated on the basis of the obtained predictions.

[4] Because of the limited spare, only the results of the chosen 4 series from the acquired results for 12 companies have been presented.

[5] Both the AR-GARCH and the MS-AR-GARCH models were estimated with the assumption of the normal distribution or of the t-Student distribution of residuals.

Table 2. Results of he AR(r)-GARCH(p,q) estimation

Parameters	AGORA	TPSA	KGHM	PKNOrlen
α_0	-	-	0.13295 [0.023]	-
α_1	0.09884 [0.000]	-	0.04839 [0.054]	-
β_0	1.15782	1.03842	1.49264	1.3602
β_1	0.08046 [0.007]	0.05786 [0.000]	0.04725 [0.000]	0.03674 [0.001]
γ_1	0.89019 [0.000]	0.92608 [0.000]	0.92901 [0.000]	0.92491 [0.000]
df	-	11.90	7.45	10.05
Assymetry (residual)	0.0097	0.2795	-0.3142	0.1881
Kurtosis (residual)	4.277	3.6316	4.333	3.9781
Jarque-Bera test (residual)	108.59 [0.000]	47.40 [0.000]	144.68 [0.000]	73.16 [0.000]
Ljunga-Box test (autocorelati-on-residual)– Q(20)	13.205 [0.868]	12.9198 [0.881]	28.63 [0.095]	17.46 [0.623]
Ljung-Box test (effect Arch-residual)-Q(20)	17.1686 [0.642]	15.3737 [0.755]	15.59 [0.742]	10.12 [0.966]
LL	-3425.78	-3404.88	-3680.77	-3261.35
SC	-3440.54	-3419.64	-3702.9	-3276.11
RMSE error	10.33	5.94	12.04	3.82

Source: Calculations in TSM programme, p-values have been presented in brackets.

Table 3 presents results of the MS-AR-GARCH model estimation where the parameters are statistically significant. The obtained residuals have similar characteristics to the AR-GARCH models. In the case of the all companies the ARCH effect has been eliminated. The Schwarz information criterion and RMSE values are presented in the last two rows of Table 3.

5. Conclusions

Analyzing[6] the fitting of the AR-GARCH and MS-AR-GARCH models one can easily notice the comparable Schwarz information criterion values. This means that both models have been fitted similarly to empirical data. Next, comparing the values of the RMSE for the examined time series one can ascertain that the errors for both types of models are also comparable. It is testified by the fact that on the basis of 12 analyzed companies included in the WIG20 index, it cannot be ascertained that the MS-AR-GARCH model acquires better volatility prediction properties than the AR-GARCH model.

[6] The summary presents the results gained from all 12 time series chosen from 20 companies of the WIG20 index.

Table 3. Results of the MS-AR(r)-GARCH(p,q) switching models estimation

Parameters		AGORA		TPSA		KGHM		PKN Orlen	
α_0		-		-		0.1404 [0.016]		-	
α_1		0.08592 [0.001]		-		0.04817 [0.055]		-	
p_{11} / p_{12}		0.8969	0.1031	0.9930	0.0070	0.9908	0.0092	0.9836	0.0164
p_{22} / p_{21}		0.0419	0.9581	0.0070	0.9930	0.0070	0.9930	0.0085	0.9915
MS-GARCH ~t-Student	$\sqrt{\beta_0}^{(1)}$	2.74468		1.75942		1.40453		1.56296	
	$\sqrt{\beta_0}^{(2)}$	1.01255		1.0982		2.04273		1.69574	
	$\beta_1^{(1)}$	0.00839 [0.05]		0.04358 [0.011]		0.03507 [0.008]		0.02836 [0.368]	
	$\gamma_1^{(1)}$	0.97836 [0.000]		0.9270 [0.000]		0.93034 [0.000]		0.91998 [0.000]	
	$df^{(1)}$	-		15.05		7.84		11.56	
	$df^{(2)}$	-		-		-		12.53	
Assymetry (residual)		0.0358		0.2341		-0.2858		0.1955	
Kurtosis (residual)		2.7618		3.3364		4.1235		3.8872	
Jarque-Bera test (residual)		4.1205 [0.127]		22.12 [0.000]		105.79 [0.000]		62.6367 [0.000]	
Ljunga-Box test (autocorelation-residual)– Q(20)		11.61 [0.929]		12.6715 [0.891]		28.2988 [0.103]		17.00 [0.653]	
Ljung-Box test (effect Arch-residual)-Q(20)		25.45 [0.185]		15.34 [0.756]		17.052 [0.65]		10.46 [0.959]	
LL		-3400.79		-3399.93		-3675.92		-3260.54	
SC		-3426.61		-3425.75		-3709.12		-3290.05	
RMSE error		9.28		5.83		11.55		3.82	

Source: Calculations in TSM programme, p-values have been presented in brackets.

It may be concluded that in the situation of similar models' quality with respect to the description of empirical data, as well as to the volatility predictions of financial time series, it would be worth to choose the AR-GARCH model as less complicated model. What is also important is the usage of the MS-AR-GARCH model enables gaining additional information on the transition mechanism and dynamics of the process in the each state. Since all models are highly stable in regimes, the average time to each regime and the average time of process duration can be fixed, what additionally increases the prediction properties of Markov model.

References

Anderson, T. G., Bollerslev, T. (1998), Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts, *International Economic Review*, vol. 39.

Bollerslev, T. (1986), Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*, vol. 31.

Cai, J. (1994), A Markov Model of Unconditional Variance in ARCH, *Journal of Business and Economic Statistics*, vol. 12.

Davidson, J. (1994), Forecasting Markov-Switching Dynamic, Conditionally Heteroscedastic Proceses, *Statistics and Probability Letters*

Dempster, A. P., Laird, N. M., Rubin, D. B. (1977), Maximum Likelihood from Incomplete Data via the EM Algorithm, *Journal of the Royal Statistical Society*, vol. 39.

Doman, R. (2004), Forecasting the Polish Financial Market Volatility with Markov Switching Models, *Macromodels'2003*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź 2004.

Doman, M., Doman, R. (2004), *Econometric Modeling of the Polish Financial Market Dynamic*, Wydawnictwo Akademii Ekonomicznej, Poznań.

Gray, S. (1996), Modeling the Conditional Distribution of Interest Rates As a Regime-Switching Process, *Journal of Financial Economics*, vol. 42.

Hamilton, J.D. (1989), A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle, *Econometrics* ,vol. 57.

Hamilton, J. D., Susmel, R. (1994), Autoregressive Conditional Heteroskedasticity and Changes in Regime, *Journal of Econometrics*, vol 64.

Kim, C.-J. (1994), Dynamic Linear Models with Markov-Switching, *Journal of Econometrics*, vol. 60.

Klassen, F. (2002), Improving GARCH Volatility Forecasts Empirical, *Economics*, vol. 27.

Marcucci, J. (2003), Forecasting Stock Market Volatility with Regime-Switching GARCH Models, *Studies in Nonlinear Dynamics and Econometrics*, vol. 9.