ORIGINAL PAPER

Application and Diagnostic Checking of Univariate and Multivariate GARCH Models in Serbian Financial Market

Jelena Minović, Belgrade Banking Academy

KEY WORDS: Multivariate GARCH models, Ljung-Box statistics, **R**esidual-based diagnostics, Lagrange Multiplier test

UDC: 519.868 **JEL**: C30, G10

ABSTRACT *- The goal of this article is to give theoretical and empirical review for diagnostic checking of multivariate volatility processes. In theoretical part we presented three categories diagnostics for conditional heteroscedasticity models: portmanteau tests of the Ljung-Box type, residual-based diagnostics (RB) and Lagrange Multiplier (LM) tests. In our empirical analysis we used the Ljung-Box statistics (Q-test) of standardized residuals, those of its squared, as well as of the cross product of standardized residuals to check the model adequacy. Our results showed that the residual-based diagnostics provide a useful check for model adequacy. Overall result is that models perform statistically well.*

Introduction

As empirical researchers are equipped with various conditional heteroscedasticity models, checking the adequacy of a fitted model becomes an important issue for model selection. Generally, misspecification in the mean and variance results in inconsistency and loss of efficiency in the estimated parameters (Tse, 2002, 358). Since estimating Multivariate GARCH models (MGARCH) is time consuming, both in terms of computations and their programming (if needed), it is desirable to check *ex ante* whether the data present evidence of multivariate ARCH effects. *Ex post*, it is also of crucial importance to check the adequacy of the MGARCH specification (Bauwens et. al., 2006, 79). MGARCH models specify equations for how the covariance move over time (Rombouts et. al., 2004). There are also tests specifically designed for multivariate models that are applied to the vectors of residuals together. Popular univariate procedures that are also relevant for multivariate models include asymmetry tests and residual-based conditional moment tests (Mills et. al., 2006). As mentioned by Tse (2002), diagnostics for conditional heteroscedasticity models applied in the literature can be divided into three categories: portmanteau tests of the Box-Pierce-Ljung type, residual-based diagnostics (RB) and Lagrange Multiplier (LM) tests (Bauwens et. al., 2006, 79). The Box–Pierce–Ljung portmanteau statistic is perhaps the most widely used diagnostic (Tse, 2002, 358).

We considered trivariate time series models for some selected securities listed at the Belgrade stock exchange (www.belex.co.yu). Prior to that, we had to perform univariate GARCH analysis for each of the analyzed return series. Then, we checked the fitted models carefully: the Ljung-Box statistics of standardized residuals and its squared values showed that models are adequate for describing the conditional heteroscedasticity of the data. Multivariate (trivariate) GARCH models which will be covered in this paper are restricted version of BEKK (named after Baba, Engle, Kraft and Kroner, 1995), model, the diagonal VEC (DVEC, initially due to Bollerslev, Engle and Wooldridge, 1988) model and Conditional Correlation Model (CCC, Bollerslev, 1990)). For trivariate version of restricted BEKK, DVEC and CCC representations we estimated covariances among daily log returns of BELEX15 index, Hemofarm and Energoprojekt stocks. For estimation of parameters in the univaraite and trivariate GARCH models we used EViews program, Version 4.1. The methods for estimation parameters which we used are maximum log-likehood and two-step approach. Particularly, we tested how the covariances between chosen securities move over time in trivariate case. Overall, our results show that the residual-based diagnostics provide a useful check for model adequacy.

The rest of the paper is organized as follows. In Section 2, we present different diagnostics for univariate GARCH models which are used in our analysis. Then, we report results of diagnostic checking for univariate GARCH models. In Section 3 misspecification tests for multivariate GARCH models are described. Then, we report results of diagnostic checking for trivariate GARCH models. Finally, Section 4 concludes.

Univariate case

Diagnostic checking of univariate GARCH models (theoretical part)

Here we briefly reviewed of different test-statistics and diagnostics for univariate GARCH models which are used in this paper and in references (Minović, 2007).

Lagrange multiplier (LM) Test: This test is used to investigate whether the standardized residuals exhibit ARCH behaviour. If the variance equation of ARCH model is correctly specified, there should be no ARCH effect left in the standardized residuals (EViews 5 User's Guide). It is calculated as $LM = TR^2$, where R^2 is defined above. Reject H_0 if $LM > \chi^2_*(m)$, where m is order of ARCH effect (Greene, 2003).

Q Test (Ljung and Box, 1978): It is used to test for the presence of autocorrelation of order m in residuals and it is calculated as

$$Q(m) = T(T+2)\sum_{l=1}^{m}\frac{\hat{\rho}_l^2}{T-l}. \tag{2.1}$$

The function $\hat{\rho}_1$, $\hat{\rho}_2$, ... is called the sample autocorrelation function (SACF) of r_t (Tsay, 2005). Q test has a χ^2-distribution with m degrees of freedom under the null hypothesis (Vogelvang, 2005). The Ljung–Box statistics of the residuals can be used to check the adequacy of a fitted model (Tsay, 2005).

Q^2 Test: It is used to test for the presence of autocorrelation of order m in squared residuals. Then, it is used to test for remaining ARCH effect in the variance equation and to check the specification of the variance equation (EViews 5 User's Guide).

The *F*-statistic reported in the regression output is used to test the null hypothesis that *all* of the slope coefficients (excluding the constant, or intercept) in a regression are zero. For ordinary least squares models, the *F*-statistic is computed as:

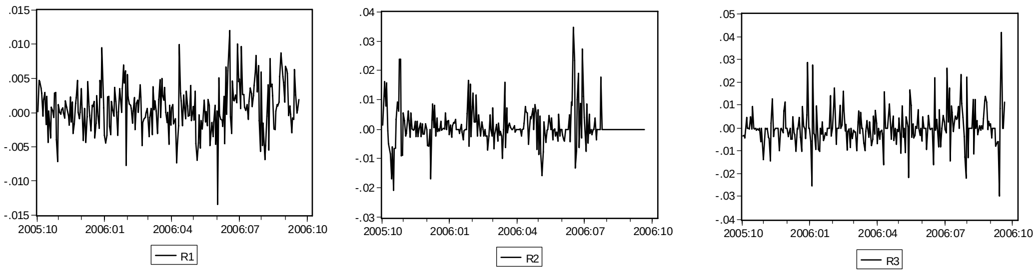
$$F = \frac{R^2/(k-1)}{(1-R^2)/(T-k)}. \tag{2.2}$$

Under the null hypothesis with normally distributed errors, this statistic has an *F*-distribution with (k - 1) and (T - k) degrees of freedom. The total number of regressors is termed as k; number of restrictions is termed as k – 1; T is number of observations (Brooks, 2002), (EViews 5 User's Guide). The p-value given just below the *F*-statistic, denoted Prob(*F*-statistic), is the marginal significance level of the *F*-test (EViews 5 User's Guide).

Diagnostic checking of univariate GARCH models (empirical part)

In this part we consider univariate case and we use data of daily log returns for BELEX15 index, Hemofarm stock and Energoprojekt stock, respectively. Our data cover the period from October 3, 2005 to October 6, 2006 (Minović, 2007). We applied log-difference transformation to convert data into continuously compounded returns, because the price series (log values) of both stocks and index is not stationary and are stationary when they are first differenced. Let r_{1t}, r_{2t}, and r_{3t} be the log return series (Figure 1) corrected for autocorrelation in the mean of BELEX15 index, Hemofarm and Energoprojekt stocks, respectively [8].

Figure 1: The graphs of daily log returns of BELEX15 index (r_1), Hemofarm (r_2) and Energoprojekt (r_3) stocks, respectively.



We observe from Figure 1 that the log returns of BELEX15 index, Hemofarm and Energoprojekt stocks evidence the well known the volatility clustering effect. It is tendency for volatility in financial markets to appear in bunches. Thus large returns (of either sign) are expected to follow large returns, and small returns (of either sign) to follow small returns (Brooks, 2002), (Minović, 2007).

We use four steps for building a volatility model for each of the log return series. The first step is to specify a mean equation by testing for serial dependence in the data and building an ARMA model for the log return series to remove any linear dependence. Then, in the second step, we use the residuals of the mean equation to test for ARCH effects. Next, in the third step, we specify a volatility model if ARCH effects are statistically significant and perform a joint estimation of the mean and volatility equations. Finally, in the fourth step we check the fitted model carefully (Tsay, 2005).

In univariate case, obviously, the residuals have to be tested for the presence of autocorrelation. In time-series terminology this is called 'diagnostic checking'. With the Ljung-Box (Q) test, we tested whether the residuals behave like a white noise process (Vogelvang, 2005). Table 1 reports the Q(m) and Q²(m) statistics for each series. We applied the Lagrange multiplier (LM) test on our series in order to investigate whether the standardized residuals exhibit ARCH behaviour (EViews 5 User's Guide), (Bauwens et. al., 2006, 79), (Minović, 2007).

Table 1: The Ljung-Box statistics of standardized residuals and squared standardized residuals in ARMA models and test for ARCH effect.

	The Ljung-Box Statistics								ARCH-LM(5) test	
series	**Q(2)**	**Q(5)**	**Q(9)**	**Q(36)**	**Q²(2)**	**Q²(5)**	**Q²(9)**	**Q²(36)**	**F-stat**	**Obs*R^2**
BELEX15	0.655 (0.721)	2.780 (0.734)	4.290 (0.891)	26.559 (0.874)	2.726 (0.256)	4.121 (0.532)	6.496 (0.689)	26.178 (0.885)	0.726 (0.604)	3.666 (0.598)
Hemofarm	1.453 (0.484)	7.483 (0.187)	11.491 (0.244)	34.857 (0.523)	33.866 (0.000)	47.446 (0.000)	72.915 (0.000)	86.635 (0.000)	6.613 (0.000)	29.775 (0.000)
Energoprojekt	1.848 (0.397)	2.209 (0.820)	3.406 (0.946)	26.659 (0.871)	10.437 (0.005)	13.065 (0.023)	13.852 (0.128)	26.716 (0.870)	3.138 (0.009)	15.099 (0.010)

We see from table above that is the significant Q-statistics for squared residuals across all lag lengths for Hemofarm stock and we infer the presence of ARCH effects. The Q-statistics for squared residuals across all lag lengths for BELEX15 index is not significant, it ignores the existence of ARCH effects. But the heteroscedasticity in BELEX15 and Hemofarm is also observed in the plots of the actual values of residuals (see Minović, 2007). From Table 2.1 we see that is only one significant autocorrelation on lag 2 in squared residuals in ARMA model for Energoprojekt stock. It is evident that exist ARCH effect for Energoprojekt stock. On the other hand, the Lagrange multiplier (LM) test (Table 2.1) shows strong ARCH effects for Hemofarm stock with test statistic F = 6.613, the *p*-value of which is zero; and ARCH effect for Energoprojekt stock with test statistic F = 3.138, the *p*-value 0.009. Then this test shows no ARCH effect for BELEX15 index with test statistic F = 0.726 and the *p*-value 0.604 (Minović, 2007).

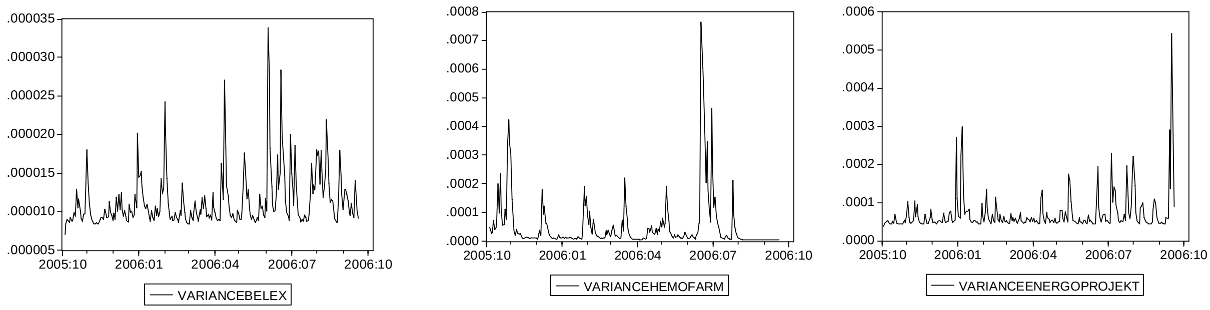
We built a volatility model for each asset returns and we inferred that right model for BELEX15 index is the ARMA(1,1)-GARCH(1,1), then for Hemofarm stock is the ARMA(2,2)-IGARCH(1,1) and for Energoprojekt stock ARMA(0,0)-GARCH(1,1) model (for detail see Minović, 2007). However, Hemofarm stock follows Integreted GARCH model (IGARCH). In order to examine IGARCH process, we applyed Wald test. Finally, we checked the fitted models carefully. The Ljung-Box statistics (Table 2.2) of standardized residuals and those of its squared showed that models are adequate for describing the conditional heteroscedasticity of the data . We applied the ARCH test on the standardized residuals to see if there ware any ARCH effects left. Both the F-statistic and the LM-statistic are very insignificant, suggesting no ARCH effect up to order 5 or 10 for each series BELEX15 index, Hemofarm and Energoprojekt stocks (see Table 2) (Minović, 2007).

Table 2: The Ljung-Box statistics and ARCH-LM test of order 5 and 10.

series	*The Ljung-Box Statistics* Q(36)	$Q^2(36)$	*ARCH-LM(5) test* F-stat	Obs*R^2	*ARCH-LM(10) test* F-stat	Obs*R^2
BELEX15	25.590 (0.901)	26.319 (0.881)	0.265 (0.932)	1.351 (0.930)	0.423 (0.934)	4.356 (0.930)
Hemofarm	29.038 (0.788)	23.685 (0.943)	0.084 (0.995)	0.432 (0.994)	0.146 (0.999)	1.518 (0.999)
Energoprojekt	26.447 (0.878)	22.156 (0.966)	0.799 (0.551)	4.028 (0.545)	0.606 (0.808)	6.183 (0.800)

On Figure 2 we plot the GARCH variances for BELEX15 index, Hemofarm and Energoprojekt stock. We see on this figure that all variances are unstable over time.

Figure 2: The GARCH variance series for BELEX15 index, Hemofarm and Energoprojekt stocks.



We found that correlation coefficients between log returns of BELEX15 index and Hemofarm stock is 0.49; between log returns of BELEX15 index and Energoprojekt stock is 0.40; and between log returns of Hemofarm and Energoprojekt stocks about 0.02 and we conclude that these two stocks are practically noncorrelated (Minović, 2007).

In addition to visual inspection Figure 2.2 tell us that GARCH variance series exhibit significant changes over time for both stocks and index. Therefore, these variances are very unstable over time. A plot of GARCH variances of BELEX15 index, Hemofarm and Energoprojekt stocks reveals that BELEX15 index has been more volatile than Hemofarm and Energoprojekt stocks. On graph for variance of Hemofarm stock, we observe the greatest peak in period June-July 2006, it was when company Schtada was bought stocks of Hemofarm and Schtada was became of major. We see significant autocorrelation on this graph which occur because Hemofarm in univariate case follow IGARCH process. On graph for variance of Energoprojekt, we see that the first peak was in February 2006, when Energoprojekt company signed contract in Nigeria valued 151 million euros (Minović, 2007).

Multivariate case

Theoretical review of misspecification tests for multivariate GARCH models

Asymmetry tests

Engle and Ng (1993) have proposed a set of tests for asymmetry in volatility, known as sign and size bias tests. The Engle and Ng tests should thus be used to determine whether an asymmetric model is required for a given series, or whether a symmetric GARCH model can be deemed adequate. In practice the Engle and Ng tests are usually applied to the residuals of a univariate GARCH fit to individual series. Denote an individual series of disturbances as ε_t, and define S_{t-1}^{-} as an indicator dummy that takes the value 1 if $\hat{\varepsilon}_t < 0$ and zero otherwise. Then the test for sign bias is based on the significance or otherwise of ϕ_1 in the regression

$$\hat{\varepsilon}_t^2 = \phi_0 + \phi_1 S_{t-1}^{-} + e_t \tag{3.1}$$

where e_t is an i.i.d. (independently and identically distributed) error term. If positive and negative shocks to $\hat{\varepsilon}_{t-1}$ impact differently upon the conditional variance, then ϕ_1 will be statistically significant (Mills et. al., 2006).

It could also be the case that the magnitude or size of the shock will affect whether the response of volatility to shocks is symmetric or not. In this case, a negative sign bias test would be conducted, based on a regression in which S_{t-1}^{-} is now used as a slope dummy variable. Negative sign bias is argued to be present if ϕ_1 is statistically significant in the regression

$$\hat{\varepsilon}_t^2 = \phi_0 + \phi_1 S_{t-1}^{-} \varepsilon_{t-1} + e_t\,. \tag{3.2}$$

Finally, defining $S_{t-1}^{+} = 1 - S_{t-1}^{-}$, so that S_{t-1}^{+} picks out the observations with positive innovations, Engle and Ng propose a joint test for sign and size bias based on the regression

$$\hat{\varepsilon}_t^2 = \phi_0 + \phi_1 S_{t-1}^{-} + \phi_2 S_{t-1}^{-} \varepsilon_{t-1} + \phi_3 S_{t-1}^{+} \varepsilon_{t-1} + e_t\,. \tag{3.3}$$

Significance of ϕ_1 indicates the presence of sign bias, where positive and negative shocks have differing impacts upon future volatility, compared with the symmetric response required by the standard GARCH formulation. On the other hand, the significance of ϕ_2 or ϕ_3 would suggest the presence of size bias, where not only the sign, but also the magnitude, of the shock is important. A joint test statistic is formulated in the standard fashion by calculating TR^2 from regression

(3.3), which will asymptotically follows a χ^2 distribution with 3 degrees of freedom under the null hypothesis of no asymmetric effects (Mills et. al., 2006).

Residual-based conditional moment tests

If the model is correctly specified, and represents an adequate characterization of the data, certain moment relationship should hold on an appropriately standardized measure of the residuals. Let θ denote a $k \times 1$ parameter vector (containing all model parameters) and $r(\theta)$ denote the restrictions function required for the test. The null hypothesis is:

$$H_0 : r(\theta) = 0 . \tag{3.4}$$

Let $\Omega = \text{var}(r(\hat{\theta}))$, (3.5)

where $\hat{\theta}$ is vector of estimated parameters.

Then the Wald test statistic is given by

$$W = r(\hat{\theta})^T \Omega^{-1} r(\hat{\theta}) . \tag{3.6}$$

Under the null, W^{α} : χ_J^2, and so the null hypothesis is rejected if $W > \chi_{J,\alpha}^2$, where α is the size of the test and J is the number of restrictions. The variance-covariance in (3.5), $\hat{\Omega}$, may be calculated from the residual sum of squares of the regression of $\hat{m}$ (the values of elements of the moment restriction function by observation) on $\hat{d}_1, \hat{d}_2, ..., \hat{d}_k$, the values of each of the derivatives of the log-likelihood, observation by observation. The residual sum of squares of this regression is given by

$$R = \hat{m}^T \hat{m} - \hat{m}^T D \left(D^T D \right)^{-1} D^T \hat{m} . \tag{3.7}$$

The required variance is then

$$\hat{\Omega} = \frac{R}{T^2} \text{ (Mills et. al., 2006).} \tag{3.8}$$

Multivariate model diagnostics

Among the specific multivariate model diagnostics, Bauwens (2003) propose the use of a multivariate version of the **Ljung-Box test** due to Hosking (1980) (Mills et. al., 2006). This test is the most widely used diagnostics to detect ARCH effects (Bauwens et. al., 2006, 79). Let $\hat{z}_t = \hat{H}_t^{-1/2} \hat{\varepsilon}_t$ denote the *N*-vector of standardized residuals, the test statistic is given by

$$HM(M) = T^2 \sum_{t=1}^{M} (T-j)^{-1} tr \left\{ C_{Z_t}^{-1}(0) C_{Z_t}(j) C_{Z_t}^{-1}(0) C_{Z_t}^{'}(j) \right\}, \tag{3.9}$$

where $Z_t = VECH\left(\hat{z}_t \hat{z}_t^{'}\right)$ and $C_{Z_t}(j)$ is the sample autocovariance matrix of order *j* given by:

$$C_{Z_t}(j) = T^{-1} \sum_{t=j+1}^{n} \left(Z_t - \overline{Z}\right)\left(Z_{t-j} - \overline{Z}\right)^{'}, \quad j = 0, ..., T-1, \tag{3.10}$$

with $\overline{Z} = \left(Z_1 + ... + Z_T\right) / T$. Under the null hypothesis of no dependence in the standardized residuals (i.e. no ARCH effects), the test statistic is asymptotically distributed as a χ^2 with $N^2 M$ degrees of freedom (Bauwens et. al., 2006, 79), (Mills et. al., 2006).

The Box-Pierce portmanteau statistics have been used as the benchmark for detecting model inadequacy in multivariate conditional heteroscedasticity models. This test is based on cross-products of the standardized residuals often provides a useful diagnostic. Denoting

$\hat{\varepsilon}_t = \left(\hat{\varepsilon}_{t1}, ..., \hat{\varepsilon}_{tk}\right)'$ and the elements of $\hat{\Sigma}_t$ by $\hat{\sigma}_{ij,t}$, we define the *i*-th standardized residuals at time *t* as

$$z_{it} = \hat{\varepsilon}_{it} / \sqrt{\hat{\sigma}_{iit}} \,. \tag{3.11}$$

Let $\hat{\rho}_{ij,t}$ be the estimated conditional correlation coefficient defined by $\hat{\rho}_{ij,t} = \hat{\sigma}_{ij,t} / \sqrt{\hat{\sigma}_{ii,t}\hat{\sigma}_{jj,t}}$; we consider $c_{ij,t}$ defined by

$$c_{ij,t} = \begin{cases} z_{it}^2 - 1 & i = j \\ z_{it}z_{jt} - \hat{\rho}_{ij,t} & i \neq j \end{cases} \tag{3.12}$$

for $i, j = 1, ..., k$. When the constant-correlation or the no-correlation models are estimated, $\hat{\rho}_{ij,t}$ is a constant with respect to *t*. Under correct model specification, $c_{ij,t}$ is asymptotically serially uncorrelated and $E\left(c_{ij,t} \mid \Phi_{t-1}\right) \to 0$ as $n \to \infty$. Thus, a diagnostic can be constructed based on the Box-Pierce statistic of the squared lag autocorrelation coefficient of $c_{ij,t}$. Specifically, we denote r_{hij} as the lag-*h* autocorrelation coefficient of $c_{ij,t}$ and define

$$Q(i, j; m) = n\sum_{h=1}^{m} r_{hij}^2 \,. \tag{3.13}$$

If the multivariate conditional heteroscedasticity model fits the data, $c_{ij,t}$ should be serially uncorrelated for *i* and *j*. An excessive value of *Q* would suggest model inadequacy. The test has been widely used in the empirical literature for diagnosing both univariate and multivariate conditional heteroscedasticity models (Tse et. al., 1999, 679).

Ling and Li (1997) develop another diagnostic test for unparametrized heteroskedasticity in the standardized residuals (Mills et. al., 2006). The Ling-Li test represents a rigorous approach to conducting tests for multivariate conditional heteroscedasticity, providing the justification for the asymptotic null distribution (Tse et. al., 1999, 679). Their formulation resembles a multivariate version of the Durbin-Watson test applied to the squares of the standardized residuals. It is defined as:

$$LL(M) = T\sum_{h=1}^{M} \tilde{R}(h) \tag{3.14}$$

which asymptotically follows a $\chi^2(M)$ under the null of no conditional heteroscedasticity, and where (Bauwens et. al., 2006, 79), (Mills et. al., 2006)

$$\tilde{R}(h) = \frac{\sum_{t=h+1}^{T}\left(\hat{\varepsilon}_t'\hat{\Sigma}_t^{-1}\hat{\varepsilon}_t - N\right)\left(\hat{\varepsilon}_{t-h}'\hat{\Sigma}_{t-h}^{-1}\hat{\varepsilon}_{t-h} - N\right)}{\sum_{t=h+1}^{T}\left(\hat{\varepsilon}_t'\hat{\Sigma}_t^{-1}\hat{\varepsilon}_t - N\right)^2} \,. \tag{3.15}$$

In particular, imposing the restriction that the variance matrix Σ_t is diagonal, Ling and Li reported that the LL(M) statistic has power against situations in which the orders of the conditional variance equations are misspecified (Tse et. al., 1999, 679).

In the derivation of the asymptotic results, conditional normality of the innovation process is not assumed. The statistic is thus robust with regard to the conditional distribution choice. Tse and Tsui (1999) show that there is a loss of information in the transformation of the residuals $\hat{\varepsilon}_t'\hat{\Sigma}_t^{-1}\hat{\varepsilon}_t$ and the test may suffer from a power reduction (Bauwens et. al., 2006, 79). In other words, the

Ling-Li test may not have good power if the misspecification occurs only in the conditional covariance term (Tse et. al., 1999, 679). Furthermore, Duchesne and Lalancette (2003) argue that if an inappropriate choice of *M* is selected, the resulting test statistic may be quite inefficient. For this reasons, these authors propose a more powerful version of the LL(M) test based on the spectral density of the stochastic process $\left\{\hat{\varepsilon}_t^{'}\hat{\Sigma}_t^{-1}\hat{\varepsilon}_t, t \in Z\right\}$ which is i.i.d. under null of homoscedasticity. Interestingly, since their test is based on a spectral density estimator, a data-dependent choice of *M* is available (Bauwens et. al., 2006, 79). Ling and Li (1997) further developed this work and derived the asymptotic distribution of the portmanteau statistic in the multivariate case. The Ling–Li statistic is based on the serial correlation coefficients of the transformed vector of residuals (Tse, 2002, 358).

Lagrange Multiplier (LM) tests are also very widespread in the GARCH literature. Generally, they have an advantage over Ljung-Box and Ling and Li tests due to their efficiency when the alternative is correct (although they can be asymptotically equivalent in certain cases). Bollerslev, Engle, and Wooldridge (1998) and Engle and Kroner (1995), among others, have developed LM tests for MGARCH models (Bauwens et. al., 2006, 79).

To reduce the number of parameters in the estimation of MGARCH models, it is a common practice to introduce restrictions. For instance, the CCC model of Bollerslev (1990) assumes that the conditional correlation matrix is constant over time. It is then desirable to test this assumption afterwards. Tse (2000) proposes a test for constant correlations. The null is $\sigma_{ij,t} = \rho_{ij}\sqrt{\sigma_{ii,t}\sigma_{jj,t}}$ where the conditional variances are GARCH(1,1), while the alternative is $\sigma_{ij,t} = \rho_{ij,t}\sqrt{\sigma_{ii,t}\sigma_{jj,t}}$. The test statistic is a LM statistic which under the null is asymptotically $\chi^2\left(N\left(N-1\right)/2\right)$ (Bauwens et. al., 2006, 79).

Engle and Sheppard (2001) propose an alternative procedure to test the constant correlation hypothesis, in the spirit if the DCC models (for detail about these models see Minović, 2007). The null $H_0 : \rho_t = \overline{\rho} \quad \forall t = 1,...,T$ is tested against the alternative $H_1 : VECH\left(\rho_t\right) = VECH\left(\overline{\rho}\right) + \beta_1^* VECH\left(\rho_{t-1}\right) + ... + \beta_p^* VECH\left(\rho_{t-p}\right)$. The test is easy to implement since H_0 implies the nullity of all coefficients in the regression $X_t = \beta_0^* + \beta_1^* X_{t-1} + ... + \beta_p^* X_{t-p} + u_t^*$, where $X_t = VECH^u\left(\hat{z}_t\hat{z}_t^{'} - I_N\right)$, $VECH^u$ is like the VECH operator but it only selects the elements under the main diagonal, $\hat{z}_t = \hat{\overline{\rho}}^{-1/2}\hat{D}_t^{-1}\hat{\varepsilon}_t$ is the $N \times 1$ vector of standardized residuals (under the null), and $D_t = diag\left(\sqrt{\sigma_{11,t}},...,\sqrt{\sigma_{NN,t}}\right)$ (Bauwens et. al., 2006, 79).

The residual-based *F* test (RBF(*i,j*)) described by Pagan and Hall (1983) consists of running a regression of the cross-products of the standardized residuals on some 'information variables' and examining the statistical significance of these variables. Bollerslev (1990) incorporated the lagged values of the cross-products of the standardized residuals. The diagnostic is then based on the F statistic for the joint significance of the two regressors (Tse et. al., 1999, 679).

Y. K. Tse and Albert K. C. Tsui (1999) considered several tests for model misspecification after a multivariate conditional heteroscedasticity model has been fitted. They examined the performance of the recent test due to Ling and Li, the Box-Pierce test, and the residual-based *F* test using Monte Carlo methods. They found that there were situations in which the Ling-Li test had very weak power. The residual-based diagnostics demonstrated significant under-rejection under the null. In contrast, the Box-Pierce test based on the cross-products of the standardized residuals often provided a useful diagnostic that has reliable empirical size as well as good power against the alternatives considered (DeGroot et.al., 2003).

Empirical analysis of trivariate GARCH models

For estimation of parameters in the univaraite and trivariate GARCH models we used EViews program, Version 4.1. We use program for modeling restricted version of trivariate BEKK model (named after Baba, Engle, Kraft and Kroner), and we extend this program on trivariate case of DVEC (diagonal vector ARCH model) and CCC (Constant Conditional Correlation Model) models (Minović, 2007). A first simple method to estimate the parameters of a trivariate GARCH models is the Berndt-Hall-Hall-Hausman (BHHH) algorithm. This algorithm uses the first derivatives of the quasi-maximum likelihood (QML) with respect to the number of parameters that are contained in multivariate GARCH models. This is an iterative procedure, the BHHH algorithm needs suitable initial parameters (Franke et.al., 2005). For all calculations in our programs number of iteration is 100 and convergence criterion is $1 \cdot 10^{-5}$ which suggests about high precision (Minović, 2007).

Modeling of restricted BEKK, DVEC and CCC models in trivariate version

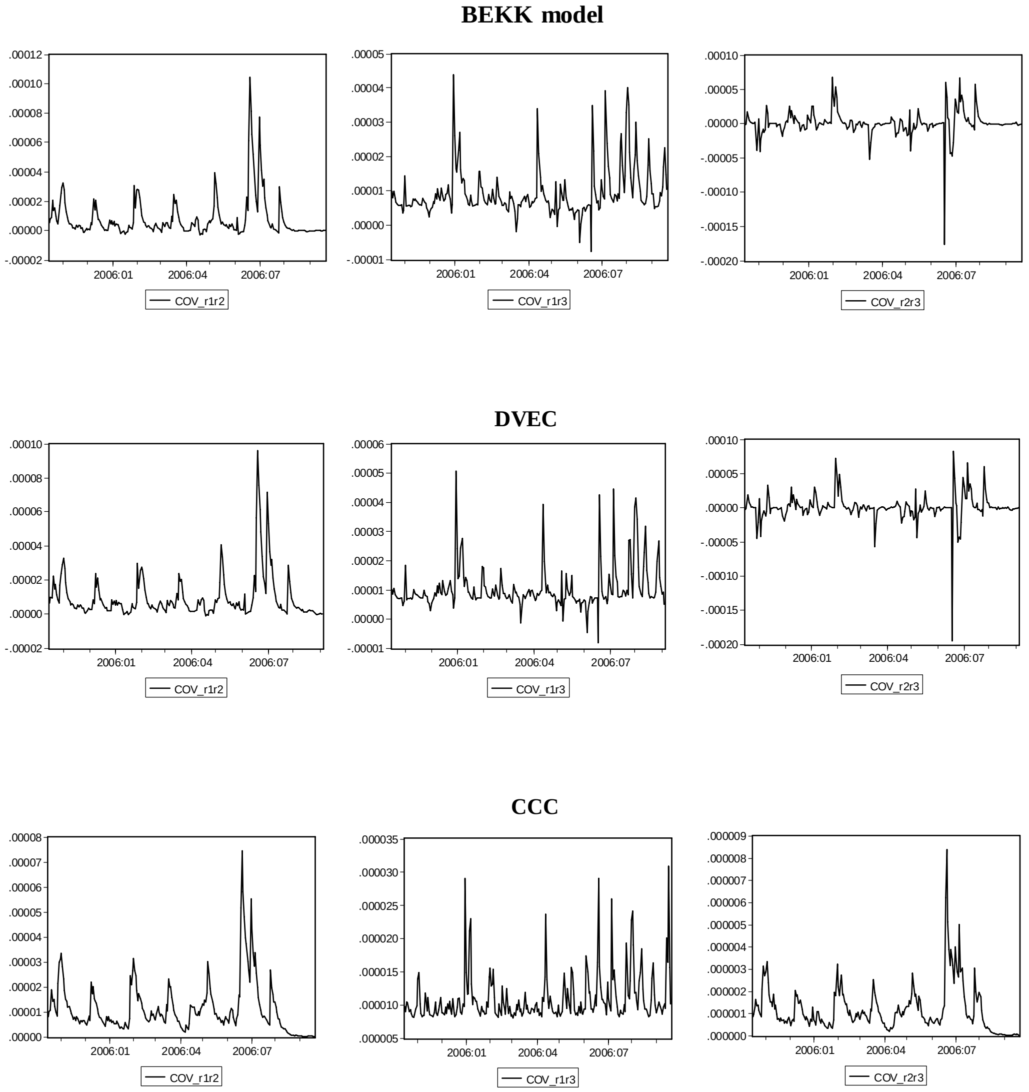
Table 3 contains the coefficients, standard errors, z-statistics, log-likelihood and information criteria for trivariate BEKK, DVEC and CCC model. The methods for estimation parameters which we use are maximum log-likehood and two-step approach. Although maximum log-likehood method can be used for all three models (BEKK, DVEC and CCC), for CCC representation we will estimate parameters using the first step of two-step approach. It is enough because CCC model uses constant correlation coefficient, and second step should be used only when correlation coefficient is time dependent (Minović, 2007).

Table 3: Estimated parameters of trivariate BEKK, DVEC and CCC models

	BEKK			DVEC			CCC		
	Coeff.	S.E.	z-Stat	Coeff.	S.E.	z-Stat	Coeff.	S.E.	z-Stat
MU(1)	-0.0003	0.0002	-1.4536	-0.0001	0.0002	-0.3216	-0.0001	0.0002	-0.6120
MU(2)	-0.0006	0.0001	-8.1946	-0.0003	0.0002	-1.6585	-0.0006	0.0001	-8.4620
MU(3)	-0.0005	0.0007	-0.7305	-0.0003	0.0006	-0.5244	-0.0003	0.0006	-0.5453
OMEGA(1)	0.0017	0.0004	4.8070	0.0000	0.0000	1.9624	0.0000	0.0000	1.3449
BETA(1)	0.7826	0.0879	8.9069	0.5216	0.1549	3.3679	0.4973	0.3058	1.6260
ALPHA(1)	0.3495	0.0739	4.7325	0.2006	0.0709	2.8288	0.1479	0.0913	1.6192
OMEGA(2)	0.0000	0.0000	0.0519	0.0000	0.0000	3.0426	0.0000	0.0000	-9.3533
BETA(2)	0.8416	0.0058	144.5231	0.6827	0.0168	40.5688	0.7085	0.0010	70.8697
ALPHA(2)	0.7682	0.0373	20.6101	0.5952	0.0623	9.5593	0.5820	0.0534	10.8965
OMEGA(3)	0.0016	0.0006	2.4940	0.0000	0.0000	2.6601	0.0000	0.0000	5.1316
BETA(3)	0.6814	0.1607	4.2397	0.3052	0.2135	1.4291	0.2269	0.1382	1.6417
ALPHA(3)	0.4283	0.0975	4.3925	0.2479	0.1152	2.1518	0.2539	0.1057	2.4011
OMEGA(4)	0.0001	0.0002	0.2752	0.0000	0.0000	5.6805	-	-	-
BETA(4)	-	-	-	0.6431	0.0690	9.3175	-	-	-
ALPHA(4)	-	-	-	0.2545	0.0692	3.6794	-	-	-
OMEGA(5)	-0.0055	0.0210	-0.2643	0.0000	0.0000	2.2551	-	-	-
BETA(5)	-	-	-	0.3899	0.2062	1.8905	-	-	-
ALPHA(5)	-	-	-	0.1765	0.0716	2.4647	-	-	-
OMEGA(6)	0.0003	0.3462	0.0010	0.0000	0.0000	-0.1434	-	-	-
BETA(6)	-	-	-	0.4663	0.2539	1.8363	-	-	-
ALPHA(6)	-	-	-	0.3630	0.1137	3.1918	-	-	-
Log likehood	2886.269			2697.955			3534.373		
Avg. log likelihood	11.8290			11.5792			14.4851		
Number of coeff.	15			21			12		

In Figure 3 we plot conditional covariances for daily log returns of BELEX15 index - Hemofarm stock (cov_r_1r_2); BELEX15 index - Energoprojekt stock (cov_r_1r_3); Hemofarm - Energoprojekt stocks (cov_r_2r_3), respectively in restricted BEKK, DVEC and CCC models.

Figure 3: Estimated conditional covariance for daily log returns of BELEX15 index and Hemorfarm stock; BELEX15 index and Energoprojekt stock; Hemofarm and Energoprojekt stocks in the trivariate BEKK, DVEC and CCC models, respectively.

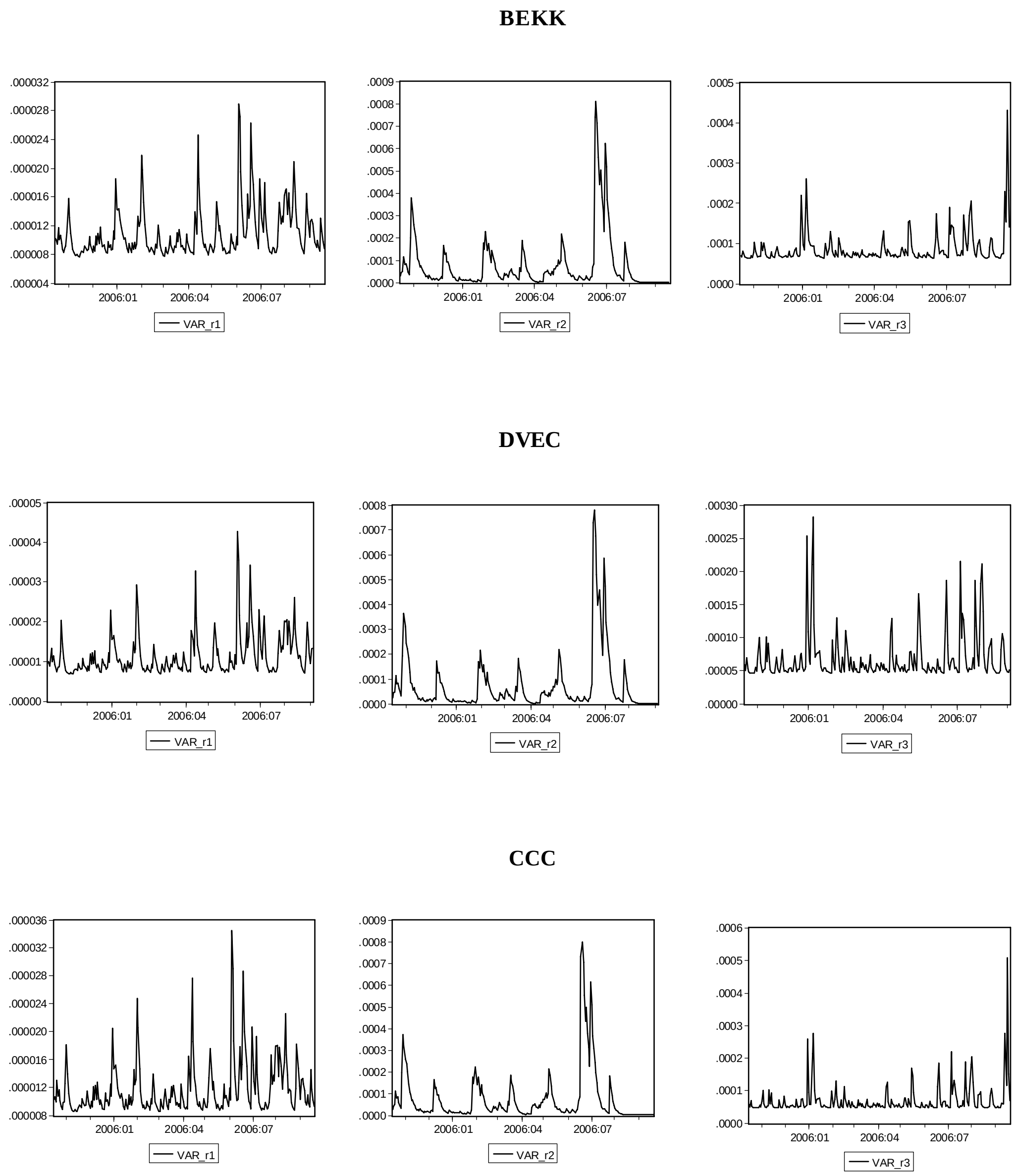


We observe from these pictures on Figure 3.1 that the restricted BEKK and DVEC results have similar behaviour for all pair log returns of stocks and index, but very different behaviour in CCC model where the covariance is positive and of not negligible magnitude especially in case Hemofarm and Energoprojekt stocks. It is because CCC model reduces to three univariate GARCH(1,1) models (covariance equations do not contain terms with cross-product of residuals). It is evidently that correlations between log returns of stocks and index are very unstable over time. Then, from Figures 3.1 we observe that Hemofram and Energoprojekt stocks are noncorrelated, this plots around zero on the graph (Minović, 2007).

On all figures in cases of BELEX15 index and Hemofarm stock as well as Hemofarm and Energoprojekt stocks we see that relationship between these stocks changes dramatically in period June-July 2006 (it was when company Schtada was bought stocks of Hemofarm). Since on all figures in case BELEX15 index and Energoprojekt stock we see that the greatest peak match with time when Energoprojekt company signed contract in Nigeria, February 2006 (Minović, 2007).

In Figure 4 we plot conditional variances for daily log returns of BELEX15 index (var_r1), Hemofram (var_r2) and Energoprojekt (var_r3) stocks, respectively in our three considering models.

Figure 4: Estimated conditional variances of daily log returns on BELEX15 index, Hemofarm stock and Energoprojekt stock, respectively in the trivariate BEKK, DVEC and CCC models.



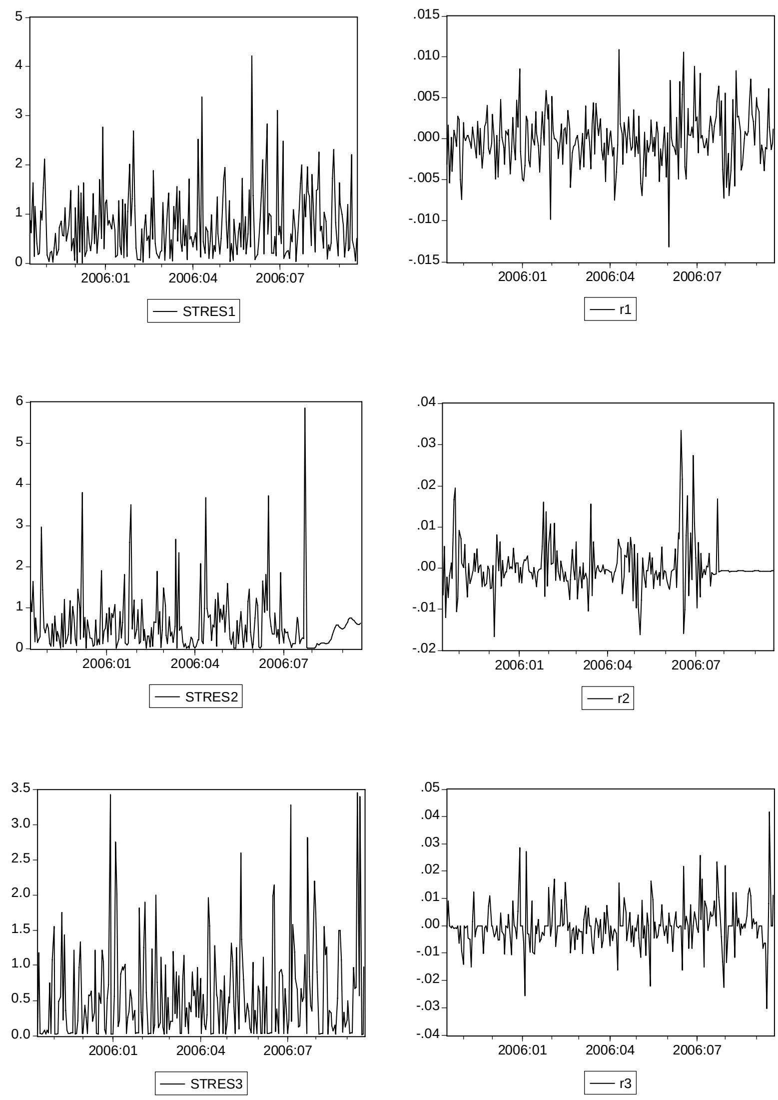
We observe from figure above that the restricted BEKK, DVEC and CCC results exhibit rather similar behaviour for each considering stocks and index. Figure 3.2 shows that BELEX15 index has always been more volatile than Hemofarm and Energoprojekt stock. On the second picture in all three models we see that time of the greatest peak match with time when Hemofarm was sold. We observe from figures that exist significant autocorrelations in data of Hemofarm stock, it is because in univariate case Hemofarm follow IGARCH process. On the third picture in all three models we observe the first peak in February 2006, when Energoprojekt company signed contract in Nigeria valued 151 million euros. The graphs of conditional variances for daily log returns of BELEX15 index, Hemofram and Energoprojekt stocks in our three considering models are

very similar to graphs in univariate case. All variances in all three models are highly unstable (Minović, 2007).
However, in order to choose the best model, diagnostic tests should be calculated.

Diagnostic checking

In the multivariate case we propose to examine the standardized residuals, squared standardized residuals as well as the cross products of the standardized residuals. Our results show that the residual-based diagnostics provide a useful check for model adequacy (for detail see reference Minović, 2007) (Tse, 2002, 358).

Figure 5: The standardized residuals of BELEX15 index (stres1), Hemofarm (stres2) and Energoprojekt (stres3) stocks versus its log returns in trivariate GARCH models.



The standardized residuals for log return of BELEX15 index, log returns of Hemofarm and Energoprojekt stocks are calculated as:

$$\hat{z}_i = \left(r_i - \hat{\mu}_i\right) / \sqrt{\sigma_i}\text{ , where } i = 1, 2, 3. \qquad (3.16)$$

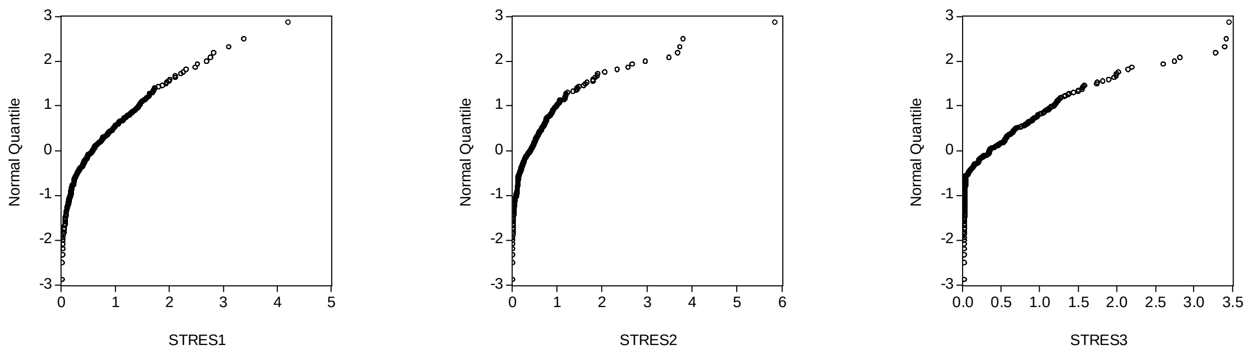
The cross product of residuals for log returns of BELEX15 index – Hemofarm stock; log returns of BELEX15 index – Energoprojekt stock; log returns of Hemofarm – Energoprojekt stocks are calculated as:

$$\hat{z}_i\hat{z}_j = \left(r_i - \hat{\mu}_i\right)\left(r_j - \hat{\mu}_j\right) / \sqrt{\sigma_{ii}\sigma_{jj}} \ . \qquad (3.17)$$

where $i = 1, 2, 3,\ j = 1, 2, 3$ and in the equation $i \neq j$. If the model is correctly specified, the standardized residuals should be uncorrelated, and identically distributed random variables with mean zero and variance one (EViews 5 User's Guide).

The goodness-of-fit of a multivariate GARCH model can also be assessed by calling the generic plot function on a fitted "mgarch" object. There is significant deviation in the tails from the normal QQ-line for both residuals, which shown earlier. Thus it seems that the normality assumption for the residuals may not be appropriate (Zivot et.al., 2006).

Figure 6: The QQ-plot of standardized residuals of BELEX15 index (stres1), Hemofarm (stres2) and Energoprojekt stocks (stres3) vs. normal distribution in trivariate GARCH models.



From Figure 6 we see that all curves have concave shape indicating that the distributions of standardized residuals are positively skewed with long right tail. Thus, there is significant deviation in the tails from the normal QQ-line for all three standardized residuals and estimates are still consistent under quasi-maximum likelihood (QML) assumptions.

For diagnostic checking we used the Ljung-Box statistics of standardized residuals and those of its squared, and of cross product of standardized residuals (Table 4 and 5). We observed that in trivariate case we have ARCH effect in variance equation of Hemofarm stock, except for DVEC model. The Q-statistics for checking whether there are any ARCH effects left in the residuals show that autocorrelation is not significant in variance equations for log returns of BELEX15 index and Energoprojekt stock (Minović, 2007).

Table 4: The Ljung-Box statistics of standardized residuals and those of its squared for log return of BELEX15 index, log return of Hemofarm and log return of Energoprojekt stocks, where the number in parentheses denotes p-value.

Q(36)	BELEX15	Hemofarm	Energoprojekt
BEKK	25.628 (0.900)	44.451 (0.158)	31.161 (0.698)
DVEC	23.205 (0.951)	43.256 (0.189)	34.341 (0.548)
CCC	22.976 (0.955)	40.311 (0.247)	26.675 (0.871)
$Q^2(36)$			
BEKK	29.221 (0.781)	58.530 (0.010)	38.323 (0.365)
DVEC	24.128 (0.935)	47.444 (0.096)	40.719 (0.270)
CCC	26.468 (0.877)	83.197 (0.000)	34.328 (0.548)

Table 5: The Ljung-Box statistics of cross product of standardized residuals, where the number in parentheses denotes p-value.

$Q^2(36)$	BELEX15-Hemofarm	BELEX15-Energoprojekt	Hemofarm- Energoprojekt
BEKK	30.429 (0.730)	32.122 (0.654)	34.498 (0.540)
DVEC	31.550 (0.680)	27.746 (0.836)	32.244 (0.648)

From Table 5 it is evident that there are no ARCH effect in covariance equations for BEKK and DVEC models for pairs BELEX15-Hemofarm; BELEX15-Energoprojekt, Hemofarm-Energoprojekt. Thus, the check of the models shows that the models are appropriate i.e. Q-statistics show that models are adequate for describing the conditional heteroscedasticity of the data.

Final conclusion: It is interesting to note that DVEC model would be the most convinient model, because only that one does not show ARCH effect for Hemofarm stock. BEKK and CCC model have smaller number of parameters and they are much easier to estimate than DVEC model. Thus, we found that the most 'complicated' model is the best model (Minović, 2007).

Conclusion

This article presents theoretical and empirical calculation for diagnostic checking of univariate and multivariate GARCH models. We illustrated our empirical approach by applying it to daily returns of the BELEX15 index, Hemofarm and Energoprojekt stocks. We presented different diagnostics for univariate GARCH models which are used in our analysis (Lagrange Multiplier test, Ljung-Box test, F-test) and we reported results of analysis. Overall, our results showed that the residual-based diagnostics provide a useful check for model adequacy. In theoretical part for multivariate case we presented three categories diagnostics for conditional heteroscedasticity models: portmanteau tests of the Box-Pierce-Ljung type, residual-based diagnostics (RB) and Lagrange Multiplier (LM) tests. The Box–Pierce–Ljung portmanteau statistic is the most widely used diagnostic. In empirical part, after a trivariate conditional heteroscedasticity model had been fitted, we used the Ljung-Box statistics (Q-test) of standardized residuals, those of its squared, as well as of the cross product of standardized residuals to check the model adequacy. Overall result is that models perform statistically well.

Literature

Bauwens, L., Laurent S. and Rombouts J. V. K. February 2006. "Multivariate GARCH models: A survey". Journal of Applied Econometrics: 79-109

Brooks, Chris. 2002. Introductory Econometrics for Finance. Cambridge University Press

DeGroot, Morris H. and Schervish Mark J. 2002. Probability and Statistics.Third Edition. Adddison-Wesley

Franke, Jürgen, Härdle Wolfgang, Hafner Christian. July 2005. Introduction to Statistics of Financial Markets. http://www.quantlet.com/mdstat/scripts/sfe/html/sfeframe131.html

Greene, William H. 2003. Econometric analysis, fifth edition. Prentice Hall

Mills, Terence C., Kerry Patterson. 2006. Handbook of Econometrics, Vol. 1, Palgrave

Minović, Jelena Z. (2007) "Multivariate GARCH models: Theoretical Survey and Model Application", Master Thesis.

Minović, Jelena. 2007. 'Application of multivariate GARCH models in Serbian financial market analysis', International Scientific Conference, Economic Faculty, Belgrade, Serbia

Minović, Jelena. 2007 'Univariate GARCH models: theoretical survey and application', BALCOR 07, Zlatibor, Serbia

Mladenović, Zorica, Pavle Petrović. 2002. Uvod u ekonometriju. Ekonomski fakultet, Beograd.

Rombouts, Jeroen V.K., Verbeek Marno. December 8, 2004. "Evaluating Portfolio Value-at-Risk using Semi-Parametric GARCH Models". http://ideas.repec.org/p/iea/carech/0414.html

Tsay, Ruey S. 2005. Analysis of Financial Time Series, Wiley, New Jersey.

Tse, Y. K. 2002. "Residual-based diagnostics for conditional heteroscedasticity models". Econometrics Journal. Vol. 5, pp. 358-373.

Tse, Y. K., and Tsui Albert K. C. 1999. "A note on diagnosing multivariate conditional heteroscedasticity models". Journal of time series analysis. Vol. 20, No. 6: 679-691

Tse, Y. K., and Tsui Albert K. C. December 1998. "A Multivariate GARCH Model with Time-Varying Correlations". http://www.econometricsociety.org/meetings/wc00/pdf/0250.pdf

Vogelvang, Ben. 2005. Econometrics: Theory and Applications with EViews. Prentice Hall.

Yang, W., David E. A., "Multivariate GARCH hedge ratios and hedging effectiveness in Australian futures market", *Accounting and Financ*, 45 (2004) 301-321.

Zivot, Eric and Jiahui Wang. 2006. Modelling Financial Time Series with S-PLUS. Springer

Quantitative Micro Software. EViews 5 User's Guide