Modeling Multivariate Volatility Processes: Theory and Evidence

▪

Jelena Z. Minović
Belgrade Banking Academy, Union University, Serbia
jminovic@ien.bg.ac.yu

Abstract. *This article presents theoretical and empirical methodology for estimation and modeling of multivariate volatility processes. It surveys the model specifications and the estimation methods. Multivariate GARCH models covered are VEC (initially due to Bollerslev, Engle and Wooldridge, 1988), diagonal VEC (DVEC), BEKK (named after Baba, Engle, Kraft and Kroner, 1995), Constant Conditional Correlation Model (CCC, Bollerslev, 1990), Dynamic Conditional Correlation Model (DCC models of Tse and Tsui, 2002, and Engle, 2002). I illustrate approach by applying it to daily data from the Belgrade stock exchange, I examine two pairs of daily log returns for stocks and index, report the results obtained, and compare them with the restricted version of BEKK, DVEC and CCC representations. The methods for estimation parameters used are maximum log-likehood (in BEKK and DVEC models) and two-step approach (in CCC model).*

Keywords: covariance; multivariate GARCH models; maximum likelihood estimation; two-step estimation; modeling.

▪

JEL Codes: C1, C3, C5.
REL Codes: 9B, 9F, 9G.

1. Introduction

Autoregressive conditional heteroscedasticity (ARCH), and the generalized ARCH (GARCH) models proved to be successful in capturing the time-varying variances of economic data in the univariate case. This has motivated many researchers to extend these models to the multivariate dimension (Tse, 2000). Multivariate volatilities have many important financial applications. About all, they play an important role in portfolio selection and asset allocation, and they can be used to compute the value at risk of a financial position consisting of multiple assets (Tsay, 2005). So, the application of multivariate GARCH (MGARCH) models is very wide. Some of the typical applications are: portfolio optimization (Kroner, Claessens, 1991), pricing of assets (Hafner, Herwartz, 1998) and derivatives, computation of the Value at Risk (Rombouts, Verbeek, 2004, Bauwens, Laurent, 2004), futures hedging (Park, Switzer, 1995, Yang, Allen, 2004, Bera et al., 1997, Lien, Luo, 1994), volatility transmitting (Karolyi, 1995) and asset allocation, estimation systemic risk in banking (Schröder, Schüler, 2003), determing of the leverage effect (De Goeij, Marquering, 2004, Kroner, Ng, 1998), estimation of the volatility impulse response function (Hafner, Herwartz 1998, 2006, Elder, 2003), nonlinear programming (Altay-Salih, Pinar, Leyffer, 2003), hedging the currency exposure risk (Kroner, Sultan, 1991, Valiani, 2004), calculation of the minimum capital risk requirements for portfolio of assets (Brooks et al., 2002), determing misspecification tests for MGARCH models (Tse, Tsui, 1999), modeling of the changing variance structure in an exchange rate regime (Bollerslev, 1990), applying MGARCH models in the analysis of the individual financial markets (Minović, 2007b).

The focus of this article is to introduce the basic concept of multivariate volatility (GARCH) modeling. To first give a theoretical survey of these models. Then, I present the estimations (maximum log likelihood and two-step approach) of these models and give the comparative analysis. MGARCH models explain how the covariances move over time. Modeling a covariance matrix is difficult because of the likely high dimensionality of the problem and the constraint that a covariance matrix must be positive definite (Pourahmadi, 1999). The crucial stage in MGARCH modeling is to provide a realistic but parsimonious specification of the variance matrix ensuring its positivity. Obviously, a disadvantage of the multivariate approach is that the number of parameters to be estimated in the GARCH equation increases rapidly. This limits the number of assets that can be included (De Goeij et al. 2004). Models covered are the VEC (initially due to Bollerslev, Engle and Wooldridge, 1988), the diagonal VEC (DVEC), the BEKK (named after Baba, Engle, Kraft and Kroner, 1995), Constant Conditional Correlation Model (CCC, Bollerslev, 1990), Dynamic Conditional Correlation Model (DCC models of Tse and Tsui, 2002, and Engle, 2002). I calculate volatility for some selected securities listed at the Belgrade stock exchange (www.belex.rs). Trivariate GARCH models are estimated using daily data from the Belgrade stock exchange for two pairs of daily log returns for both stocks and index.

I also estimate the restricted version of BEKK, DVEC, and CCC models.

The rest of the paper is organized as follows. At the beginning of the second section I analyze the basic theory of multivariate time series. Modeling of dynamic interdependent variables is conducted with multivariate time series. I define time-varying means, variances and covariances for the N components-the conditional variance-covariance matrix. The rest of Section 2 presents a theoretical survey of multivariate GARCH formulations, containing the following models: VEC, diagonal VEC, and BEKK. Section 3 describes class of models that includes: Constant Conditional Correlation (CCC) and Dynamic Conditional Correlation (DCC) models. For each of the models classes, a theoretical review is provided, following a comparative analysis discussing their advantages and disadvantages. We presented two estimation methods: maximum log likelihood (this is estimator for the first class of the models, Section 2) and two-step approach (for the second class of the models, Section 3). Section 4 presents the empirical results from applying MGARCH (trivariate version of BEKK, DVEC, and CCC) models at the Belgrade stock exchange. In particular, I test how the covariances between chosen securities move over time. The Section 5 concludes.

2. Multivariate GARCH (MGARCH) models

2.1. Multivariate time series

Often the current value of a variable depends not only on its past values, but also on past and/or current values of other variables (Schmidth, 2005). Price movements in one market can spread easily and instantly to another market. Financial markets are more dependent on each other than ever before. Consequently, knowing how the markets are interrelated is of great importance in finance. For an investor or a financial institution holding multiple assets, the dynamic relationship between returns on the assets play an important role in decision making (Tsay, 2005). Modeling of dynamic interdependent variables is done using with multivariate time series. A multivariate time series $r_t = (r_{1t}, r_{2t}, \ldots, r_{Nt})'$ is a vector of N processes that have data available for the same moments in time (Schmidth, 2005).

Multivariate time series r_t is *weakly stationary* if its first and second moments are time-invariant. In particular, the mean vector and covariance matrix of a weakly stationary series are constant over time. For a weakly stationary time series r_t, we define its mean vector and covariance matrix as

$$\mu = E(r_t) = \begin{bmatrix} E(r_{1t}) \\ \vdots \\ E(r_{Nt}) \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_N \end{bmatrix} = const. \tag{2.1}$$

$$\Gamma_0(k) = E\left[(r_t - \mu)(r_{t-k} - \mu)'\right] = \begin{bmatrix} \gamma_0 & \gamma_{12}(k) & \cdots & \gamma_{1N}(k) \\ \gamma_{21}(k) & \gamma_1 & \cdots & \gamma_{2N}(k) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{N1}(k) & \gamma_{N2}(k) & \cdots & \gamma_N \end{bmatrix}, \tag{2.2}$$

where the expectation is taken element by element over the joint distribution of r_t. The mean μ is an N-dimensional vector consisting of unconditional expectations of the components of r_t. The covariance matrix Γ_0 is an $N \times N$ matrix. The i^{th} diagonal element of Γ_0 is the variance of r_{it}, whereas the $(i, j)^{\text{th}}$ element of Γ_0 is the covariance between r_{it} and r_{jt}, and it is a function of k (Tsay, 2005).

A dynamic model with time-varying means, variances and covariances for the N components of $r_t = (r_{1t}, r_{2t}, \ldots, r_{Nt})'$ is:

$$r_t = \mu_t(\theta) + \varepsilon_t. \tag{2.3}$$

Here, θ is a finite vector of parameters, $\mu_t(\theta)$ is the conditional mean vector, and ε_t is an $N \times 1$ vector of shock, or innovation, of the series at time t equal to:

$$\varepsilon_t = \Sigma_t^{1/2}(\theta) z_t, \tag{2.4}$$

where $\Sigma_t^{1/2}(\theta)$ is an $N \times N$ positive definite matrix. Furthermore, we assume the $N \times 1$ random vector z_t to have the following first two moments:

$$E(z_t) = 0, \quad Var(z_t) = I_N, \tag{2.5}$$

where I_N is the identity matrix of order N. The conditional mean vector has the form:

$$\mu_t = E(r_t \mid I_{t-1}) = E_{t-1}(r_t), \tag{2.6}$$

where I_{t-1} is the information available at time t-1, at least containing $(r_{t-1}, r_{t-2}, \ldots)$. To make this clear we calculate the conditional variance matrix of r_t:

$$\begin{aligned} Var(r_t \mid I_{t-1}) = Var_{t-1}(r_t) &= Var_{t-1}(\varepsilon_t) \\ &= \Sigma_t^{1/2} Var_{t-1}(z_t)\left(\Sigma_t^{1/2}\right)' \\ &= \Sigma_t. \end{aligned} \tag{2.7}$$

Hence, $\Sigma_t^{1/2}$ is any $N \times N$ positive definite matrix such that Σ_t is the conditional variance matrix of r_t. Both Σ_t and μ_t depend on the unknown parameter vector θ, which can be split in most cases into two disjoint parts, one for μ_t and one for Σ_t (Bauwens, 2005; Bauwens et al., 2006). Multivariate volatility modeling is concerned with the time evolution of Σ_t. We refer to a model for the $\{\Sigma_t\}$ process as a volatility model for the return series r_t (Tsay, 2005).

The goal is to model the conditional variance-covariance matrix Σ_t which is an $N \times N$ non-negative definite matrix. Different models for Σ_t have been proposed over the last two decades (Wang, Yao, 2005).

2.2. The VEC model

First, I introduce the VEC operator:

- VEC is the operator that stacks a matrix as a column vector:

$$VEC(\Sigma_t) = \left(\sigma_{11t}, \sigma_{21t}, \ldots, \sigma_{N1t}, \sigma_{12t}, \sigma_{22t} \ldots, \sigma_{NNt}\right)'. \tag{2.8}$$

- A useful property is:

$$VEC(ABC) = (C' \otimes A) VEC(B) \text{(Bauwens, 2005).} \tag{2.9}$$

The general multivariate GARCH(p,q) model is given as:

$$VEC(\Sigma_t) = C + \sum_{i=1}^{q} A_i \cdot VEC\left(\varepsilon_{t-i}\varepsilon_{t-i}'\right) + \sum_{j=1}^{p} B_j \cdot VEC\left(\Sigma_{t-j}\right), \tag{2.10}$$

where A_i and B_j are parameter matrices containing $(N^*)^2$ parameters [with $N^*=N(N+1)/2$], whereas the vector C contains N^* coefficients. We will assume that all eigenvalues of the matrix $\sum_{i=1}^{q} A_i + \sum_{j=1}^{p} B_j$ have modulus smaller than one, in which case the vector process e_t is covariance stationary with unconditional covariance matrix given by Σ_t (Hafner, Herwartz, 2006).

A potentially serious issue with the unrestricted VEC model described by equation (2.10) is that it requires the estimation of a large number of parameters. This over-parameterization led to the development of the simplified diagonal VEC model, by Bollerslev, Engle and Wooldridge (1988), where the A and B matrices are forced to be diagonal. The result is a reduction of the number of parameters in the variance and covariance equations to 18 for the trivariate case (Brooks et al., 2003).

2.3. The diagonal VEC model

Because of the simplification that it provides, the diagonal VEC model is frequently used. Each of its variance-covariance terms is postulated to follow a GARCH-type equation. The model can be written as follows (Tse, Tsui,1999):

$$\sigma_{ij,t} = c_{ij} + \sum_{h=1}^{p} a_{hij}\varepsilon_{t-h,i}\varepsilon_{t-h,j} + \sum_{h=1}^{q} b_{hij}\sigma_{t-h,ij} \qquad 1 \le i \le j \le k, \tag{2.11}$$

where c_{ij}, a_{hij} and b_{hij} are parameters. The diagonal VEC multivariate GARCH model could also be expressed as an infinite order multivariate ARCH model, where the covariance is expressed as a geometrically declining weighted average of past cross products of unexpected returns, with recent observations carrying higher weights. An alternative solution to the dimensionality problem would be to use orthogonal GARCH or factor GARCH models (Brooks, 2002).

Now, we will present the diagonal VEC model in the form:

$$\Sigma_t = C_0^* + \sum_{i=1}^{m} A_i^* \odot \left(\varepsilon_{t-i}\varepsilon_{t-i}^{'}\right) + \sum_{j=1}^{s} B_j^* \odot \Sigma_{t-j}, \tag{2.12}$$

where m and s are non-negative integers, and $\odot$ denotes Hadamard product[(1)] (element by element matrix multiplication) (Tsay, 2005). Let us define the symmetric $N \times N$ matrices A_i^* and B_j^* as the matrices implied by the relations $A = diag\left[vec\left(A^*\right)\right]$[(2)], $B = diag\left[vec\left(B^*\right)\right]$, and C_0^* as given by $C = vec\left(C_0^*\right)$ (Bauwens et al., 2006). The model which is represented by Eq. (2.12) is DVEC(m,s) model (Tsay, 2005). Σ_t must be parameter matrices, and only the lower portions of these matrices need to be parameterized and estimated. For example, Silberberg and Pafka (2001) prove that a sufficient condition to ensure the positive definiteness of the covariance matrix Σ_t in Eq. (2.12) is that the constant term C_0^* is positive definite and all the other coefficient matrices, A_i^* and B_j^*, are positive semidefinite (De Goeij et al., 2004).

Each element of Σ_t depends only on its own past value and the corresponding product term in $\varepsilon_{t-1}\varepsilon_{t-1}^{'}$. That is, each element of a DVEC model follows a GARCH(1,1) type model. The model is simple, but it may not produce a positive-definite covariance matrix. Furthermore, the model does not allow for dynamic dependence between volatility series (Tsay, 2005).

We can construct a scalar VEC model: $A = aU$ and $B = bU$, where a and b scalars and U is matrix of ones (Bauwens, 2005).

2.4. The BEKK model

In order for an estimated multivariate GARCH model to be plausible, Σ_t is required to be positive definite for all values of the disturbances. Veryfying that this holds is a non-trivial issue even for VEC or diagonal VEC models of moderate size. To circumvent this problem, Engle and Kroner (1995) proposed a quadratic formulation for the parameters that ensured positive definiteness. This became known as the BEKK model (Brooks et al., 2003). Its number of parameters grows linearly with the number of assets. Therefore, this model is relatively parsimonious and suitable for a large set of assets (De Goeij et al., 2004). The BEKK model is in the form:

$$\Sigma_t = C_0 C_0^{'} + \sum_{k=1}^{K}\sum_{i=1}^{q} A_{ki}^{'}\varepsilon_{t-i}\varepsilon_{t-i}^{'}A_{ki} + \sum_{k=1}^{K}\sum_{i=1}^{p} B_{ki}^{'}\Sigma_{t-i}B_{ki}, \tag{2.13}$$

where C_0 is a lower triangular matrix and A_{ki} and B_{ki} are $N \times N$ parameter matrices. The BEKK representation in Eq. (2.13) is a special case of Eq. (2.10) (Hafner, Herwartz, 2006). Based on the symmetric parameterization of the model, Σ_t is almost surely positive definite provided that $C_0 \times C_0$' is positive definite (Tsay, 2005).

Engle and Kroner (1995) proved that the necessary condition for the covariance stationarity of the BEKK model is that the eigenvalues, that is the characteristic roots of $\sum_{i=1}^{q}\sum_{k=1}^{K}(A_{ik}^{*}\otimes A_{ik}^{*}) + \sum_{i=1}^{p}\sum_{k=1}^{K}(B_{ik}^{*}\otimes B_{ik}^{*})$, should be less than one in absolute value. Hence, the process can still render stationary even if there exists an element with a value greater than one in the matrix. Obviously, this condition is different from the stationarity condition required by univariate GARCH model: that the sum of ARCH and GARCH terms has to be less than one (Pang et al., 2002).

The BEKK(1,1,K) model is defined as:

$$\Sigma_t = C_0 C_0^{'} + \sum_{k=1}^{K} A_k^{'}\varepsilon_{t-1}\varepsilon_{t-1}^{'}A_k + \sum_{k=1}^{K} B_k^{'}\Sigma_{t-1}B_k, \tag{2.14}$$

where C_0, A_k and B_k are $N \times N$ matrices of parameters, but C_0 is upper triangular. One can also write $C_0 \times C_0^{'} = \Omega > 0$. Positivity of Σ_t is guaranteed if $\Sigma_0 \geq 0$. Here, there are 11 parameters, against 21 in the VEC model (Bauwens, 2005). This model allows for dynamic dependence between the volatility series (Tsay, 2005).

The diagonal and scalar BEKK models can be defined as follows:

- The diagonal BEKK model.

Take A_k and B_k as diagonal matrices. For this case, the BEKK model is a restricted version of the VEC model with diagonal matrices (Bauwens, 2005; Franke et al., 2005).

- The scalar BEKK model.

$A_k = a_k \times U$, $B_k = b_k \times U$, where a and b scalars and U is a matrix of ones (Bauwens, 2005).

The diagonal BEKK model is given by the following equations:

$$\sigma_{11,t} = c_{11}^2 + a_{11}^2 \varepsilon_{1,t-1}^2 + b_{11}^2 \sigma_{11,t-1} \quad (2.15)$$

$$\sigma_{22,t} = c_{11}^2 + c_{22}^2 + a_{22}^2 \varepsilon_{2,t-1}^2 + b_{22}^2 \sigma_{22,t-1} \quad (2.16)$$

$$\sigma_{12,t} = \sigma_{21,t} = c_{11}c_{22} + a_{11}a_{22}\varepsilon_{1,t-1}\varepsilon_{2,t-1} + b_{11}b_{22}\sigma_{12,t-1} \quad (2.17)$$

This model exhibits essentially the same problems as the full BEKK model there is no parameter in any equation that exclusively governs a particular covariance equation. Hence, it is not clear whether the parameters for α_{12} are just the result of the parameter estimates for α_{11} and α_{22}, or if the covariance equation alters the parameter estimates of the variance equations. In addition, the model is not very flexible and can therefore be misspecified. However, if the covariance exhibits a different degree of persistence than the volatilities, it is clear that either the volatility or the covariance process is misspecified (Baur, 2004).

The BEKK(1,1,1) model $\Sigma_t = \Omega + A'\varepsilon_{t-1}\varepsilon_{t-1}'A + B'\Sigma_{t-1}B$ can be written as a VEC model (subject to restrictions) using formula (2.9):

$$VEC(\Sigma_t) = VEC(\Omega) + (A \otimes A)' VEC(\varepsilon_{t-1}\varepsilon_{t-1}') + (B \otimes B)' VEC(\Sigma_{t-1}). \quad (2.18)$$

Hence, the BEKK model is weakly stationary if the eigenvalues of $(A \otimes A) + (B \otimes B)$ are smaller than one in modulus, and then

$$VEC(\Sigma) = \left(I_{N^2} - (A \otimes A)' - (B \otimes B)'\right)^{-1} VEC(\Omega) \text{ (Bauwens, 2005).} \quad (2.19)$$

2.5. Multivariate GARCH model estimation

Suppose the vector stochastic process $\{r_t\}$ (for $t = 1,\ldots,T$) has conditional mean, conditional variance matrix and conditional distribution $\mu_t(\theta_0)$, $\Sigma_t(\theta_0)$ and $p(r_t \mid \varsigma_0, I_{t-1})$ respectively. Here, $\varsigma_0 = (\theta_0 \quad \eta_0)$ is a k-dimensional parameter vector, and η_0 is the vector that contains the parameters of the distribution of the innovations z_t. Importantly, to justify the choice of the estimation procedure, we assume that the model to be estimated encompasses the true formulations of $\mu_t(\theta_0)$ and $\Sigma_t(\theta_0)$ (Bauwens et al., 2006).

The procedure used most often in estimating θ_0 involves the maximization of a likelihood function constructed under the assumption of an i.i.d. distribution for the standardized innovations z_t. The likelihood function for the i.i.d. case can then be viewed as a quasi-likelihood function (Bauwens et al., 2006).

Consequently, one has to make an additional assumption on the innovation process by choosing a density function, denoted $g(z_t(\theta);\eta)$, where η is a vector of nuisance parameters. Thus, the problem to solve is to maximize the sample log-likehood function $L_T(\theta,\eta)$ for the T observations, with respect to the vector of parameters $\varsigma=(\theta,\eta)$, where

$$L_T(\varsigma)=\sum_{t=1}^{T}\log f(r_t\mid\varsigma,I_{t-1}), \qquad (2.20)$$

with

$$f(r_t\mid\varsigma,I_{t-1})=|\Sigma_t|^{-1/2}\,g\left(\Sigma_t^{-1/2}(r_t-\mu_t)\mid\eta\right), \qquad (2.21)$$

and the dependence with respect to θ occurs through μ_t and Σ_t. The term $|\Sigma_t|^{-1/2}$ is the Jacobian that arises in the transformation from the innovations to the observables. Note that as long as $g(.)$ belongs to the class of elliptical distributions, it is a function of $z_t^{'}z_t$, the maximum likelihood estimator is independent of the decomposition choice for $\Sigma_t^{1/2}$. This is because $z_t^{'}z_t=(r_t-\mu_t)^{'}\,\Sigma_t^{-1}(r_t-\mu_t)$ (Bauwens et al., 2006).

The most commonly employed distribution in the literature is the multivariate normal, uniquely determined by its first two moments (so that $\varsigma=\theta$ since η is empty). In this case, the sample log-likehood, defined up to a constant, is:

$$L_T(\theta)=-\frac{1}{2}\sum_{t=1}^{T}\log|\Sigma_t|-\frac{1}{2}\sum_{t=1}^{T}(r_t-\mu_t)^{'}\,\Sigma_t^{-1}(r_t-\mu_t) \text{ (Bauwens et al., 2006).} \qquad (2.22)$$

Under the assumption of conditional normality, the parameters of the multivariate GARCH models of any of the above specifications can be estimated by maximizing the log likelihood function l:

$$\ell(\theta)=-\frac{TN}{2}\log 2\pi-\frac{1}{2}\sum_{t=1}^{T}\left(\log|\Sigma_t|+\varepsilon_t^{'}\,\Sigma_t^{-1}\,\varepsilon_t\right), \qquad (2.23)$$

where θ denotes all the unknown parameters to be estimated, N is the number of assets (the number of series in the system), and T is the number of observations, $\varepsilon_t=r_t-\mu_t$, and all other notation is as above. The maximum-likelihood estimate for θ is asymptotically normal. This makes the traditional procedures for statistical inference applicable (Brooks, 2002). Maximizing the log-likelihood function requires nonlinear maximization methods, because it involves only first order derivatives. The algorithm introduced by Berndt (1974) is easily implemented and particularly useful for the estimation of multivariate GARCH processes (Cízek et al., 2005).

It is well-known that the normality of the innovations is rejected in most applications dealing with high-frequency data. In particular, the kurtosis of most financial asset returns is larger than three which means that they have too many extreme values to be normally distributed. Moreover, their unconditional distribution has often fatter tails than what is implied by a conditional normal distribution: the increase of the kurtosis coefficient brought by the dynamics of the conditional variance is not usually sufficient to match adequately the unconditional kurtosis of the data (Bauwens et al., 2006).

If the conditional distribution of e_t is not normal, then maximizing Eq. (2.23) is interpreted as quasi maximum-likelihood (QML) (Hafner, Herwartz, 2006). The QML-estimator is consistent under the main assumption that the considered multivariate process is strictly stationary and ergodic (Cízek et al., 2005). This quasi-maximum likelihood (QML) estimator is suitable for models which specify conditional covariances and variances because it correctly specifies the conditional mean and the conditional variance (Bauwens, Laurent, 2002).

Estimation of multivariate GARCH models is troublesome, because the number of parameters may be large even for a moderate vector dimension N. Suppose there is enough data available for estimation, the likelihood might still be relatively "flat" as a function of many parameters. Thus, it might be hard for optimization routines to find the global maximum. Therefore, constraints on the parameter space are in many cases indispensible (Deistler, 2006).

2.6. The comparative analysis

The difficulty in estimating a VEC, or even a BEKK, model is the high number of unknown parameters. Imposing several restrictions might only help moderately. It is not surprising that these models are rarely used when the number of series is larger than 3 or 4 (Bauwens et al., 2006).

Even in the simple case of two assets, the conditional variance and covariance equations for the unrestricted VEC model contain 21 parameters. As the number of assets employed in the model increases, the estimation of the VEC model quickly becomes infeasible (Brooks, 2002). Hence, to reduce the number of parameters, Bollerslev, Engle and Wooldridge (1988) suggest the diagonal VEC (DVEC) model in which the A and B matrices are diagonal. Each variance σ_{iit} depends only on its own past squares error $\varepsilon^2_{i,t-1}$ and its own lag $\sigma_{ii,t-1}$. Each covariance σ_{ijt} depends only on its own past cross-products of errors $\varepsilon_{i,t-1}\varepsilon_{j,t-1}$, and its own lag (Bauwens, 2005). This reduces the number of parameters to be estimated from 21 to 9 when $N = 2$; and from of 78 to 18 when $N = 3$. The reduction is quite restrictive, and implies that there is no "spillover effect" (Brooks, 2002, Bauwens, 2005). In the bivariate case: DVEC model contains 9 parameters while BEKK model contains only 7 parameters. This occures because the parameters governing the dynamics of the covariance equation in BEKK model are the products of the some model's parameters that come from its two variance equations. Another way to reduce the number of parameters is to use a scalar BEKK model. This means to have A_k and B_k equal to a scalar times a matrix of ones (Bauwens et al., 2006). BEKK model allows for dependence of conditional variances of one variable on the lagged values of another variable. Causalities in variances can then be modeled (Franke et al., 2005). From a numerical optimization point of view, the BEKK model also increases the nonlinearity of the constraints, by utilizing a higher order polynomial representation compared to VEC specification (Altay-Salih et al., 2003). For each BEKK model there is an equivalent VEC representation, but not vice versa, so that the BEKK model is a special case of the VEC model (Franke et al., 2005).

The crucial point in MGARCH modeling is to provide a realistic but parsimonious specification of the variance matrix ensuring its positivity. There is a tradeoff between flexibility and parsimony. BEKK models are flexible but require too many parameters for multiple time series of more than four elements. Diagonal VEC and BEKK models are much more parsimonious, but very restrictive for the cross dynamics. They are not suitable if volatility transmission is the object of interest, but they do a good job in representing the dynamics of variances and covariances (Bauwens et al., 2006).

2.7. Advantages and disadvantages

Obviously, a disadvantage of the multivariate approach is that the number of parameters to be estimated in the GARCH equation increases rapidly, limiting the number of assets that can be included. In order to reduce the number of parameters to be estimated, it is advisable to impose some restrictions on A and B without lowering, the explanatory power of the model significantly (De Goeij et al., 2004). A disadvantage of the VEC model is that there is no guarantee of a positive semi-definite covariance matrix. A variance-covariance or correlation matrix must always be „positive semi-definite". Among other things, this means that the variance-covariance matrix will have all positive numbers on the leading diagonal and will be symmetric around this leading diagonal (Brooks, 2002). By diagonalizing the model we constrain its dynamic dependence and may introduce bias in the estimates of the other parameters. For instance, only shocks in asset i can influence the conditional variance of asset i. This assumption is quite restrictive and is obviously a disadvantage of the diagonal VEC model (De Goeij et al., 2004). An advantage of the way DVEC models are formulated is that the intuition of the GARCH model, found to be very successful, is retained (Tse, 2000).

An advantage of the BEKK model is that the conditional-variance matrices are always positive definite. This is an important advantage in simulation studies (Tse, 2000). The BEKK model allows for dynamic dependence between the volatility series. On the other hand, the model has several disadvantages. First, the parameters in A_{ki} and B_{ki} do not have direct interpretations concerning lagged values of volatilities or shocks. Second, the number of parameters employed is $N^2(p+q)+N(N+1)/2$, increasing rapidly with p and q. Limited experience shows that many of the estimated parameters are statistically insignificant that introduces complications in modeling (Tsay, 2005).

Modeling of multivariate GARCH models is challenging for several reasons: the model structure should be parsimonious but still flexible; positivity of the conditional variances should be ensured for all sample paths; and the vector GARCH process ε_t should be (wide sense) stationary (Deistler, 2006).

3. A new class of multivariate GARCH models

In the previous models, we specify the conditional covariances, in addition to the variances. Next, I review models where we specify the conditional correlations, in addition to the variances. This allows some flexibility in the specifications of the variances: they need not be the same for each component. However, we face problem if we want to specify a positive-definite conditional correlation matrix (Bauwens, 2005). This section collects models that may be viewed as nonlinear combinations of univariate GARCH models. These models are Constant Conditional Correlation Model (CCC, Bollerslev, 1990), and Dynamic Conditional Correlation Models (DCC models of Tse and Tsui, 2002, and Engle, 2002). This allows for models where one can specify separately, on the one hand, the individual conditional variances, and on the other hand, the conditional correlation matrix or another measure of dependence between the individual series. The models in this category are less demanding in terms of parameters than the models of the first category ("direct generalizations of the univariate GARCH models"). This makes their estimation easier (Bauwens et al., 2006).

3.1. Constant conditional correlation model (CCC)

The first reparametrization of Σ_t is to use the conditional coefficients and variances of ε_t. Specifically, we write Σ_t as

$$\Sigma_t \equiv \left[\sigma_{ij,t}\right] = D_t \rho_t D_t , \tag{3.1}$$

where ρ_t is the $N \times N$ conditional correlation matrix of ε_t, and D_t is $N \times N$ diagonal matrix consisting of the conditional standard deviations of elements of ε_t (i.e. $D_t = diag\left\{\sqrt{\sigma_{11,t}},\ldots,\sqrt{\sigma_{NN,t}}\right\}$).

Because ρ_t is symmetric with unit diagonal elements, the time evolution of Σ_t is governed by that of the conditional variances $\sigma_{ii,t}$ and the elements $\rho_{ij,t}$ of ρ_t, where $j < i$ and $1 \le i \le N$. Therefore, to model the volatility of ε_t, it suffices to consider the conditional variances and correlation coefficients of ε_{it} (Tsay, 2005).

The conditional variances $\sigma_{ii,t}$ are modeled by a univariate GARCH model. Hence,

$$\sigma_{ij,t} = \rho_{ij,t}\sqrt{\sigma_{ii,t}\sigma_{jj,t}} \quad \forall i \neq j . \tag{3.2}$$

Positivity of S_t follows from the positivity of ρ_t and that of each $\sigma_{ii,t}$ (Bauwens, 2005).

As we said before, it is often difficult to verify the condition that the conditional-variance matrix of an estimated multivariate GARCH model is positive definite. Furthermore, such conditions are often very difficult to impose during the optimization of the log-likelihood function. However, if we postulate the simple assumption that the correlations are time invariant, these difficulties elegantly disappear (Tse, 2000).

Bollerslev (1990) suggested a multivariate GARCH model in which all conditional correlations are constant and the conditional variances are modeled by univariate GARCH models. This is so-called CCC model (constant conditional correlation) (Franke et al., 2005).

The constant conditional correlation model (CCC) is defined as

$$\rho_t = \rho = \left[\rho_{ij}\right], \ \rho_{ii} = 1. \tag{3.3}$$

Hence,

$$\sigma_{ij,t} = \rho_{ij}\sqrt{\sigma_{ii,t}\sigma_{jj,t}} \quad \forall i \neq j , \tag{3.4}$$

and the dynamics of the covariances is determined only by the dynamics of the two conditional variances. There are $\frac{N(N-1)}{2}$ parameters in ρ (Bauwens, 2005).

Because of its simplicity, the CCC model has been very popular in empirical applications. A specific member of the group of CCC models is obtained by further constraining the correlations to be zero. This model is denoted as the nocorrelation (NC) model. Thus the CCC model is given by

$$\sigma_{ii,t} = c_i + \sum_{h=1}^{p} a_{hi}\varepsilon_{t-h,i}^2 + \sum_{h=1}^{q} b_{hi}\sigma_{t-h,i} \qquad i = 1,\dots,k \tag{3.5}$$

$$\sigma_{ij,t} = \rho_{ij}\sqrt{\sigma_{ii,t}\sigma_{jj,t}} \quad 1 \le i < j \le k \tag{3.6}$$

and the NC model is its special case with $\rho_{ij} = 0$ (Tse, Tsui,1999).

The restriction that the constant conditional correlations, and thus the conditional covariances, are proportional to the product of the corresponding conditional standard deviations highly reduces the number of unknown parameters and thus simplifies estimation (Bauwens et al., 2006).

3.2. Dynamic conditional correlation model (DCC)

A new class of multivariate models called dynamic conditional correlation (DCC) model is proposed by Tse, Tsui, and Engle (using the parameterization in Eq. (3.1)). They are flexibile like univariate GARCH models and parsimonious parametric models for the correlations. They are not linear but can often be very simply estimated using univariate or two step likelihood function based methods. DCC models perform well in a variety of situations (Engle, 2002). The models of Tse and Tsui (2002) and Engle (2002) are genuinely multivariate and are useful for modeling high dimensional data sets (Bauwens et al., 2006).

For *N*-dimensional returns, Tse and Tsui (2002) assume that the conditional correlation matrix ρ_t follows the model $DCC_T(M)$ (Tsay, 2005):

$$\rho_t = (1-\theta_1-\theta_2)\rho + \theta_1\psi_{t-1} + \theta_2\rho_{t-1} \tag{3.7}$$

$$\psi_{ij,t-1} = \frac{\sum_{m=1}^{M} e_{i,t-m}e_{j,t-m}}{\sqrt{\left(\sum_{m=1}^{M} e_{i,t-m}^2\right)\left(\sum_{m=1}^{M} e_{j,t-m}^2\right)}} \tag{3.8}$$

$$e_{it} = \varepsilon_{it} / \sqrt{\sigma_{iit}} \tag{3.9}$$

where θ_1 and θ_2 are scalar parameters, $\theta_1, \theta_2 > 0$ and $\theta_1 + \theta_2 < 1$. ρ is like in CCC, an $N \times N$ positive-definite matrix with unit diagonal elements, and ψ_{t-1} is the $N \times N$ sample correlation matrix using shocks (e_t) from $t = t - M,\ t - M + 1,\ \dots,\ t - 1$ for a prespecified M. A necessary condition to ensure positivity of ψ_{t-1} is that $M \ge N$. Notice that $\psi_{ii,t-1} = 1$ for each i by construction (Tsay, 2005, Bauwens, 2005). ρ_t is a weighted average of correlation matrices (ρ, ψ_{t-1}, ρ_{t-1}). Hence, $\rho_t > 0$ if any of three components is greater than zero (Bauwens, 2005). Notice further that the CCC model is nested in this model. Therefore, one can test $\theta_1 = \theta_2 = 0$ to check whether imposing constant conditional correlations is empirically relevant (Bauwens et al., 2006).

Estimation of the two scalar parameters θ_1 and θ_2 requires special constraints to ensure positive definiteness of the correlation matrix. This is a parsimonious model, but it can be difficult to apply it. The choice of ρ and *M* deserves a careful investigation (Tsay, 2005).

Engle (2002) proposes the model $DCC_E(1,1)$:

$$\rho_t = \left(diag\, Q_t\right)^{-1/2} Q_t \left(diag\, Q_t\right)^{-1/2} \quad (3.10)$$

where $Q_t = (q_{ij,t})$ is an $N \times N$ matrix, symmetric and positive, given by

$$Q_t = (1-\theta_1-\theta_2)\bar{Q} + \theta_1 e_{t-1} e_{t-1}' + \theta_2 Q_{t-1}, \quad (3.11)$$

where $e_t = (e_{1t}, e_{2t}, \ldots, e_{Nt})'$ is the standardized innovation vector with elements $e_{it} = \varepsilon_{it} / \sqrt{\sigma_{iit}}$. $\bar{Q}$ is the unconditional covariance matrix of e_t, it is an $N \times N$ matrix, symmetric and positive. θ_1 and θ_2 are non-negative scalar parameters satisfying $0 < \theta_1 + \theta_2 < 1$, which implies that $Q_t > 0$ and $\psi_t > 0$ (Tsay, 2005, Bauwens, 2005). Q_t is the covariance matrix of e_t, since q_{iit} is not equal to one by construction. If $\theta_1 = \theta_2 = 0$, and $\bar{q}_{ii} = 1$, we obtain the CCC model. Hence, one can test for CCC against $DCC_E(1,1)$ (Bauwens, 2005).

In both DCC models all the conditional correlations have the same dynamics. This saves a lot of parameters compared to VEC and BEKK models, but is quite restrictive (especially when *N* is large) (Bauwens, 2005).

To conclude, DCC models enable us to use flexible GARCH specifications in the variance part. Indeed, as the conditional variances (together with the conditional means) can be estimated using *N* univariate models, one can easily extend the DCC-GARCH models to more complex GARCH-type structures (Bauwens et al., 2006).

3.3. Estimation

Univariate GARCH models have met widespread empirical success. However, the problems associated with the estimation of multivariate GARCH models with time-varying correlations have constrained the researchers in estimating only the models with limited scope or considerable restrictions. Large time-varying covariance matrices are needed in portfolio management and optimization, models of the term structure of treasuries or commodities, and large vector autoregressions (Engle et al., 2001).

Maximum likelihood method can be used for all these conditional correlation models. Nevertheless, some two-step estimation approaches have been developed to increase the computational efficiency, and have been used more often in practice (Wang, Yao, 2005). A feature allows for easy empirical application of the DCC models presented above is that they can be estimated consistently using a two-step approach. This makes the approach feasible when *N* is high. Engle and Sheppard (2001) show that in the case of a DCC_E model, log-likelihood can be written as the sum of a mean and volatility part (depending on a set of unknown parameters θ_1^*) and a correlation part (depending on θ_2^*) (Bauwens et al., 2006). The separation shows that a two-step estimation procedure is feasible and that variances and correlations can be estimated separately. The two-stage approach mainly has the advantage that the dimensionality of the maximization problem is reduced, accelerating the maximization process (Baur, 2004). Indeed, recalling that the conditional variance matrix of a DCC model can be expressed as $\Sigma_t = D_t \rho_t D_t$, we get an inefficient but consistent estimator of the parameter θ_1^* when the identity matrix replace ρ_t (2.24). In this case, the quasi-likelihood function corresponds to the sum of likelihood functions of N univariate models:

$$QL1_T\left(\theta_1^*\right) = -\frac{1}{2}\sum_{t=1}^{T}\sum_{i=1}^{N}\left[\log\left(\sigma_{iit}\right) + \frac{\left(r_{it}-\mu_{it}\right)^2}{\sigma_{iit}}\right]. \tag{3.12}$$

Given θ_1^*, and under appropriate regularity conditions, we get a consistent, but inefficient, estimator of θ_2^* by maximizing:

$$QL2_T\left(\theta_2^* \mid \theta_1^*\right) = -\frac{1}{2}\sum_{t=1}^{T}\left(\log|\rho_t| + e_t'\rho_t^{-1}e_t\right), \tag{3.13}$$

where $e_t = D_t^{-1}\left(r_t - \mu_t\right)$. The sum of the likelihood functions in (3.12) and (3.13), plus one half of the total sum of squared standardized residuals ($\sum_t e_t' e_t / 2$) (which is almost exactly equal to $NT/2$), is equal to the log-likelihood (2.24). This enables us to compare the log-likelihood of the two-step approach with that of the one-step approach, and that of other models (Bauwens et al., 2006).

Engle and Sheppard (2001) explain why the estimators $\hat{\theta}_1^*$ and $\hat{\theta}_2^*$, obtained by maximizing (3.12) and (3.13) separately, are not fully efficient. This is because they are limited information estimators, and it is true even if e_t is normally distributed. However, one iteration of a Newton-Raphson algorithm[3] applied to the total likelihood (2.24), starting at ($\hat{\theta}_1^*$, $\hat{\theta}_2^*$), provides an estimator that is asymptotically efficient (Bauwens et al., 2006).

3.4. The comparative analysis

The assumption of constant correlation in the CCC model makes the estimation of a large model feasible and ensures that the estimator is positive definite. This is achieved by simply requiring each univariate conditional variance to be non-zero and the correlation matrix to be of full rank (Engle et al., 2001). The CCC model has a smaller number of parameters which makes it simple to estimate.

To show explicitly the difference between DCC_T and DCC_E, we write the expression of the correlation coefficient in the bivariate case: for the $DCC_T(M)$ (Bauwens, 2005, Bauwens et al., 2006),

$$\rho_{12t} = \left(1-\theta_1-\theta_2\right)\rho_{12} + \theta_2\rho_{12,t-1} + \theta_1\frac{\sum_{m=1}^{M} e_{1,t-m}e_{2,t-m}}{\sqrt{\left(\sum_{m=1}^{M} e_{1,t-m}^2\right)\left(\sum_{m=1}^{M} e_{2,t-m}^2\right)}} \tag{3.14}$$

and for the $DCC_E(1,1)$,

$$\rho_{12t} = \frac{(1-\theta_1-\theta_2)\bar{q}_{12} + \theta_1 e_{1,t-1} e_{2,t-1} + \theta_2 q_{12,t-1}}{\sqrt{\left((1-\theta_1-\theta_2)\bar{q}_{11} + \theta_1 e_{1,t-1}^2 + \theta_2 q_{11,t-1}\right)\left((1-\theta_1-\theta_2)\bar{q}_{22} + \theta_1 e_{2,t-1}^2 + \theta_2 q_{22,t-1}\right)}}. \tag{3.15}$$

Unlike DCC_T, DCC_E does not model the conditional correlation as a weighted sum of past correlations. Indeed, the matrix Q_t is written like a GARCH equation, and then transformed into a correlation matrix. Dynamic conditional correlation (DCC) models allow for different persistence between variances and correlations, but impose common persistence in the latter (although this restriction may be relaxed). They allow for modeling a very large number of series. Easy to estimate extension of the CCC model, they have been used in studies of volatility transmission as an alternative to the BEKK model (Bauwens et al., 2006).

3.5. Advantages and disadvantages

The parameterization given by Eq. (3.1) is useful because it models covariances and correlation directly. Yet, it has several weaknesses. First, the likelihood function becomes complicated when $N \geq 3$. Second, the approach requires a constrained maximization in estimation to ensure the positive definiteness of S_t. The constraint becomes complicated when N is large (Tsay, 2005).

CCC model has the advantage of applicability without restriction to large systems of time series. On the other hand, the assumption of constant correlation is possibly quite restrictive. For example, in the empirical analysis of financial markets one typically observes increasing correlation in times of crisis or crash (Franke et al., 2005). The assumption that the conditional correlations are constant may indeed seem unrealistic in many empirical applications. A group of authors have proposed a generalization of the CCC model by making the conditional correlation matrix time dependent (DCC model). An additional difficulty is that the time dependent conditional correlation matrix has to be positive definite for all *t*. The DCC models guarantee this when simple conditions are imposed on the parameters (Bauwens et al., 2006). That multivariate and univariate volatility forecasts are consistent with each other is a desirable practical feature of the DCC models. When new variables are added to the system, the volatility forecasts of the original assets will be unchanged. Correlations may also remain unchanged depending on how the model is revised. These DCC models, which parameterize the conditional correlations directly, are naturally estimated in two steps – the first is a series of univariate GARCH estimates and the second the correlation estimate. These methods have clear computational advantages over multivariate GARCH models. The number of parameters to be estimated in the correlation process is independent of the number of series to be correlated. This allows for estimation of very large correlation matrices (Engle, 2002).

4. Empirical results

In this section, I apply trivariate time series models. I estimate the multivariate

GARCH models using EViews program, Version 4.1, using daily data from the Belgrade stock exchange for two pairs of daily log returns for stocks and index. Daily log returns data are for BELEX15 index, Hemofarm stock, and Energoprojekt stock. They cover the period from October 3, 2005 to October 6, 2006. A first simple method to estimate the parameters of univariate and 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, thus BHHH algorithm needs suitable initial parameters (Franke et al., 2005). Additionally, I use computer program for modeling restricted version of trivariate BEKK model, and extend it on the trivariate case of DVEC and CCC models. For all calculations in the programs used, the number of iterations is 100 and the convergence criterion is 1×10^{-5} which ought to be considered a procedure with high precision.

4.1. Data

I apply log-difference transformation to convert data into continuously compounded returns. The price series (log values) of both stocks and index are not stationary and but become stationary when they are first differenced. Let r_{1t}, r_{2t}, r_{3t} and be the log return series corrected for autocorrelation in the mean of BELEX15 index, Hemofarm and Energoprojekt stocks, respectively. Figure 4.1 shows the plots of daily log returns.

BELEX15 tracks free float capitalization of 15 most liquid, continuously traded stocks. Maximum weight for each component is limited to 20%. BELEX15 is calculated and published both with intraday and closing values. Index base period is October 1, 2005, and base value was 1,000.00 index points (http://www.sinteza.net/). BELEX15 is not adjusted for paid dividends, and is not protected from dilution effect, that appears as a result of dividends' payout (http://www.belex.rs/index-e.php).

"Hemofarm Concern" is the biggest pharmaceutical company in Serbia. We observe from Figure 4.1 that in July 2006 price of Hemofarm stock rapidly grew. This happened because German Schtada Company bought 67% stocks of Hemofarm. After that price curve is flat, which means that the price of Hemofarm stocks was not changing.

"Energoprojekt" company is one of the most important firms in Serbia in the construction industry. Company got a job in Nigeria valued at 151 million euros in February 2006. We see that as the first peak of price jump on Figure 4.1. After that price again goes down until July 2006, when positive business results raise investment in Energoprojekt stock, with constant positive price trend from than on. Possible reasons for this are previous underpricing of Energoprojekt's fundamentals compared to similar companies in the Balkan region and the entrance of new funds to the Belgrade Stock Exchange without a broad offer of high quality stocks. Additionally, management of "Energoprojekt" was focused on activities which supported its stock price.

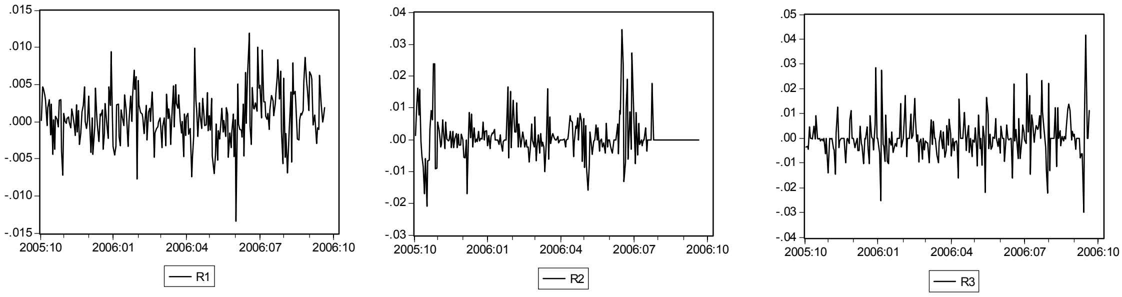


Figure 4.1. *The graphs of daily log returns of BELEX15 index (r_1), Hemofarm (r_2) and Energoprojekt (r_3) stocks, respectively*

We observe on Figure 4.1 that the log returns of BELEX15 index, Hemofarm and Energoprojekt stocks offer evidence of the well known volatility clustering effect. It is a tendency for volatility in financial markets to appear in bunches. Large returns (of either sign) are expected to follow large returns, and small returns (of either sign) to follow small returns (Brooks, 2002).

I found that the correlation coefficients (only the first measure of correlation) between log returns of BELEX15 index and Hemofarm stock is 0.49; that between log returns of BELEX15 index and Energoprojekt stock is 0.40; and that between log returns of Hemofarm and Energoprojekt stocks about is 0.02. This means that these two stocks are unncorrelated.

Prior to modeling multivariate volatility processes, univariate GARCH analysis was performed. I used four steps in building a volatility model for each of the analyzed return series. The first step was to specify a mean equation by testing for serial dependence in the data, and building an ARMA model for the return series so as to remove any linear dependence. The second step, we to use the residuals of the mean equation to test for ARCH effects. The third step was to specify a volatility model when the ARCH effects were statistically significant and perform a joint estimation of the mean and volatility equations. This allowed us to conclude that the right model for BELEX15 index is the ARMA(1,1)-GARCH(1,1), for Hemofarm stock is the ARMA(2,2)-IGARCH(1,1), and for Energoprojekt stock ARMA(0,0)-GARCH(1,1) model. Finally, in the fourth step, we checked the fitted models carefully: the Ljung-Box statistic of standardized residuals and its squared values showed that models are adequate for describing the conditional heteroscedasticity of the data (for detail see Minović, 2007a).

4.2. Modeling of trivariate volatility processes

The methods for the estimation of parameters which I 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 the first step of two-step approach is used. This is enough because CCC model has constant correlation coefficient, and the second step needs to be used only when the correlation coefficient is time dependent.

Using values from Table A.1 in Appendix A we have restricted trivariate BEKK model:

$$\hat{\sigma}_{11,t} = \underset{(0.0004)}{0.0017}^2 + \underset{(0.0739)}{0.3495}^2\, \varepsilon_{1,t-1}^2 + \underset{(0.0879)}{0.7826}^2\, \sigma_{11,t-1}, \tag{4.1}$$

$$\hat{\sigma}_{22,t} = \underset{(0.0002)}{0.0001}^2 + \underset{(0.0373)}{0.7682}^2\, \varepsilon_{2,t-1}^2 + \underset{(0.0058)}{0.8416}^2\, \sigma_{22,t-1}, \tag{4.2}$$

$$\hat{\sigma}_{33,t} = \underset{(0.0006)}{0.0016}^2 - \underset{(0.0210)}{0.0055}^2 + \underset{(0.3462)}{0.0003}^2 + \underset{(0.0975)}{0.4283}^2\, \varepsilon_{3,t-1}^2 + \underset{(0.1607)}{0.6814}^2\, \sigma_{33,t-1}, \tag{4.3}$$

$$\hat{\sigma}_{12,t} = \underset{(0.0739\cdot 0.0373)}{0.3495\cdot 0.7682}\, \varepsilon_{1,t-1}\varepsilon_{2,t-1} + \underset{(0.0879\cdot 0.0058)}{0.7826\cdot 0.8416}\, \sigma_{12,t-1}, \tag{4.4}$$

$$\hat{\sigma}_{13,t} = \underset{(0.0004\cdot 0.0006)}{0.0017\cdot 0.0016} + \underset{(0.0739\cdot 0.0975)}{0.3495\cdot 0.4283}\, \varepsilon_{1,t-1}\varepsilon_{3,t-1} + \underset{(0.0879\cdot 0.1607)}{0.7826\cdot 0.6814}\, \sigma_{13,t-1}, \tag{4.5}$$

$$\hat{\sigma}_{23,t} = -\underset{(0.0002\cdot 0.0210)}{0.0001\cdot 0.0055} + \underset{(0.0373\cdot 0.0975)}{0.7682\cdot 0.4283}\, \varepsilon_{2,t-1}\varepsilon_{3,t-1} + \underset{(0.0058\cdot 0.1607)}{0.8416\cdot 0.6814}\, \sigma_{23,t-1}. \tag{4.6}$$

Similarly, the trivariate DVEC model has the form:

$$\hat{\sigma}_{11,t} = \underset{(0.0709)}{0.2006}\, \varepsilon_{1,t-1}^2 + \underset{(0.1549)}{0.5216}\, \sigma_{11,t-1}, \tag{4.7}$$

$$\hat{\sigma}_{22,t} = \underset{(0.0623)}{0.5952}\, \varepsilon_{2,t-1}^2 + \underset{(0.0168)}{0.6827}\, \sigma_{22,t-1}, \tag{4.8}$$

$$\hat{\sigma}_{33,t} = \underset{(0.1152)}{0.2479}\, \varepsilon_{3,t-1}^2 + \underset{(0.2135)}{0.3052}\, \sigma_{33,t-1}, \tag{4.9}$$

$$\hat{\sigma}_{12,t} = \underset{(0.0692)}{0.2545}\, \varepsilon_{1,t-1}\varepsilon_{2,t-1} + \underset{(0.0690)}{0.6431}\, \sigma_{12,t-1}, \tag{4.10}$$

$$\hat{\sigma}_{13,t} = \underset{(0.0716)}{0.1765}\, \varepsilon_{1,t-1}\varepsilon_{3,t-1} + \underset{(0.2062)}{0.3899}\, \sigma_{13,t-1}, \tag{4.11}$$

$$\hat{\sigma}_{23,t} = \underset{(0.1137)}{0.3630}\, \varepsilon_{2,t-1}\varepsilon_{3,t-1} + \underset{(0.2539)}{0.4663}\, \sigma_{23,t-1}. \tag{4.12}$$

Finally, the trivariate CCC model has the form:

$$\hat{\sigma}_{11,t} = \underset{(0.0913)}{0.1479}\, \varepsilon_{1,t-1}^2 + \underset{(0.3058)}{0.4973}\, \sigma_{11,t-1}, \tag{4.13}$$

$$\hat{\sigma}_{22,t} = \underset{(0.0534)}{0.5820}\, \varepsilon_{2,t-1}^2 + \underset{(0.0010)}{0.7085}\, \sigma_{22,t-1}, \tag{4.14}$$

$$\hat{\sigma}_{33,t} = \underset{(0.1057)}{0.2539}\, \varepsilon_{3,t-1}^2 + \underset{(0.1382)}{0.2269}\, \sigma_{33,t-1}, \tag{4.15}$$

$$\hat{\sigma}_{12,t} = \rho_{12}\sqrt{\sigma_{11,t}\sigma_{22,t}}, \text{ where } \rho_{12} = 0.4917, \tag{4.16}$$

$$\hat{\sigma}_{13,t} = \rho_{13}\sqrt{\sigma_{11,t}\sigma_{33,t}}, \text{ where } \rho_{13} = 0.3992, \tag{4.17}$$

$$\hat{\sigma}_{23,t} = \rho_{23}\sqrt{\sigma_{22,t}\sigma_{33,t}}, \text{ where } \rho_{23} = 0.0218. \tag{4.18}$$

Equations (4.1)-(4.6), (4.7)-(4.12) and (4.13)-(4.18) show that these models do not allow for dynamic dependence between the volatility series. From the equations for CCC model (from (4.13) to (4.18)) we can see that trivariate CCC model reduces to three univariate GARCH(1,1) models, where univariate GARCH models are estimated for each asset and then the correlation matrix is estimated via one step approach.

It is evident from the behaviour of conditional covariances (Figure 4.2) that correlation between log returns for BELEX15 index-Hemofarm stock and between BELEX15 index-Energoprojekt stock is very unstable over time.

Then, all figures that show ploted covariances (Figure 4.2) as well as all those with variances (Figure A.1 in Appendix A) of daily log returns of Hemofarm stock, we can see significant autocorrelation. This is because Hemofarm, in the univariate case, follows Integrated GARCH (IGARCH) process.

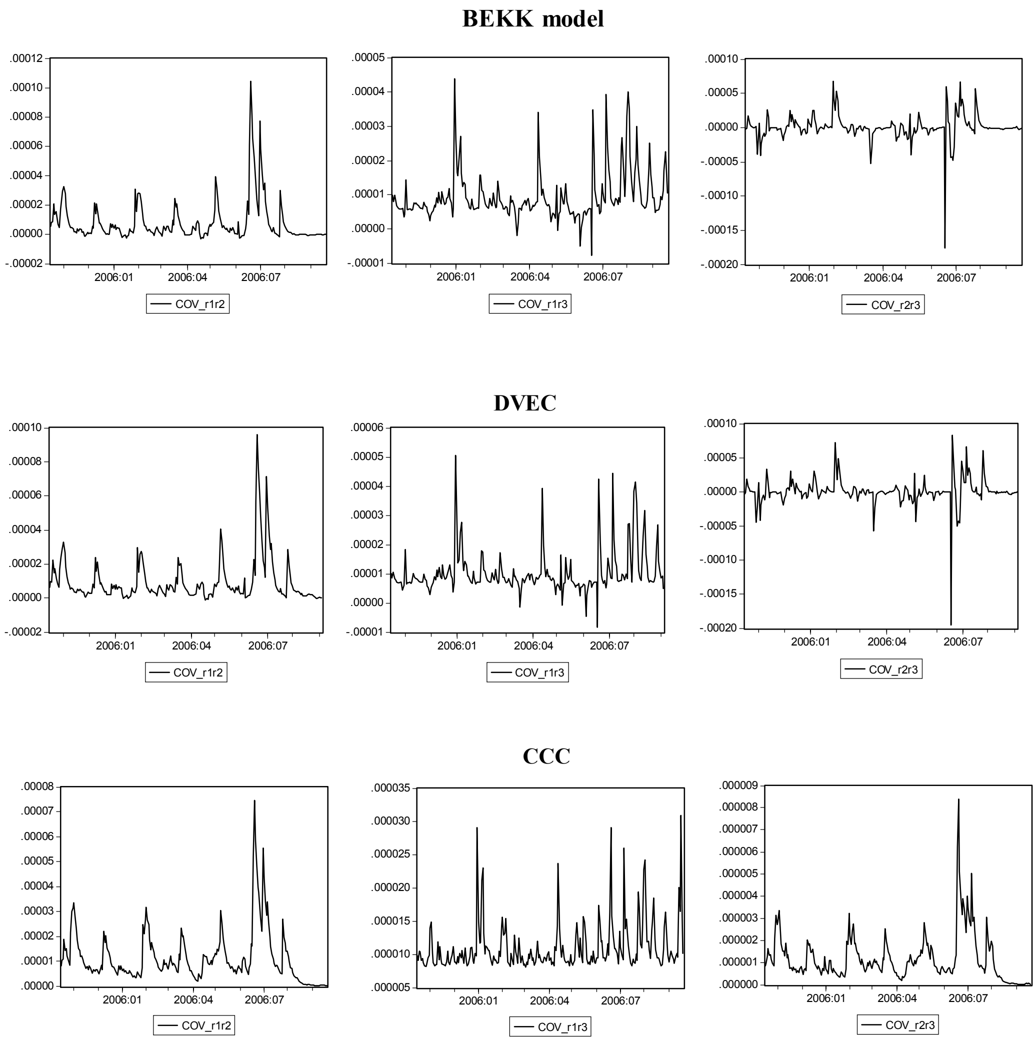


Figure 4.2. *Estimated conditional covariance for daily log returns of BELEX15 index and Hemorfarm stock ($cov\_r_1r_2$); BELEX15 index and Energoprojekt stock ($cov\_r_1r_3$); Hemofarm and Energoprojekt stocks ($cov\_r_2r_3$) in the trivariate BEKK, DVEC and CCC models, respectively*

On the other hand, it has been frequently observed in multivariate estimation frameworks that volatility changes over time (see Figure A.1 in Appendix A). I showed in the trivariate case that BELEX15 is more volatile than Hemofarm and Energoprojekt stocks. In all the models, estimated conditional variance of Hemofarm stock has the greatest peak at the time that Hemofarm was sold. In the case of variance of Energoprojekt stock, the first peak was in February 2006. This coincides with the time a significant contract was signed. However, in order to choose the best model, diagnostic tests should be calculated.

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

Table 4.1

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)

The Ljung-Box statistics of cross product of standardized residuals, where the number in parentheses denotes *p*-value

Table 4.2

$Q^2(36)$	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)

4.3. Analysis of the results

For diagnostic checking we used the Ljung-Box statistics of standardized residuals and those of its squared, and of cross product of standardized residuals (Tables 4.1 and 4.2). We observe 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.

From Table 4.2 it is evident that there are no ARCH effects in covariance equations for BEKK and DVEC models for pairs BELEX15-Hemofarm; BELEX15-Energoprojekt, and Hemofarm-Energoprojekt. Thus, check of the models shows that they are appropriate: Q-statistics show that the models are adequate for describing the conditional heteroscedasticity of the data.

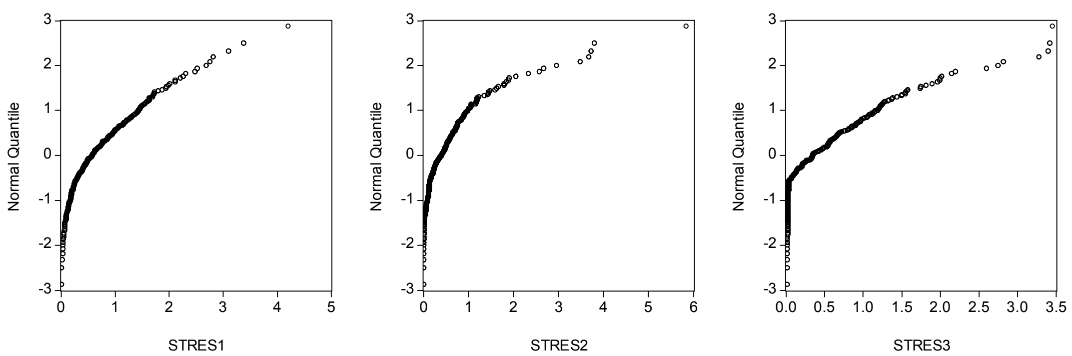


Figure 4.3. *The QQ-plot of standardized residuals of BELEX15 index (stres1), Hemofarm (stres2) and Energoprojekt stocks (stres3) plotted against normal distribution in trivariate GARCH models*

From Figure 4.3 we see that all the curves have a concave shape. This indicates that the distributions of standardized residuals are positively skewed with a long right tail. Thus, there is a 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.

It is important to note that DVEC model would be the most appropriate model, because only that one does not show ARCH effect for the Hemofarm stock. BEKK and CCC model have a smaller number of parameters and they are much easier to estimate than DVEC model. Thus, the most "complicated" model proved to be the best model.

5. Conclusion

Volatility plays an important role in controlling and forecasting risks in various financial operations. For a univariate return series, volatility is often represented in terms of conditional variances or conditional standard deviations. Many statistical models have been developed for modeling univariate conditional variance processes (Wang, Yao, 2005). The univariate volatility models have a limitation – they model the conditional variance of each series independently of all other series. This could be a significant limitation for two reasons. First, to the extent that there could be "volatility spilovers" between markets or assets, the univariate model would be misspecified. Second, the covariances between series are of interest, as well as the variances of the individual series themselves, as is often the case in finance. While univariate descriptions are useful and important, problems of risk assessment, asset allocation, hedging in futures markets and options pricing, portfolio Value at Risk estimates, CAPM betas, and so on require a multivariate framework. All the problems mentioned above require covariances as inputs. Multivariate GARCH models can potentially overcome both of these deficiencies of their univariate counterparts. In addition, there are many situations when empirical multivariate models of conditional heteroscedasticity can be used fruitfully (Brooks, 2002).

In this article I presented a summary of theoretical and empirical modeling with multivariate GARCH models. Because of heteroscedasticity, multivariate GARCH models are suitable for the estimation of volatility and correlation of multivariate time series. For the empirical work, the restricted BEKK, DVEC, and CCC models are preferable. This is because they are much easier to estimate, while maintaining a sufficient level of generality. These models are relatively simple in comparison to other unrestricted models version, which allows one to achieve reliable estimates of variances and covariances. We showed that even DVEC, CCC, and restricted version of BEKK models with a reduced number of parameters give results that are accurate to an acceptable degree. All analyzed models (BEKK, DVEC, and CCC) show similar behavior in variances and covariances.

The main finding is that conditional covariances exhibit significant changes over time for both stocks and index. Thus, I conclude that these models overcome the usual concept of the time invariant correlation coefficient. In the trivariate case, the restricted BEKK and DVEC, give results that are similar for each pair of log returns of stocks and index, but those of CCC model are very different, covariance being positive and not of a negligible magnitude.

After a trivariate conditional heteroscedasticity model had been fitted, we used the Ljung-Box statistics (Q-test) of standardized residuals, those of the squared residuals, residuals, and of the cross product of standardized residuals to check for the model's adequacy. The overall result is that models perform well statistically.

Acknowledgements

The author wishes to thank Professor Zorica Mladenoviæ from Faculty of Economics, University of Belgrade. This article benefited from discussions with her, her helpful comments and suggestions. Many thanks also to Mrdjan Mladjan, from University Pompeu Fabra in Barcelona, who offered valuable advice on the content and the English of this paper.

Notes

(1) If X = (x_{ij}) and Y = (y_{ij}) are both $m \times n$ matrices, then $X \odot Y$ is the $m \times n$ matrix containing elementwise products ($x_{ij} \times y_{ij}$) (Bauwens et al., 2006).

(2) If v is a vector of dimensions m then $diag(v) = I_m \times v$ (Bauwens et al., 2006).

(3) The algorithm used to obtain parameter estimates is Newton-Raphson algorithm. This algorithm uses the matrix of analytic second derivatives of the log likelihood in forming iteration updates and in computing the estimated covariance matrix of the coefficients. Candidate values for the parameters $\theta_{(1)}$ may be obtained using the method of Newton- Raphson by linearizing the first order conditions $\partial F / \partial \theta$ at the current parameter values $\theta_{(i)}$:

$$g_{(i)} + H_{(i)}\left(\theta_{(i+1)} - \theta_{(i)}\right) = 0$$
$$\theta_{(i+1)} = \theta_{(i)} - H_{(i)}^{-1} g_{(i)}$$

where g is the gradient vector $\partial F / \partial \theta$, and H is the Hessian matrix $\partial^2 F / \partial \theta^2$. If the function is quadratic, Newton-Raphson will find the maximum in a single iteration. If the function is not quadratic, the success of the algorithm will depend on how well a local quadratic approximation captures the shape of the function (EViews 5 User's Guide).

References

Altay-Salih, A., Pinar M., Leyffer S., "Constrained Nonlinear Programming for Volatility Estimation with GARCH Models", August 11, 2003, http://www.siam.org/journals/sirev/45-3/40011.html

Baur, D. October 2004. "A Flexible Dynamic Correlation Model". Working Paper. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=377722

Bauwens, L., "MGARCH-slides-LB-print", November 2005, Universit ´e catholique de Louvain

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

Bauwens, L., Laurent S., "A New Class of Multivariate Skew Densities, with Application to GARCH Models", April 2002, revised April 30, 2004, http://www.core.ucl.ac.be/econometrics/Bauwens/papers/2002-20-JBESfinal.pdf

Brooks, Ch. (2002). *Introductory Econometrics for Finance*, Cambridge University Press

Brooks, Ch., Burke, S., Persand, G., "Multivariate GARCH Models: Software Choice and Estimation Issues", *Journal of Applied Econometrics*, John Wiley & Sons, Ltd., Vol. 18(6), April 2003, pp. 725-734

Cízek, P., Wolfgang, H., Rafa, W., "Statistical Tools for Finance and Insurance", March 3, 2005, http://www.xplore-stat.de/tutorials/

De Goeij, P., Marquering, W., "Modeling the Conditional Covariance Between Stock and Bond Returns: A Multivariate GARCH Approach", *Journal of Financial Econometrics,* Vol. 2, 2004, No. 4, pp. 531-564

Deistler, M., "Multivariate Modeling of Financial Time Series", Dissertation, Mai 2006, http://publik.tuwien.ac.at/files/pub-tm_4874.pdf

Engle, R., "Dynamic conditional correlation – A simple class of multivariate GARCH models", Forthcoming *Journal of Business and Economic Statistics,* Vol. 17, No. 5, 2002

Engle, R.F., Sheppard, K., "Theoretical and Empirical properties of Dynamic Conditional Correlation Multivariate GARCH", September 7, 2001, http://ideas.repec.org/r/nbr/nberwo/8554.html

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

Hafner, Ch., Herwartz, H. "Volatility impulse response for multivariate GARCH models: An exchange rate illustration", *Journal of International Money and Finance*, 2006, pp. 1-22

Minović, Jelena, "Univariate GARCH models: theoretical survey and application", *BALCOR 07*, Proceedings of International Conference, 2007, Zlatibor, Serbia.

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

Pang, Siu, Au-Yeung, Gannon, G., "Regulatory change and structural effects in HSIF and HSI volatility", October 2002, http://wwwdocs.fce. unsw.edu.au/banking/seminar/2002/Gannon.pdf

Pearce, D.W. (1992). *The MIT Dictonary of modern economics,* MIT Press

Pourahmadi, M. "Joint mean-covariance with applications to longitudinal data: Unconstrained parameterization", *Biometrica* 86, 3: 1999, pp. 677-690

Schmidth, A. (2005). *Quantitative Finance for Physicists-An Introduction*, Academic Press

Tsay, R.S. (2005). *Analysis of Financial Time Series*, Wiley, New Jersey

Tse, Y.K., "A test for constant correlations in a multivariate GARCH model", *Journal of Econometrics* 98: 2000, pp. 107-127

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

Wang, M., Yao, Q., "Modelling multivariate volatilities: An ad hoc method", October 18, 2005, http://stats.lse.ac.uk/q.yao/qyao.links/paper/wy04.pdf

Quantitative Micro Software. EViews 5 User's Guide

Official the Belgrade stock exchange web site, http://www.belex.rs

Brokerage web site, http://www.sinteza.net/

http://www2.gsu.edu/~mkteer/continuo.html#elliptical

Appendix A

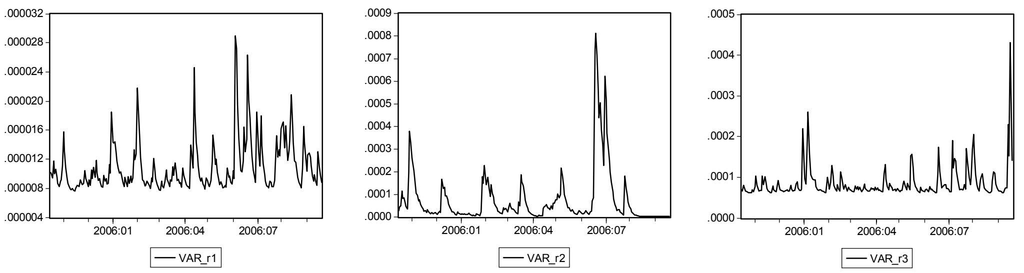
Estimated parameters of trivariate BEKK, DVEC and CCC models. This table contains the coefficients, standard errors, z-statistics, log-likelihood and information criteria for trivariate BEKK, DVEC and CCC model

Table A.1

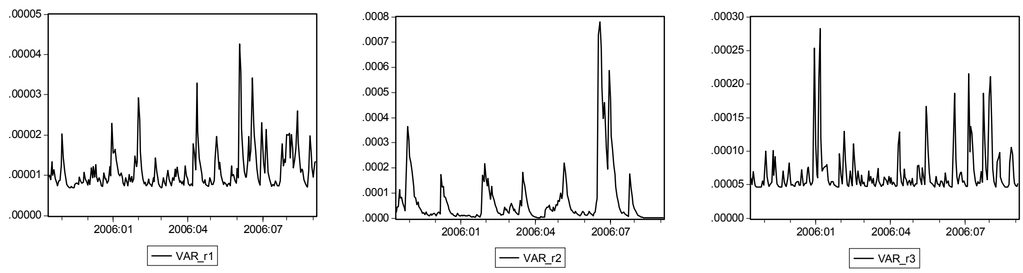
	BEKK			DVEC			CCC		
	Coeff.	S.E.	z-Stat	Coeff.	S.E.	z-Stat	Coeff.	S.E.	z-Stat
MU(1)	-0.0003	0.0002	-1.4536	-0.0001	0.0002	-0.3216	-0.0001	0.0002	-0.6120
MU(2)	-0.0006	0.0001	-8.1946	-0.0003	0.0002	-1.6585	-0.0006	0.0001	-8.4620
MU(3)	-0.0005	0.0007	-0.7305	-0.0003	0.0006	-0.5244	-0.0003	0.0006	-0.5453
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

	BEKK			DVEC			CCC		
	Coeff.	S.E.	z-Stat	Coeff.	S.E.	z-Stat	Coeff.	S.E.	z-Stat
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**	

BEKK



DVEC



CCC

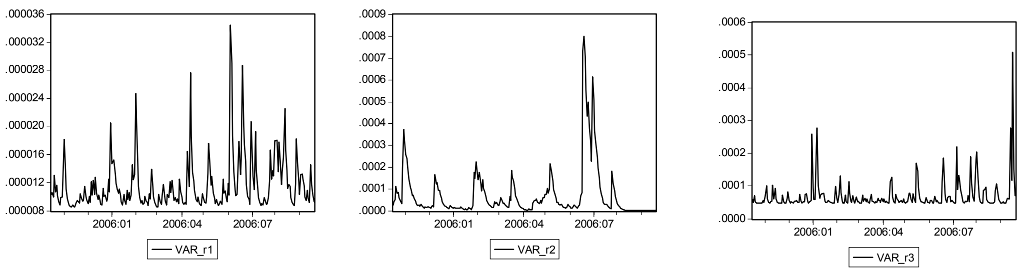


Figure A.1. *Estimated conditional variances of daily log returns on BELEX15 index (var_r$_1$), Hemofarm stock (var_r$_2$) and Energoprojekt stock (var_r$_3$), respectively in the trivariate BEKK, DVEC and CCC models*