Jacek Osiewalski*, Mateusz Pipień**

BAYESIAN ANALYSIS OF DYNAMIC CONDITIONAL CORRELATION USING BIVARIATE GARCH MODELS***

Abstract. Multivariate ARCH-type specifications provide a theoretically promising framework for analyses of correlation among financial instruments because they can model time-varying conditional covariance matrices. However, general VechGARCH models are too heavily parameterized and, thus, impractical for more than 2- or 3-dimensional vector time series. A simple t-BEKK(1,1) specification seems a good compromise between parsimony and generality. Unfortunately, Bollerslev's constant conditional correlation (CCC) model cannot be nested within VECH or BEKK GARCH structures. Recently, Engle (2002) proposed a parsimoniously parameterized generalization of the CCC model; this dynamic conditional correlation (DCC) specification may outperform many older multivariate GARCH models. In this paper we consider Bayesian analysis of the conditional correlation coefficient within different bivariate GARCH models, which are compared using Bayes factors and posterior odds. For daily growth rates of PLN/USD and PLN/DEM (6.02.1996-28.12.2001) we show that the t-BEKK(1, 1) specification fits the bivariate series much better than DCC models, but the posterior means of conditional correlation coefficients obtained within different models are very highly correlated.

Keywords: model comparison, Bayes factors, multivariate GARCH processes, BEKK models, DCC models, exchange rates.

JEL Classification: C11, C32, C52.

1. INTRODUCTION

Appropriate statistical modeling of correlation among financial instruments is crucial for any application of portfolio analysis and for empirical research on dependencies between financial markets. Multivariate

* Prof. dr hab. (Professor), Department of Econometrics, Cracow University of Economics.

** Dr (Ph.D., Assistant Professor), Department of Econometrics, Cracow University of Economics.

[213]

ARCH-type specifications provide a theoretically promising framework as they can model time-varying conditional covariance matrices. However, general VechGARCH models presented by Engle and Kroner (1995) and Gourieroux (1997, Chapter 6) are too heavily parameterized. The number of free parameters of multivariate ARCH-type models can increase very fast as the dimension k of the vector time series grows. In the general version of the k-variate VechGARCH(p, q) (or VECH(p, q)) model, this number is a fourth order polynomial of k, making even VECH(1, 1) impractical for $k > 2$. Thus, within ARCH-type models, interest focuses on restricted GARCH specifications or on factor ARCH models (e.g. Diebold and Nerlove 1989, King, et al. 1994), Gourieroux 1997, Chapter 8). However, factor GARCH models can be not only difficult to estimate (due to the presence of latent variables), but also inadequate (inflexible in modeling complicated dynamics of the conditional covariance matrix). A simple t-BEKK(1, 1) model, based on specifications proposed by Baba et al. (1989) and corresponding to certain non-linear restrictions in t-VECH(1, 1) – cf. Osiewalski and Pipień (2002), seems a good compromise between parsimony and generality. However, this BEKK(1, 1) model inherits some inflexibility of the VECH(1, 1) covariance structure; namely, Bollerslev's (1990) constant conditional correlation (CCC) model cannot be nested within VECH or BEKK GARCH structures. Recently, Engle (2002) proposed a parsimoniously parameterized generalization of the CCC model; his dynamic conditional correlation (DCC) specification may outperform many older multivariate GARCH models. Hence, it is of great interest to empirically check the explanatory power of DCC models.

In order to illustrate a formal Bayesian comparison of various bivariate ARCH-type models through their Bayes factors, Osiewalski and Pipień (2004a, b) used two exchange rates that were most important for the Polish economy till the end of 2001, namely the zloty (PLN) values of the US dollar and German mark. The data consisted of the official daily exchange rates of the National Bank of Poland (NBP fixing rates), starting from February 1, 1996. By restricting to only bivariate VAR(1) models with GARCH(1, 1) disturbances, it was possible to estimate unparsimoniously parameterized specifications, such as VechGARCH models. Those first comparisons focused on older multivariate GARCH structures, proposed prior to 2001. Thus, the class of models did not contain more recent models proposed by Tse and Tsui (2002), van der Weide (2002) and, in particular, the DCC models of Engle (2002).

The main result of Osiewalski and Pipień (2004a, b) is that the simple t-BEKK(1, 1) model wins model comparison. In its k-variate version, it has $O(k^2)$ free parameters, much less than the k-variate general version of VECH(1, 1), requiring $O(k^4)$ free parameters. In this paper we focus on

$O(k^2)$ specifications, in particular on variants of Engle's DCC structures. Our aim is to compare these models, which were not considered in our previous Bayesian works, to the winner from those studies. We show that, for our data set, the unrestricted t-BEKK(1, 1) model describes the time-varying conditional covariance matrix still much, much better than quite sophisticated (and very elegant) DCC structures, specially designed to model dynamic conditional correlation.

In Section 2 we briefly present our Bayesian statistical methodology and the numerical tools we use. Section 3 is mainly devoted to the description of the competing model specifications and the results of their formal comparison using Bayes factors. In Section 4 the sequences of estimates of the conditional correlation coefficients (representing dynamics of the relationship between our two series) and standard deviations (measuring volatility of each series) are presented and compared.

2. STATISTICAL METHODOLOGY AND NUMERICAL TOOLS

We consider several competing parametric Bayesian models for the same observation matrix y. The i-th Bayesian model (M_i) is characterised by the joint density function:

(1) $p(y,\theta_{(i)}|M_i,y_{(0)}) = p(y|M_i,\theta_{(i)},y_{(0)})p(\theta_{(i)}|M_i)$ $(i = 1, \ldots, m)$,

where $y_{(0)}$ denotes initial conditions and $p(y|M_i,\theta_{(i)},y_{(0)})$, $p(\theta_{(i)}|M_i)$ are the sampling density function and the prior density function under M_i, respectively. $\theta_{(i)}$, the parameter vector in M_i, groups parameters common to all m models and model-specific parameters. For the purposes of inference within M_i and model comparison, we use the obvious decomposition

$$p(y,\theta_{(i)}|M_i,y_{(0)}) = p(y|M_i,y_{(0)})p(\theta_{(i)}|y,M_i,y_{(0)}),$$

where $p(\theta_{(i)}|y,M_i,y_{(0)})$ is the posterior density function in M_i and

$$p(y|M_i,y_{(0)}) = \int_{\theta_i} p(y|M_i,\theta_{(i)},y_{(0)})p(\theta_{(i)})d\theta_{(i)}$$

is the marginal data density in the i-th Bayesian model. Competing models are compared pair-wise through the Bayes factor $B_{ij} = $ $= p(y|M_i,y_{(0)})/p(y|M_j,y_{(0)})$, which, together with the prior odds ratio $P(M_i)/P(M_j)$, determines the posterior odds of M_i against M_j:

(2)
$$\frac{P(M_i|y,y_{(0)})}{P(M_j|y,y_{(0)})} = \frac{P(M_i)}{P(M_j)} B_{ij},$$

where $P(M_h)$ and $P(M_h|y,y_{(0)})$ are, respectively, the prior and posterior probability of M_h (e.g. O'Hagan 1994). The crucial role of the Bayes factor in model comparison means that computing marginal data densities under competing models is the main numerical task. Direct evaluation of the integral defining the marginal data density (as well as of integrals related to posterior inferences) – through either numerical quadratures or Monte Carlo sampling from the prior density – is not efficient (or even not feasible) when the dimension of the parameter space is as high as in the models considered in this paper. Thus we have to resort to other numerical tools, based on good exploration of the parameter space through sampling from the posterior. Here we use Metropolis-Hastings (M-H) Markov chains (e.g. O'Hagan 1994), Gamerman (1997).

Using simple identities, we can write the marginal data density in the form

(3)
$$p(y|M_i,y_{(0)}) = \left\{ \int_{\theta_i} [p(y|M_i,\theta_{(i)},y_{(0)})]^{-1} dP(\theta_{(i)}|M_i,y,y_{(0)}) \right\}^{-1},$$

where $P(\theta_{(i)}|M_i,y,y_{(0)})$ denotes the posterior cumulative distribution function. This formula is the basis of the method by Newton and Raftery (1994), which approximates the marginal data density by the harmonic mean of the values $p(y|M_i,\theta_{(i)},y_{(0)})$, calculated for the observed y and for $\theta_{(i)}$ drawn from the posterior distribution. The N-R harmonic mean estimator is consistent, but without finite asymptotic variance. Despite this serious theoretical weakness, the N-R estimator (very easy to compute) was quite stable for all our models; (cf. Osiewalski and Pipień 2004a for more discussion of computational aspects).

In order to sample from the posterior distribution in a model with the parameter vector θ, we use a sequential version of the M-H algorithm, where the proposal density $q(\theta|\theta^{(m-1)})$ for the next value of θ given the previous draw $\theta^{(m-1)}$ is proportional to $f_S(\theta|3,\theta^{(m-1)},C)$, a Student t density with 3 degrees of freedom, mean $\theta^{(m-1)}$ and a fixed covariance matrix C (approximating the posterior covariance matrix). This Student-t density (symmetric in θ and $\theta^{(m-1)}$) is truncated by the inequality restrictions described in Section 3, i.e.

(4)
$$q(\theta|\theta^{(m-1)}) = a_q(\theta^{(m-1)})f_S(\theta|3,\theta^{(m-1)},C)I(\theta \in \Theta),$$

$$a^q(\theta^{(m-1)}) = \left[\int_{\Theta_i} f_S(\theta|3,\theta^{(m-1)},C)d\theta \right]^{-1}.$$

This leads to the M-H Markov chain with the following acceptance probability:

(5) $\qquad \alpha(\theta;\theta^{(m-1)}) = \min\{(g_y(\theta)a_q(\theta))/(g_y(\theta^{(m-1)})a_q(\theta^{(m-1)})), \; 1\},$

where $g_y(.)$ denotes the kernel of the posterior density. Thus, given the previous state of the chain, $\theta^{(m-1)}$, the current state $\theta^{(m)}$ is equal to the candidate value θ^* (drawn from the truncated Student-t distribution discussed above) with probability $\alpha(\theta^*;\theta^{(m-1)})$ or $\theta^{(m)} = \theta^{(m-1)}$ with probability $1 - \alpha(\theta^*;\theta^{lm-1})$. Our results, presented in next sections, are based on 500 000 states of the Markov chain, generated after 10 000 burnt-in states.

3. THE DATA AND COMPETING MODELS

In order to compare competing bivariate ARCH-type specifications we use the growth rates of PLN/USD and PLN/DEM. Our original data set consists of 1485 daily observations on the exchange rates themselves, PLN/USD (x_{1t}) and PLN/DEM (x_{2t}). It covers the period from 1.02.1996 till 28.12.2001. The first three observations from 1996 (February 1, 2 and 5) are used to construct initial conditions. Thus T, the length of the modeled vector time series of daily growth rates of x_{1t} and x_{2t} is equal to 1482.

We denote our modeled bivariate observations as $y_t = (y_{1t}, y_{2t})'$, where y_{1t} is the daily growth (or return) rate of the PLN value of US dollar and y_{2t} is the daily growth (or return) rate of the PLN value of German mark, both expressed in percentage points and obtained from the daily exchange rates x_{it} ($i = 1, 2$) by the formula $y_{it} = 100 \ln(x_{it}/x_{i,t-1})$. Osiewalski and Pipień (2004a) used only a short part of this bivariate series, till the end of 1997 ($T = 475$). Now we base our results on all $T = 1482$ observations, as Osiewalski and Pipień (2004b), but we do not use any exogenous variables in the conditional mean specification. Thus we stay within the pure VAR-GARCH framework, like Osiewalski and Pipień (2004a).

We model the data using the basic VAR(1) framework:

$$y_t - \delta = R(y_{i-1} - \delta) + \varepsilon_t$$

with the error described by competing bivariate GARCH specifications. More specifically,

(7) $\quad \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} - \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} \left\{ \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} - \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} \right\} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}, \qquad t = 1, \, ..., \, T.$

The elements of δ and R are common parameters, which we treat as a priori independent of all other (mainly model-specific) parameters and assume for them the multivariate standardized normal prior $N(0, I_6)$, truncated by the restriction that all eigenvalues of R lie inside the unit circle. We assume that the conditional distribution of ε_t (given its past, ψ_{t-1}) is Student-t with zero location vector, inverse precision matrix H_t and unknown degrees of freedom $v > 2$, i.e.

$$(8) \qquad \varepsilon_t|\psi_{t-1} \sim t(0_{(2x1)}, H_t, v), \qquad H_t = \begin{bmatrix} h_{11,t} & h_{12,t} \\ h_{12,t} & h_{22,t} \end{bmatrix}.$$

As regards initial conditions for H_t, we take $H_0 = h_0 I_2$ and treat h_0 as an additional parameter. We assume prior independence for v, h_0 (which are common) and the remaining parameters; v follows the exponential distribution with mean 10, $Exp(10)$, truncated by the condition $v > 2$; h_0 has the exponential prior with mean 1, $Exp(1)$.

The conditional covariance matrix of ε_t given ψ_{t-1} is $(v-2)^{-1}vH_t$. Competing bivariate GARCH models are defined by imposing different structures on H_t. That is, model-specific parameters are the ones describing H_t in a given model. The sampling density function in each model is always the product of T conditional bivariate Student-t densities (for $y_{(t)}$) with v degrees of freedom, mean $\delta + R(y_{t-1} - \delta)$ and covariance matrix $(v-2)^{-1}vH_t$.

The first specification considered here is the very parsimonious constant conditional correlation (CCC) model of Bollerslev (1990); it imposes the following structure on H_t:

$$(9) \qquad h_{11,t} = a_{10} + a_{11}\varepsilon_{1,t-1}^2 + b_{11}h_{11,t-1},$$

$$h_{22,t} = a_{20} + a_{22}\varepsilon_{2,t-1}^2 + b_{22}h_{22,t-1},$$

$$h_{12,t} = \rho_{12}\sqrt{h_{11,t}h_{22,t}},$$

where ρ_{12} is the time-invariant conditional correlation coefficient. This simple structure of H_t amounts to modeling each conditional variance by a different GARCH(1, 1) process and making the conditional covariance a simple function of the variances. In its k-variate version, the CCC model describes ε_t using only $2 + 3k + k(k-1)/2$ free parameters; so we have 9 parameters when $k = 2$. For the model-specific parameters we take the following priors:

(10) $a_{10} \sim Exp(1), \quad a_{20} \sim Exp(1), \quad (a_{11}, a_{22}, b_{11}, b_{22}) \sim U([0, 1]^4),$

$$\rho_{12} \sim U([-1, 1]),$$

where $U(A)$ denotes the uniform distribution over A. Osiewalski and Pipień (2004a, b) show that, for our data, the CCC model is inadequate – it is much worse than heavily parameterized VechGARCH specifications and than more parsimonious BEKK structures, which all assume time-varying conditional correlations. It seems that modeling dynamic correlation with almost as few parameters as in the CCC model would be the most welcome solution.

The simple CCC specification (under conditional normality, i.e. with $\nu = +\infty$) has been generalized by Engle (2002) in such a way as to make conditional correlations fully dynamic, keeping the conditional covariance formula basically unchanged. Engle's dynamic conditional correlation (DCC) models describe the diagonal elements of H_t in the same way as in CCC, but assume that

(11) $h_{12,t} = \rho_{12,t}\sqrt{h_{11,t} h_{22,t}},$

where $\rho_{12,t}$ is the time-varying conditional correlation coefficient, modeled as

$$\rho_{12,t} = q_{12,t}/\sqrt{q_{11,t} q_{22,t}},$$

with $q_{ij,t}$'s being entries of a symmetric positive definite matrix Q_t of the same order as the dimension of ε_t. A simple specification for Q_t, considered in Engle (2002), assumes that

(12) $Q_t = (1 - \alpha - \beta)S + \alpha \xi_{t-1} \xi'_{t-1} + \beta Q_{t-1}$

where α and β are nonnegative scalar parameters $(\alpha + \beta < 1)$, ξ_t is the vector of standardized errors and S is their unconditional correlation matrix. In the case of our bivariate conditionally Student-t specification, we keep Engle's basic structure and define S as a square matrix with ones on the diagonal and $s_{12} = s_{21} = \rho_{12}$, an unknown parameter from the interval $(-1, 1)$; this assures positive definiteness of S and Q_t. Also, in our case

(13) $\xi_{it} = \varepsilon_{it}\sqrt{(\nu - 2)/(\nu h_{ii,t})} \quad (i = 1, 2).$

Thus, our second specification (called DCC0) generalizes the conditionally normal basic structure proposed by Engle (2002) to the Student-t conditional error distribution. The initial condition for Q_t is $Q_0 = q_0 I_2$, where q_0 is a free parameter. In its k-variate version, DCC0 describes ε_t using

$5 + 3k + k(k-1)/2$ free parameters – only three (q_0, α and β) more than the CCC model (irrespective of k). Of course, CCC corresponds to $\alpha = \beta = 0$, so it is nested in DCC0. We follow the exact Bayesian approach, which is fully feasible in the bivariate case. Thus we do not use the approximate two-step estimation procedure suggested by Engle (2002). The three new parameters are assumed independent *a priori* of the remaining ones. The prior for q_0 is $Exp(1)$, while the one for (α, β) is uniform over the unit simplex.

The third model (called DCC1) is also of the DCC form, but the specification of Q_t is different. The previous period error terms are not standardized and there are less restrictions:

$$(14) \qquad Q_t = V + \alpha \varepsilon_{t-1} \varepsilon'_{t-1} + \beta Q_{t-1},$$

V consists of $v_{11} \sim Exp(1), v_{22} \sim Exp(1)$ and $v_{12} = v_{21} = \rho_{12}\sqrt{v_{11}v_{22}}$ with $\rho_{12} \sim U([-1,1])$, so V is positive definite with prior probability 1, and $(\alpha, \beta) \sim U([0,1]^2)$.

The fourth model (DCC2) generalizes the structure of Q_t by replacing the two scalar parameters (α and β) by two symmetric, nonnegative definite matrices (A and B):

$$(15) \qquad Q_t = V + A \circ \varepsilon_{t-1} \varepsilon'_{t-1} + B \circ Q_{t-1},$$

where $C \circ D$ is the Hadamard product of two matrices of the same size (i.e., the element-by-element multiplication). This equation resembles (24) in Engle (2002), but (as in DCC1) the previous period error terms are not standardized and there are no restrictions on $A + B$. Our Bayesian DCC2 specification uses the same V as in DCC1 and assumes that A consists of: $\alpha_{12} = \alpha_{21} = \alpha_r\sqrt{\alpha_{11}\alpha_{22}}$, $\alpha_{11}, \alpha_{22} \sim U([0,1])$ and $\alpha_r \sim U([-1,1])$; similarly for β_{ij} in B: $\beta_{12} = \beta_{21} = \beta_r\sqrt{\beta_{11}\beta_{22}}$, $\beta_{11}, \beta_{22} \sim U([0,1])$ and $\beta_r \sim U([-1,1])$. So Q_t is positive definite with prior probability 1. In its k-variate version DCC2 has $3 + 3\{2k + k(k-1)/2\}$ free parameters that enter the conditional distribution of ε_t (18 for $k = 2$).

As we have already noted, the t-CCC specification was strongly rejected by our data when compared to VECH and BEKK bivariate t-GARCH structures. Now we show the results of our Bayesian comparison between the t-CCC and each t-DCC model. The decimal logarithm of the Bayes factor in favor of the DCC0 model is 46.60, in favor of the DCC1 specification is 45.15, while for the DCC2 structure we obtain 46.65. All three DCC models are about 45 orders of magnitude better (i.e., more probable a posteriori under equal prior model probabilities) than the CCC

model ! High and almost equal values of the Bayes factors for DCC0 and DCC2 indicate that these two models describe the data equally well. The price we pay for not using standardized residuals in Q_t amounts to estimating more parameters in DCC2 than in DCC0. The parameterization in DCC1 seems not rich enough, but the difference between the Bayes factors (of this model against CCC and of DCC0 against CCC) is not large when we take into consideration sensitivity with respect to the prior distribution and numerical stability issues.

The results obtained for the DCC models seem encouraging. However, our previous results (cf. Osiewalski and Pipień 2004b) show that that the decimal log of the Bayes factor for a simple t-BEKK(1, 1) model (against CCC) is even much higher, equal to 64.13. This BEKK specification is defined by the following structure of H_t:

(16)

$$H_t = \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}' (\varepsilon_{t-1}\varepsilon_{t-1}') \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} + \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}' H_{t-1} \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix},$$

i.e. $H_t = A + B'\varepsilon_{t-1}\varepsilon_{t-1}'B + C'H_{t-1}C$.

The parameters of this structure have the following prior distributions:

$$a_{11} \sim Exp(1), \quad a_{22} \sim Exp(1), \quad a_{12} \sim N(0,1), \quad b_{11} \sim N(0.5,1),$$

$$b_{12} \sim N(0,1), \quad b_{21} \sim N(0,1), \quad b_{22} \sim N(0.5,1), \quad c_{11} \sim N(0.5,1),$$

$$c_{12} \sim N(0,1), \quad c_{21} \sim N(0,1), \quad c_{22} \sim N(0.5,1),$$

which are truncated by the restrictions of positive semi-definiteness of the symmetric (2×2) matrix A and stability of the general (2×2) matrix C (all eigenvalues of C lie inside the unit circle). Also, the conditions: $b_{11} > 0$ and $c_{11} > 0$ are imposed in order to guarantee identifiability, since B and $-B$ as well as C and $-C$ lead to the same H_t, and thus are observationally equivalent. In the k-variate version, our t-BEKK(1, 1) model describes the conditional distribution of ε_t (given its past) using $2 + k(k+1)/2 + 2k^2$ free parameters (13 for $k = 2$).

The success of the t-BEKK(1,1) model (its clear superiority over the DCC models and other specifications in explaining the time-varying conditional covariance structure) suggests further search for more parsimonious special cases of t-BEKK(1,1) that would hopefully keep its explanatory power. Some models, like t-BEKK(1,0) that assumes $C = 0$, have already been tried (cf. Osiewalski and Pipień 2004b). The decimal log of the Bayes factor of t-BEKK(1,0) relative to t-CCC is –23.71 (!). Thus, the BEKK(1, 0)

model (with an ARCH(1) structure only) is even much worse the than the CCC specification, so it will not be discussed further. Here we propose a simple "scalar t-BEKK$(1,1)$" structure, which amounts to assuming $B = bI_2$ and $C = cI_2$, where b and c are independent scalar parameters with $N(0.5, 1)$ prior distributions, truncated by the restrictions: $b > 0$ and $0 < c < 1$. So we consider

$$(17) \qquad H_t = A + b^2 \varepsilon_{t-1}\varepsilon'_{t-1} + c^2 H_{t-1},$$

which is much simpler than DCC1 (it uses the DCC1 structure of Q_t at the level of H_t). The decimal log of the Bayes factor of this scalar t-BEKK$(1,1)$ relative to t-CCC is 48.75, indicating that this restricted, extremely simple BEKK formulation can compete in dynamic correlation modeling with more sophisticated DCC structures, designed for this purpose. In fact, our scalar BEKK is about two or three orders of magnitude more probable *a posteriori* than DCC0 or DCC2 (assuming equal prior model probabilities). Of course, the unrestricted BEKK specification undoubtedly wins our model comparison for the analyzed data set, being about 15 orders of magnitude better than the second best specification.

All our results, those presented previously in Osiewalski and Pipień (2004b) and the new ones given here, indicate that the growth rates of PLN/USD and PLN/DEM strongly reject the constant conditional correlation hypothesis. These exchange rates form a bivariate time series with strong correlation dynamics, where BEKK models can (and should) be used. The fact that BEKK models do not nest the CCC case is not a problem for the Bayesian approach, which can deal with testing non-nested specifications using Bayes factors and posterior model probabilities. The results of this section are summarised in Table 1, where we rank the models by the increasing value of the decimal logarithm of the Bayes factor of BEKK$(1,1)$ against the alternative models.

Table 1. Logs of Bayes factors in favor of t-BEKK$(1,1)$ and average posterior means of $\rho_{12,t}$

Model	Number of parameters	Rank	$\log_{10}(B_{1t})$	Average $E(\rho_{12,t}|y)$
M_1, BEKK$(1,1)$	19	1	**0**	0.162
M_2, scalar BEKK$(1,1)$	13	2	**15.38**	0.122
M_3, DCC2	24	3–4	**17.48**	0.132
M_4, DCC0	18	3–4	**17.53**	0.132
M_5, DCC1	20	5	**18.99**	0.132
M_6, CCC	15	6	**64.13**	0.237

4. POSTERIOR INFERENCE ON CONDITIONAL CORRELATION COEFFICIENTS AND VOLATILITIES

In this section we compare main results for individual volatilities and the dynamic correlation structure, obtained within each model. It is important to know whether models that have different explanatory power describe this structure in a similar way.

The plots of the sampling conditional correlation coefficients $\rho_{12,t}$ (for each $t = 1, ..., T$; $T = 1482$) are presented in Figure 1, where we draw two lines: the upper one representing the posterior mean plus two posterior standard deviations and the lower one – the posterior mean minus two posterior standard deviations. We focus on typical patterns, so only two models are represented in Figure 1. It is clear that constancy of conditional correlations, which are quite tightly concentrated around their abruptly changing posterior means, is not supported by the data. This explains why the CCC model receives negligible posterior probability when compared to DCC or BEKK specifications. The last column of Table 1 presents time averages for the sequences of posterior means of the conditional correlation coefficient in each model, while Table 2 gives the empirical correlation coefficients between these sequences (the numbers above the diagonal).

Table 2. Correlation coefficients between the posterior means of the conditional correlations (upper part) and covariances (lower part)

Specification	BEKK	Scalar BEKK	DCC2	DCC0	DCC1
BEKK	X	0.9113	0.9180	0.9094	0.9125
Scalar BEKK	0.9152	x	0.9837	0.9810	0.9826
DCC2	0.9299	0.9987	x	0.9767	0.9959
DCC0	0.9172	0.9982	0.9982	x	0.9643
DCC1	0.9192	0.9996	0.9993	0.9983	X

These results show that the models of comparable explanatory power lead to almost the same inference on the dynamics of conditional correlation. For the scalar BEKK and all three DCC models, averages of $E(\rho_{12,t}|y)$ ($t = 1, ..., T$) are about $0.12 - 0.13$ and the empirical correlation coefficients between pairs of $E(\rho_{12,t}|y)$ sequences are bigger than 0.96. However, the sequence of $E(\rho_{12,t}|y)$ coming from the unrestricted BEKK is slightly less correlated with the others (about 0.91) and has a somewhat higher average (0.16). Also, the plot obtained for the unrestricted BEKK looks somewhat different (cf. Figure 1).
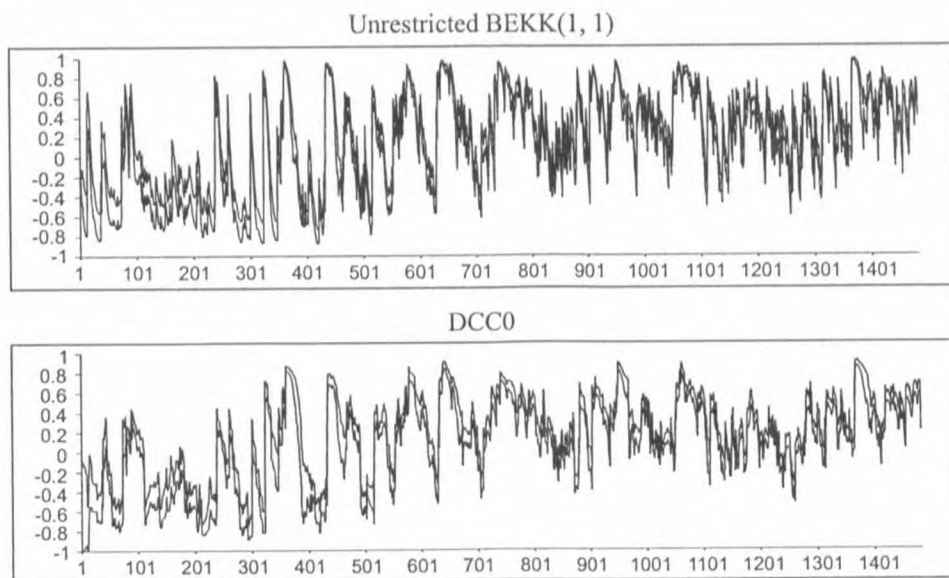
Unrestricted BEKK(1, 1)

DCC0

Fig. 1. Conditional correlations (posterior mean ± 2 standard deviations)

Very similar results (as for the conditional correlation $\rho_{12,t}$) have been obtained for the sequences of the posterior means of the conditional covariance $h_{12,t}$; empirical correlations are given in Table 2 (the numbers below the diagonal).

Individual time-varying volatility of each time series is measured by the conditional standard deviation $\sqrt{(v-2)^{-1}vh_{ii,t}}$ ($i = 1, 2$). The sequences of 1482 point estimates, obtained by inserting the posterior means of the model parameters, are plotted in Figure 2 for two models. These estimates exhibit the same dynamic pattern for all models of the same explanatory power (scalar BEKK, DCC2, DCC0, DCC1) – the empirical correlation coefficients (Table 3) are basically equal to 1. The results obtained in CCC and all DCC are also highly correlated. The empirical correlation coefficients are somewhat lower (especially for PLN/DEM) when we compare the unrestricted BEKK specification to the remaining models. Thus the best model leads to slightly different inference on volatility. As regards time averages of the sequences of estimated in-sample volatilities, they are almost the same in all models (including CCC).

Our conclusion is that inferences from the best fitting model can be approximated by the scalar BEKK or DCC specifications. Since the scalar BEKK model is the simplest, but it does not nest the CCC case, one should estimate and compare these two non-nested models. This seems a feasible strategy even for k-variate time series with $k > 2$.

Conditional standard deviations for PLN/USD
unrestricted BEKK(1, 1)

Conditional standard deviations for PLN/DEM
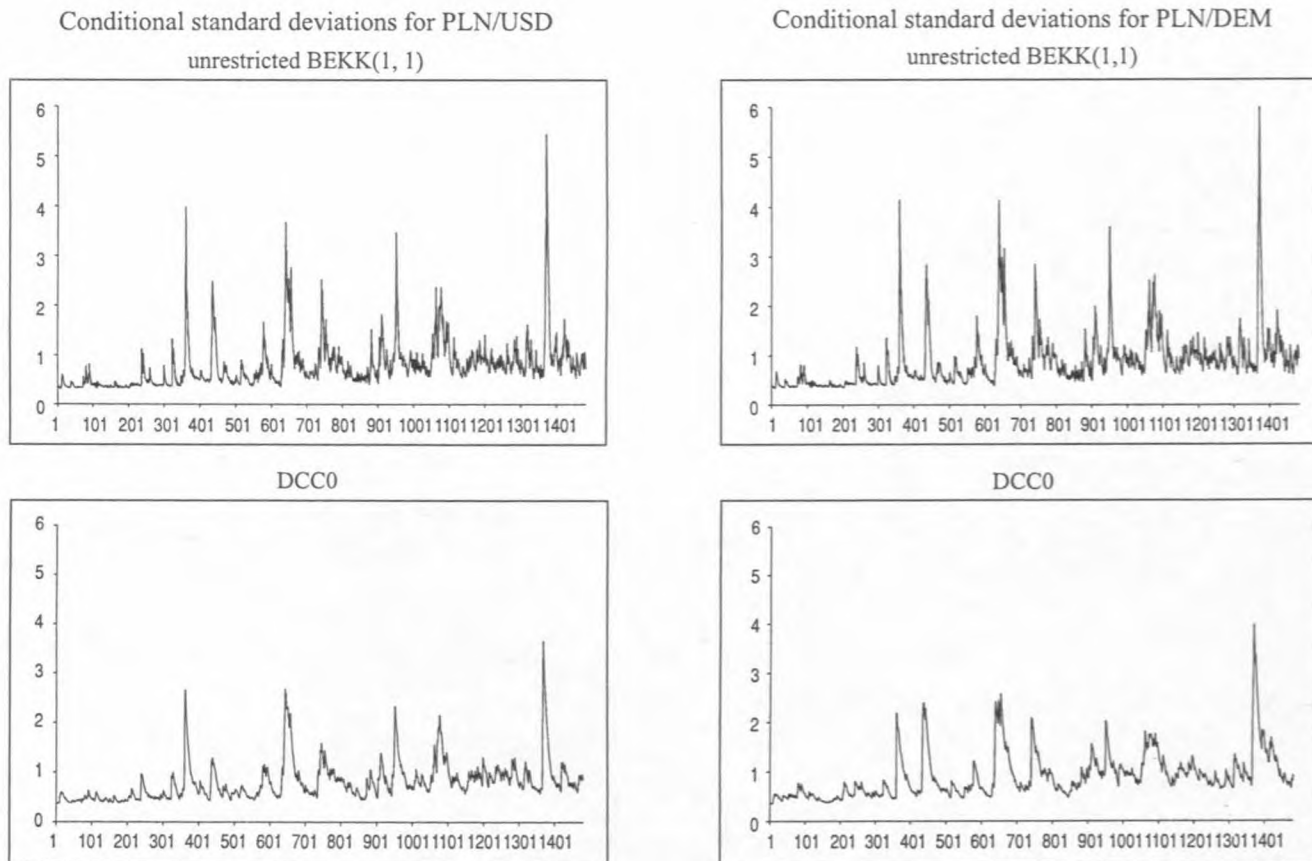unrestricted BEKK(1,1)

DCC0

DCC0

Fig. 2. Point estimates of the conditional standard deviations in two main models

Table 3. Correlation coefficients between the estimates of the conditional standard deviations
for PLN/USD (upper part) and for PLN/DEM (lower part)

Specification	BEKK	Scalar BEKK	DCC2	DCC0	DCC1	CCC
BEKK	X	0.9217	0.9234	0.9219	0.9233	0.9420
Scalar BEKK	0.8826	x	0.99988	0.99997	0.99988	0.9930
DCC2	*0.8854*	*0.99988*	X	0.99990	0.99999	0.9941
DCC0	*0.8826*	*0.99992*	*0.99989*	x	0.99988	0.9933
DCC1	*0.8860*	*0.99986*	*0.99999*	*0.99986*	x	0.9940
CCC	*0.8828*	*0.99966*	*0.99973*	*0.99974*	*0.99967*	x

REFERENCES

Baba Y., Engle, R. F., Kraft, D. and Kroner, K. (1989), *Multivariate Simultaneous Generalised ARCH*, manuscript, San Diego: University of California.

Bollerslev, T. (1990), "Modelling the Coherence in Short-run Nominal Exchange Rates: A Multivariate Generalised ARCH Model", *Review of Economics and Statistics*, 72, 498–505.

Diebold, F. and Nerlove, M. (1989), "The Dynamic of Exchange Rate Volatility: A Multivariate Latent Factor ARCH Model", *Journal of Applied Econometrics*, 4, 1–22.

Engle, R. (2002), "Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models", *Journal of Business and Economic Statistics*, 20, 339–350.

Engle, R. F. and Kroner, K. F. (1995), "Multivariate Simultaneous Generalised ARCH", *Econometric Theory*, 11, 122–150.

Gamerman, D. (1997), *Markov Chain Monte Carlo. Stochastic Simulation for Bayesian Inference*, London: Chapman and Hall.

Gourieroux, C. (1997), *ARCH Models and Financial Applications*, New York: Springer.

King, M., Sentana, E. and Wadhwani, S. (1994), "Volatility and Links Between National Stock Markets", *Econometrica*, 62, 901–934.

Newton, M. A. and Raftery, A. E. (1994), "Approximate Bayesian Inference by the Weighted Likelihood Bootstrap" (with discussion), *Journal of the Royal Statistical Society*, B 56, 3–48.

O'Hagan, A. (1994), *Bayesian Inference*, London: Edward Arnold.

Osiewalski, J. and Pipień, M. (2002), "Multivariate *t*-GARCH Models – Bayesian Analysis for Exchange Rates". In: Welfe, W. (ed.), *Modelling Economies in Transition – Proceedings of the Sixth AMFET Conference*, Łódź: Absolwent, 151–167.

Osiewalski, J. and Pipień, M., (2004a), "Bayesian Comparison of Bivariate ARCH-type Models for the Main Exchange Rates in Poland", *Journal of Econometrics*, 122 (in press).

Osiewalski, J. and Pipień, M., (2004b), "Bayesian Comparison of Bivariate GARCH Processes. The Role of the Conditional Mean Specification". In: Welfe, A. (ed.), New Directions in Macromodelling, Amsterdam: Elsevier, 173–196 (forthcoming).

Tse, Y. K. and Tsui, A. K. C. (2002) "A Multivariate Generalized Autoregressive Conditional Heteroskedasticity Model with Time-varying Correlations", *Journal of Business and Economic Statistics*, 20, 351–362.

van der Weide, R. (2002), "GO-GARCH: A Multivariate Generalized Orthogonal GARCH Model", *Journal of Applied Econometrics*, 17, 549–564.

Jacek Osiewalski, Mateusz Pipień

BAYESOWSKA ANALIZA DYNAMICZNEJ KORELACJI WARUNKOWEJ
Z WYKORZYSTANIEM DWUWYMIAROWYCH MODELI GARCH

(Streszczenie)

Wielowymiarowe specyfikacje typu ARCH stanowią teoretycznie obiecujące ramy dla analiz skorelowania instrumentów finansowych, ponieważ umożliwiają modelowanie zmiennych w czasie macierzy warunkowych kowariancji. Jednak ogólne modele VechGARCH mają zbyt wiele parametrów, są więc niepraktyczne w przypadku więcej niż 2- lub 3-wymiarowych wektorowych szeregów czasowych. Prosta specyfikacja t-BEKK(1, 1) wydaje się dobrym kompromisem pomiędzy oszczędnością parametryzacji i ogólnością modelu. Niestety model stałych korelacji warunkowych (CCC) Bollersleva nie jest szczególnym przypadkiem struktur VECH czy BEKK. Ostatnio Engle (2002) zaproponował oszczędnie sparametryzowane uogólnienie modelu CCC; ta specyfikacja o dynamicznej korelacji warunkowej (DCC) może zdominować wiele starszych wielowymiarowych modeli GARCH. W artykule rozważamy bayesowską analizę warunkowego współczynnika korelacji w ramach różnych dwuwymiarowych modeli GARCH, które są porównywane przy użyciu czynników Bayesa i ilorazów szans a posteriori. Dla dziennych stóp zmian kursów PLN/USD i PLN/DEM (6.02.1996 – 28.12.2001) wykazuje się, że specyfikacja t-BEKK(1, 1) opisuje dwuwymiarowy szereg czasowy znacznie lepiej niż modele DCC. Jednak wartości oczekiwane a posteriori warunkowych współczynników korelacji, uzyskane w ramach różnych modeli, są bardzo silnie skorelowane.