DYNAMIC ECONOMETRIC MODELS
Vol. 6 – Nicolaus Copernicus University – Toruń – 2004

Jacek Osiewalski, Mateusz Pipień
Cracow University of Economics

Bayesian Comparison of Bivariate GARCH Processes in the Presence of an Exogenous Variable*

1. Introduction

In order to illustrate a formal Bayesian comparison of various bivariate ARCH-type models through their Bayes factors, Osiewalski and Pipień (2003) used two foreign 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). By restricting to only bivariate VAR(1) models with GARCH(1,1) or ARCH(1) disturbances, it was possible to estimate unparsimoniously parameterised specifications, such as general multivariate ARCH-type models, presented by Engle and Kroner (1995) and Gourieroux (1997, ch.6). These models have much more parameters than univariate ARCH and GARCH models, proposed originally by Engle (1982) and Bollerslev (1986), and analysed using the Bayesian approach by Geweke (1989), Kleibergen and Van Dijk (1993), Bauwens and Lubrano (1998), Bauwens, Lubrano and Richard (1999), Osiewalski and Pipień (1999, 2000), Vrontos, Dellaportas, and Politis (2000) and Bos, Mahieu and Van Dijk (2000). 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 ARCH and GARCH specifications or on factor ARCH models; see e.g. Diebold and Nerlove (1989), King, Sentana and Wadhwani (1994) and Gourieroux (1997, ch. 8). Apart from

* This research was supported by a grant from the Cracow University of Economics.

the dimension of the parameter space, Osiewalski and Pipień (2003) considered other aspects of empirical ARCH-type specifications: free conditional covariances versus constant conditional covariances or correlations, direct ARCH versus latent factor ARCH models, conditional Normality versus Student *t* tails, and the ARCH(1) structure versus GARCH(1,1).

In our previous study only pure time-series models for daily data were used, without introducing any variables that would be motivated by theoretical considerations. The relationship: (PLN/USD)/(PLN/DEM) ≈ DEM/USD, which linked two Polish official exchange rates to the international FOREX market, was ignored. In this paper we assume that this approximate relation (in log terms) is a cointegration equation in the sense of Engle and Granger (1987) and that the DEM/USD rate is weakly exogenous in the Bayesian sense of Florens and Mouchart (1985) and Osiewalski and Steel (1996). We build a two-equation conditional model with the error correction mechanism (ECM) and the disturbances following one of the competing bivariate GARCH specifications. The aim of the paper is to check sensitivity of our Bayesian model comparison with respect to the presence of the third (exogenous) exchange rate.

In view of high dimensionality of the parameter spaces and non-standard forms of the posterior densities as well as their full conditionals, we use the Metropolis-Hastings (M–H) algorithm to simulate and explore the posterior distributions. The values of the marginal data densities for each model, which are the main quantities for Bayesian model comparison, are approximated by means of the Newton and Raftery's (1994) estimator, based on the harmonic mean of the likelihood values calculated at M–H draws from the posterior. Both the bivariate framework and a short time series (475 daily observations) enable us to obtain final results for all models rather quickly.

The structure of the paper is as follows. The next section shows the data and the ECM-type model framework for daily growth rates of two exchange rates. Section 3 presents all the models used for the bivariate error term of the basic specification and ranks the models using Bayes factors. Section 4 concludes.

2. The Data and Model Framework

In order to compare competing bivariate ARCH-type specifications we use the growth rates of PLN/USD and PLN/DEM. Osiewalski and Pipień (1999, 2000) modelled these two series using univariate AR(1) – *t*-GARCH(1,1) models. Our original data set consists of 478 daily observations on three exchange rates: PLN/USD (x_{1t}) PLN/DEM (x_{2t}) and DEM/USD (w_t), covering the period from February 1, 1996 till December 31, 1997. The first three observations from 1996 (February 1,2,5) are used to construct initial conditions. Thus T, the length of the modelled vector time series of daily growth rates of x_{1t} and x_{2t} is equal to 475.

We denote our modelled bivariate observations as $y_t=(y_{1t},y_{2t})$`, where y_{1t} is the daily growth (or return) rate of the US dollar and y_{2t} is the daily growth (or return) rate of the German mark, both expressed in percentage points and obtained from the daily exchange rates x_{it}, i=1,2, by the formula $y_{it}=100\ln(x_{it}/x_{it-1})$. We also define $ECM_t = \ln x_{1t}-\ln x_{2t}-\ln w_t$ and $z_t=100\ln(w_t/w_{t-1})$, and model our data using the conditional ECM-type VAR(1) framework:

$$y_t - \delta = R(y_{t-1} - \delta) + \alpha z_t + \lambda \cdot ECM_{t-1} + \varepsilon_t$$

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

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} - \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} \left\{ \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} - \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} \right\} + \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} z_t + \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} ECM_{t-1} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}, \ t=1,...,T. \tag{1}$$

The elements of δ, R, α and λ are common parameters, which we treat as a priori independent of model-specific parameters and assume for them the multivariate standardised Normal prior $N(0, I_{10})$, truncated by the restriction that all eigenvalues of R lie inside the unit circle.

In the next section we present and compare 10 different ARCH-type specifications for the disturbances of the bivariate VAR(1) model in (1).Within the Bayesian posterior odds approach, the explanatory power of the i-th model is summarised by the marginal density of the $T\times 2$ observation matrix $y=(y_1...y_T)$ (given the initial conditions $y_{(0)}$), evaluated at the actual data. This density value is calculated by integrating (averaging) the likelihood function with respect to the proper prior distribution of the parameter vector $\theta_{(i)}\in\Theta_i$:

$$p(y \mid M_i, y_{(0)}) = \int_{\Theta_i} p(y \mid M_i, \theta_{(i)}, y_{(0)}) p(\theta_{(i)}) d\theta_{(i)} . \tag{2}$$

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:

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

where $P(M_h)$ and $P(M_h \mid y, y_{(0)})$ are, respectively, the prior and posterior probability of M_h; see, e.g. O'Hagan (1994). Direct evaluation of the integral in (2) (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 Markov chains; see, e.g. O'Hagan (1994), Gamerman (1997).

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

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

where $P(\theta_{(i)}|M_i,y,y_{(0)})$ denotes the posterior cumulative distribution function. Formula (3) 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.

3. Competing specifications

In this section we present and compare 10 different ARCH-type specifications for the disturbances of the bivariate VAR(1) model (1). We try to follow the *general-to-specific* strategy and start with two non-nested, conditionally t distributed multivariate GARCH(1,1)-type processes: the Vech-GARCH specification and Bollerslev's (1990) Constant Correlation model. We then consider five simplifications of t-VECH(1,1), including a simple BEKK formulation, which explains our data best. Hence we also examine special cases of our favourite t-BEKK(1,1) specification in search of a good and even more parsimonious model.

3.1. Basic non-nested specifications

First we consider the t-VECH(1,1) model (M_1), where the conditional distribution of ε_t (given its past, denoted by ψ_{t-1}) is Student t with a zero location vector, inverse precision matrix H_t and unknown degrees of freedom $\nu > 2$, i.e.

$$\varepsilon_t \mid \psi_{t-1} \sim t(0_{[2x1]}, H_t, \nu)\ ,\ H_t = \begin{bmatrix} h_{11,t} & h_{12,t} \\ h_{12,t} & h_{22,t} \end{bmatrix},$$

where the vectorisation of the lower part of H_t is parameterised as

$$\mathrm{Vech}(H_t)=\begin{pmatrix} h_{11,t} \\ h_{12,t} \\ h_{22,t} \end{pmatrix}=\begin{pmatrix} a_{10} \\ a_{20} \\ a_{30} \end{pmatrix}+\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}\begin{pmatrix} \varepsilon_{1,t-1}^2 \\ \varepsilon_{1,t-1}\varepsilon_{2,t-1} \\ \varepsilon_{2,t-1}^2 \end{pmatrix}+\begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}\begin{pmatrix} h_{11,t-1} \\ h_{12,t-1} \\ h_{22,t-1} \end{pmatrix}, \tag{4}$$

with $H_0 = h_0 I_2$ and h_0 treated as an additional free parameter. Remind that the conditional covariance matrix of ε_t given ψ_{t-1} is $(\nu-1)^{-1}\nu H_t$. In the k-variate version of this model the number of free parameters of the error process ε_t is a fourth order polynomial of k, namely $2+[1+k(k+1)]k(k+1)/2$; this gives 23 parameters when k=2. We assume prior independence for ν, h_0 and the three groups of parameters in (4). The degrees of freedom parameter follows the Exponential distribution with mean 10, $Exp(10)$, truncated by the condition $\nu > 2$. The initial value h_0 has the Exponential prior with mean 1, $Exp(1)$. For a_{j0}'s we assume the product of the densities of the following distributions: $Exp(1)$ for a_{10} and a_{30}, and $N(0, 1)$ for a_{20}, truncated by the restriction that $a_{10}\cdot a_{30} - a_{20}^2 > 0$. The prior densities of the other parameters are the products of the densities of the following Normal distributions:

$$a_{11} \sim N(0.5;1),\ a_{33} \sim N(0.5;1),\ b_{11} \sim N(0.5;1),\ b_{33} \sim N(0.5;1),$$
$$a_{ij} \sim N(0,1) \text{ and } b_{ij} \sim N(0,1) \text{ for all other pairs } (i,j);$$

these densities are truncated by the restrictions that the matrices

$$A_1 = \begin{pmatrix} a_{11} & a_{12}/2 & a_{21} & a_{22}/2 \\ a_{12}/2 & a_{13} & a_{22}/2 & a_{23} \\ a_{21} & a_{22}/2 & a_{31} & a_{32}/2 \\ a_{22}/2 & a_{23} & a_{32}/2 & a_{33} \end{pmatrix}, \quad B_1 = \begin{pmatrix} b_{11} & b_{12}/2 & b_{21} & b_{22}/2 \\ b_{12}/2 & b_{13} & b_{22}/2 & b_{23} \\ b_{21} & b_{22}/2 & b_{31} & b_{32}/2 \\ b_{22}/2 & b_{23} & b_{32}/2 & b_{33} \end{pmatrix}$$

be nonnegative definite (Gourieroux, 1997) and the eigenvalues of B_1 lie inside the unit circle.

Bollerslev (1990) argues that, for exchange rates, the assumption of constant conditional correlation may be appropriate. Thus, we also consider the following model (M_2):

$$\varepsilon_t \mid \psi_{t-1} \sim t(0_{[2x1]}, H_t, \nu),\ H_t = \begin{bmatrix} h_{11,t} & h_{12,t} \\ h_{12,t} & h_{22,t} \end{bmatrix},$$
$$h_{11,t} = a_{10} + a_{11}\varepsilon_{1,t-1}^2 + b_{11}h_{11,t-1},\ h_{22,t} = a_{20} + a_{22}\varepsilon_{2,t-1}^2 + b_{22}h_{22,t-1},$$
$$h_{12,t} = \rho_{12}\sqrt{h_{11,t}h_{22,t}},$$

where ρ_{12} is the time-invariant conditional correlation coefficient. In M_2, as in M_1, $H_0 = h_0 I_2$, where h_0 has the Exponential prior with mean 1. For the remaining parameters we take the following priors:

$$a_{10} \sim Exp(1)\,,\ a_{20} \sim Exp(1)\,,\ (a_{11}, a_{22}, b_{11}, b_{22}) \sim U([0,1]^4),\ \ \rho_{12} \sim U([-1,1])\,,$$

where $U(A)$ denotes the uniform distribution over A. In its k-variate version, M_2 describes ε_t using only $2+3k+k(k\text{-}1)/2$ free parameters; so we have 9 parameters when k=2.

Table 1 summarises model assumptions and presents the decimal logarithms of the Bayes factors in favour of M_1, $\log_{10}(B_{1j})$ for j=1,2. The decimal logarithm of the Bayes factor of M_1 against M_2, $\log_{10}(B_{12})$=29.63, indicates that – under equal prior probabilities – M_1 is about 30 orders of magnitude more probable *a posteriori* than M_2. This means that the constant conditional correlation assumption is simply improbable *a posteriori* (relative to the VECH model with no restrictions on its conditional correlations). M_2 seems too restrictive, so its simplifications and special cases will not be considered. However, the VECH model is unparsimoniously parameterised, and thus completely impractical for $k > 2$. Hence, we consider some of its special cases in search of even better models.

Table 1. Two basic models and logs of Bayes factors in favour of M_1

Model	Description	$\log_{10}(B_{1j})$
M_1 t-VECH(1,1)	$\varepsilon_t \mid H_t, \nu \sim t(0,\nu,H_t)$ with $vech H_t$ in (4)	0
M_2 t-Constant Conditional Correlations	$\varepsilon_t \mid H_t, \nu \sim t(0,\nu,H_t)$ $h_{ii,t} = a_{i0} + a_{ii}\varepsilon^2_{i,t-1} + b_{ii}h_{ii,t-1}$ (i=1,2) $h_{12,t} = \rho_{12}\sqrt{h_{11,t}h_{22,t}}$	29.63
Priors: a_{10}, a_{30} h_0 ~ Exp(1), ν ~ Exp(10), a_{11}, a_{22}, a_{33} ~ N(0.5, 1), ρ_{12} ~ U(–1, 1), other ~ N(0, 1)		
Restrictions: see the main text		

3.2. Simplifications of *t*-VECH(1,1)

All five models considered in this section can be obtained from t-VECH(1,1) by imposing certain restrictions on its parameters; the restrictions are linear and very simple for four specifications, but non-linear in the fifth case. The prior distributions for the four simpler models (M_3, M_4, M_5, M_6) are defined as the appropriate conditional distributions from the prior distribution in M_1. Only for the last model, M_7, the prior distribution is elicited separately, without any use of conditioning. The five models (M_3–M_7) as well as the decimal logarithms of the Bayes factors in favour of M_1, $\log_{10}(B_{1j})$ for j=3,...,7, are shown in Table 2.

First we mention the t-VECH(1,1) specification with zero restrictions on α and λ, i.e. the most richly parameterised model considered by Osiewalski and Pipień (2003). Since in M_1 zero values of α are completely improbable *a posteriori*, it is not surprising that the Bayes factor of M_1 against M_3 is so high. The next model is the conditionally normal VECH(1,1) specification, N-VECH(1,1) or M_4, obtained from M_1 through conditioning on $\nu=+\infty$ (and thus loosing one free parameter).

Another simplification amounts to setting $a_{2j}= b_{2j}=0$ $j=1,2,3$, in (4). This leads to M_5, the t-VECH(1,1) model with constant conditional covariance, equal to $\nu(\nu-2)^{-1}a_{20}$ for all t. In its k-variate version this model describes ε_t with $2+(1+2k)k(k+1)/2$ unknown parameters (17 free parameters when k=2). Of course, such a specification induces variable conditional correlations (except for a_{20}=0) and thus is very different from M_2, the model with constant conditional correlations. M_4 and especially M_5 fit the data worse than M_1, but not as poorly as M_3.

The fourth simplification, defining M_6, assumes $b_{ij}=0$, $i, j = 1,2,3$, in (4). This leads to the t-VechARCH(1) or t-VECH(1,0) specification with $1+[1+k(k+1)/2]k(k+1)/2$ (13 for k=2) free parameters describing ε_t. The ARCH(1) structure does not seem enough for our bivariate series, which requires the dependence of conditional covariance matrix on the more distant past of the series, which is assured by the GARCH(1,1) structure.

All four simplifications described above have (in their k-variate versions) too many parameters to be of practical use for k>2; the number of parameters in M_3, M_4 and M_6 is $O(k^4)$, similarly as in M_1, and in M_5 it is a third order polynomial of k. Now we consider a much more sophisticated simplification, where this number is only $O(k^2)$. This parsimonious model, M_7, is a simple special case of the elegant multivariate GARCH specification proposed by Baba, Engle, Kraft and Kroner (1989), and thus called BEKK in the literature. Engle and Kroner (1995) discuss general BEKK formulations and their equivalence to VechGARCH models. We consider a simple t-BEKK(1,1) specification where the conditional distribution of ε_t (given its past, ψ_{t-1}) is Student t with zero location vector, BEKK-type inverse precision matrix H_t and unknown degrees of freedom $\nu > 2$, i.e.

$$\varepsilon_t \mid \psi_{t-1} \sim t(0_{[2x1]}, H_t, \nu),$$

$$H_t = \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}' (\varepsilon_{t-1}\varepsilon_{t-1}') \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} + \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}' H_{t-1} \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \tag{5}$$

with $H_0 = h_0 I_2$ and h_0 treated as an additional parameter. Both the degrees of freedom parameter and h_0 are a priori independent of the other parameters and follow the same prior distributions as in the previous models. The other parameters are all independent a priori and with the following prior distributions:

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

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

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

truncated by the restrictions of the positive semi-definiteness of the symmetric (2x2) matrix A consisting of a_{ij} and the stability of the general (2x2) matrix C consisting of c_{ij} (all eigenvalues of C lie inside the unit circle). There are no restrictions on the general (2x2) matrix B consisting of b_{ij}. In the k-variate version, this model describes ε_t using $2+k(k+1)/2+2k^2$ free parameters (13 parameters for k=2).

Table 2. Simplifications of the t-VECH(1,1) specification (M_1).

Model	Description	$\mathrm{Log}_{10}(B_{1j})$
M_3	t-VECH(1,1) with $\alpha=\lambda=0_{[0x2]}$ (no exogenous variable)	81.97
M_4 N-VECH(1,1)	t-VECH(1,1) with $\nu \to +\infty$	7.63
M_5 t-VECH with constant conditional covariances.	t-VECH(1,1) with the following restrictions in (4): $\begin{pmatrix} h_{11,t} \\ h_{12,t} \\ h_{22,t} \end{pmatrix} = \begin{pmatrix} a_{10} \\ a_{20} \\ a_{30} \end{pmatrix} + \begin{pmatrix} a_{11} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & a_{33} \end{pmatrix} \begin{pmatrix} \varepsilon^2_{1,t-1} \\ \varepsilon_{1,t-1}\varepsilon_{2,t-1} \\ \varepsilon^2_{2,t-1} \end{pmatrix} + \begin{pmatrix} a_{11} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & a_{33} \end{pmatrix} \begin{pmatrix} h_{11,t-1} \\ h_{12,t-1} \\ h_{22,t-1} \end{pmatrix}$	22.73
M_6 t-VECH(1,0)	t-VECH(1,1) with restrictions $b_{ij}=0$, i,j=1,2,3.	4.42
M_7 t-BEKK(1,1)	For restrictions in (4) see Osiewalski and Pipień (2001); Assumed $vech H_t$ is presented in (5).	-2.52
Priors specified through conditioning (in M_3–M_6) or independently (in M_7)		

Since M_7 is much better than M_2–M_6, and about 2–3 orders of magnitude better than M_1, let us comment on the relation between M_7 and M_1. In spite of formal incompatibility of their prior specifications, both models lead to almost the same posterior distributions of quantities of interest (common parameters or conditional covariances and correlations) and to the same predictive results. Osiewalski and Pipień (2001) show that the simple BEKK(1,1) error process can be obtained from Vech-GARCH(1,1) in 64 alternative ways, each time by imposing 10 non-linear restrictions on a_{ij}s and b_{ij}s in (4). Using a Lindley type test based on the approximate Normality of certain functions of basic parameters in the t-VECH(1,1) model (with no exogenous variables), they conclude that the data set analysed here does not give clear support to the simple BEKK specification for the disturbances, although it is not rejected either. This very weak conclusion, based on the analysis of the posterior distribution in the t-VECH(1,1) model, is quite different from the reasoning based on the posterior odds ratio, which favours parsimony and leaves no doubt about the superiority of the t-BEKK(1,1) error structure. Our data favour the t-BEKK(1,1) model over all the alternatives considered so far. It appears as flexible as the t-VECH(1,1) specification, leading to virtually the same posterior inference on quantities of interest, but it has much less free parameters. In the next subsection we show consequences of further simplifications of M_7. The main question is whether reducing the number of free parameters in M_7 can increase the marginal data density value.

3.3. Simplifications of *t*-BEKK(1,1)

There are two natural reductions of M_7. One is the t-BEKK(1,0) model, M_8, which appears as a result of imposing zero restrictions on all c_{ij}s in (5), the other is the N-BEKK(1,1) specification, M_9, obtained by taking the limit $\nu=+\infty$ for the degrees of freedom parameter. The third model, N-BEKK(1,0) or M_{10}, results from jointly imposing all these restrictions. The prior distributions for all three simpler models (M_8, M_9, M_{10}) are defined as the appropriate conditional distributions from the prior distribution in M_7. Table 3 presents the three models as well as the decimal logarithms of the Bayes factors in favour of M_7, $\log_{10}(B_{7i})$ (i=8,9,10), calculated using the N-R method. Not surprisingly, conditional normality of the error process is strongly rejected by the data, which is in full accordance with the marginal posterior distribution for the degrees of freedom parameter ν in M_7 (most of the posterior mass is concentrated in the interval [3; 6.5]). As regards the reduction of M_7 to M_8 (a model with an ARCH(1) structure, Student t conditional distribution, and only $1+k(k+1)/2+k^2$ free parameters for ε_t, i.e. 8 if k=2), t-BEKK(1,0) is about 6 orders of magnitude worse than t-BEKK(1,1), about 3–4 orders of magnitude worse than M_1, but slightly better than M_6; see also Table 4.

The overall qualitative conclusion (based on the N-R estimates of the marginal data density values) is that M_7, i.e. the t-BEKK(1,1) specification (with free α and λ), is the best model among all 10 models under consideration. The Bayes factors in favour of M_7 are so high that this particular specification would receive practically all the posterior probability mass under any reasonable prior model probabilities.

Table 3. Simplifications of the t-BEKK(1,1) specification (M_7).

Model	Description	$\log_{10}(B_{7i})$
M_8 t-BEKK(1,0)	t-BEKK(1,1) with $c_{ij}=0$ (i,j=1,2) in (5)	6.05
M_9 N-BEKK(1,1)	t-BEKK(1,1) $\nu\to+\infty$	36.82
M_{10} N-BEKK(1,0)	N-BEKK(1,1) with $c_{ij}=0$ (i,j=1,2) in (5)	39.35
Priors obtained through conditioning		

Table 4. Logs of Bayes factors in favour of *t*-BEKK(1,1).

	With exogenous variable			No exogenous variable		
Model	Number of parameters	Rank	$\log_{10}(B_{7i})$	Number of parameters	Rank	$\log_{10}(B_{7i})$
M_7 , *t*-BEKK(1,1)	23	1	0	19	1	0
M_1 , *t*-VECH(1,1)	33	2	2.52	29	2	4.50
M_8 , *t*-BEKK(1,0)	17	3	6.05	14	3	6.22
M_6 , *t*-VECH(1,0)	23	4	6.94	19	4	10.74
M_5 , *t*-VECH(1,1) ConstCovar.	27	6	25.25	23	5	11.53
M_2 , *t*-ConstCor(1,1)	19	7	32.15	15	6	27.84
M_9 , N-BEKK(1,1)	22	8	36.82	18	7	44.53
M_{10} , N-BEKK(1,0)	17	9	39.35	13	8	47.42
M_4 , N-VECH(1,1)	32	5	10.16	28	9	48.10
t-VECH(1,1) no exo	29	10	84.49	see M_1	–	–

3.4. Stability of model comparisons

In this subsection we discuss the stability of Bayes factors and model ranks with respect to the assumption $\alpha=\lambda=0$, i.e. the lack of the exogenous variable and the ECM term. Table 4 presents the decimal logarithms of the Bayes factors in favour of the *t*-BEKK(1,1) model (M_7), i.e. the values of $\log_{10}(B_{7i})$. The Table also shows the total number of free parameters of each specification, including common δ, α, λ and R from (1). While the Bayes factors (obtained in two cases) can be very different, the resulting model ranks are similar enough – they indicate the leading position of the *t*-BEKK(1,1) specification.

The results in Table 4 were obtained under specific proper prior distributions over parameter spaces of particular models. These priors are consistent with our prior knowledge and not too informative. When we make the priors much more diffuse, the marginal data density values for reasonable models change by about 2 orders of magnitude, which should be compared to the huge distance between the best and the worst specifications. However, any further sensitivity analysis was too costly in terms of computational time and effort.

4. Concluding remarks

Introducing the exogenous DEM/USD exchange rate (through our conditional ECM model) has almost no effect on the results of the Bayesian comparison of competing bivariate GARCH error processes for the pair of growth rates of PLN/USD and PLN/DEM. Obviously, the presence of DEM/USD helps enormously in explaining the modelled growth rates and thus reduces the unexplained volatility. This does not mean, however, that we suggest using such relevant exogenous variables (and conditional models) in predictive analyses like option pricing or building dynamic hedging strategies. Exogenous variables are very useful in explaining volatility *ex post*, but are uncertain *ex ante*. Hence it seems reasonable to base predictive analyses on good models of marginal processes for the forecasted financial instruments. For our data set and class of models, the simple VAR(1) – *t*-BEKK(1,1) specification considered by Osiewalski and Pipień (2003) seems a reasonable approximation of the marginal bivariate process generating PLN/USD and PLN/DEM.

References

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

Bauwens, L., Lubrano, M. (1998), Bayesian Inference on GARCH Models Using the Gibbs Sampler, *Econometrics Journal*, 1, C23–C46.

Bauwens, L., Lubrano, M., Richard, J.–F. (1999), *Bayesian Inference In Dynamic Econometric Models,* Oxford University Press, Oxford.

Bollerslev, T. (1986), Generalised Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*, 31, 307–327.

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.

Bos, Ch.S., Mahieu, R.J., Van Dijk, H.K. (2000), Daily Exchange Rate Behaviour and Hedging of Currency Risk, *Journal of Applied Econometrics*, 15, 671–696.

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

Engle, R.F. (1982), Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation, *Econometrica*, 50, 987–1007.

Engle, R.F., Granger, C.W.J. (1987), Co-integration and Error Correction: Representation, Estimation and Testing, *Econometrica*, 55, 251–276.

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

Florens, J.–P., Mouchart, M. (1985), Conditioning in Dynamic Models, *Journal of Time Series Analysis*, 6, 15–34.

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

Geweke, J. (1989), Exact Predictive Densities for Linear Models with ARCH Disturbances, *Journal of Econometrics*, 40, 63–86.

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

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

Kleibergen, F., Van Dijk, H.K. (1993), Non-stationarity in GARCH Models: A Bayesian Analysis, *Journal of Applied Econometrics*, 8, S41–S61.

Newton, M. A., 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*, Edward Arnold, London.

Osiewalski, J., Pipień, M. (1999), Bayesian Forecasting of Foreign Exchange Rates using GARCH Models with Skewed t Conditional Distributions, in: Welfe, W., ed., *MACROMODELS'98 – Conference Proceedings,* Vol. 2 (Absolwent, Łódź), 195–218.

Osiewalski, J., Pipień, M. (2000), GARCH-In-Mean Through Skewed t Conditional Distributions: Bayesian Inference for Exchange Rates, in: Welfe, W., Wdowiński, P., eds., *MACROMODELS'99 – Conference Proceedings* (Absolwent, Łódź), 354–369.

Osiewalski, J., Pipień, M. (2001), Multivariate *t*-GARCH Models – Bayesian Analysis for Exchange Rates, in: Welfe, W., ed., *MACROMODELS'2000 – Conference Proceedings* (Absolwent, Łódź), 128–148.

Osiewalski, J., Pipień, M. (2003), Bayesian Comparison of Bivariate ARCH-type Models for the Main Exchange Rates in Poland, *Journal of Econometrics (forthcoming).*

Osiewalski, J., Steel, M.F.J. (1996), A Bayesian Analysis of Exogeneity in Models Pooling Time-series and Cross-sectional Data, *Journal of Statistical Planning and Inference 50*, 187–206.

Vrontos, I.D., Dellaportas, P., Politis, D.N. (2000), Full Bayesian Inference for GARCH and EGARCH Models, *Journal of Business and Economic Statistics*, 18, 187–198.