Central European Journal of Economic Modelling and Econometrics

Bayesian Analysis for Hybrid MSF-SBEKK Models of Multivariate Volatility

Jacek Osiewalski*, Anna Pajor†

Submitted: 20.05.2009, Accepted: 3.08.2009

Abstract

The aim of this paper is to examine the empirical usefulness of two new MSF - Scalar BEKK(1,1) models of n-variate volatility. These models formally belong to the MSV class, but in fact are some hybrids of the simplest MGARCH and MSV specifications. Such hybrid structures have been proposed as feasible (yet non-trivial) tools for analyzing highly dimensional financial data (large n). This research shows Bayesian model comparison for two data sets with $n = 2$, since in bivariate cases we can obtain Bayes factors against many (even unparsimonious) MGARCH and MSV specifications. Also, for bivariate data, approximate posterior results (based on preliminary estimates of nuisance matrix parameters) are compared to the exact ones in both MSF-SBEKK models. Finally, approximate results are obtained for a large set of returns on equities ($n = 34$).

Keywords: Bayesian econometrics, Gibbs sampling, time-varying volatility, multivariate GARCH processes, multivariate SV processes

JEL Classification: C11, C32, C51.

*Department of Econometrics and Operations Research, Cracow University of Economics and Andrzej Frycz Modrzewski Cracow Academy, e-mail: eeosiewa@cyf-kr.edu.pl

†Department of Econometrics and Operations Research, Cracow University of Economics, e-mail: eopajor@cyf-kr.edu.pl

1 Introduction

Most of n-variate volatility models used in financial econometrics belong to either the MGARCH or MSV (multivariate stochastic volatility or variance) class; see, e.g., Bauwens, Laurent, Rombouts (2006), and Tsay (2005). Only a few of them are practical tools for analysing large portfolios, e.g. the Scalar BEKK (SBEKK) model and, in particular, the Dynamic Conditional Correlation (DCC) structure of Engle (2002). Both cases represent the MGARCH class and in each we can use variance targeting and approximate methods to estimate the parameter vector of dimension growing with the portfolio size; the remaining parameters, requiring more numerical effort, form a vector of fixed dimension irrespective of the number of assets. However, according to the Bayesian posterior odds criterion, MGARCH models may explain data much worse than MSV specifications; see Osiewalski, Pajor, Pipień (2007). So one should be interested in applying MSV models instead of MGARCH.

Latent AR(1) processes, used in the MSV class to describe volatility, are very efficient in dealing with outliers and, thus, in modelling tail behaviour. Since such modelling is crucial for any risk assessment, the MSV class should be kept under consideration in spite of the fact that MSV structures with many latent processes are too complicated to be practical in highly dimensional problems. Easier way of modelling was proposed by Osiewalski and Pajor (2007) through a hybrid model, based on Engle's DCC structure and the simplest MSV structure, the Multiplicative Stochastic Factor (MSF, or Stochastic Discount Factor, SDF) specification. However, the MSF-DCC (previously called SDF-DCC) model is still too complex and, thus, Osiewalski (2009) has recently proposed some MSF-SBEKK (SDF-DCC) hybrids. This paper is devoted to examining their empirical usefulness.

Main n-variate volatility models and the hybrid structures are presented in section 2. Section 3 is devoted to Bayesian model comparison for two sets of bivariate observations ($n = 2$), where we can obtain exact results considering unparsimonious structures as well. In section 4 we describe our approximate posterior inference for the MSF-SBEKK models and compare it to the exact one (using the same bivariate data sets). In section 5 approximate posterior results are obtained for the MSF-SBEKK models applied to a large set of returns on equities ($n = 34$). Conclusions are grouped in section 6.

The approach adopted in this paper is very different from the modelling strategy of Chib, Nardari, Shephard (2006), who proposed factor SV models for highly dimensional data (together with their Bayesian analysis using Markov Chain Monte Carlo tools). They did not apply any GARCH ideas and structures in their models, staying within the pure SV class, but relying on sophisticated specifications. Instead, we try to combine the simplest MSV and MGARCH ways of dealing with multivariate volatility.

2 Hybrid n-variate volatility specifications

Assume there are n assets. We denote by $r_t = (r_{1,t} \dots r_{n,t})$ n-variate observations on their logarithmic return (or growth) rates, and we model them using the basic VAR(1) framework:

$$r_t = \delta_0 + r_{t-1}\Delta + \varepsilon_t; \quad t = 1, \dots, T, \dots, T+s; \tag{1}$$

where T is the length of the observed time series and s is the forecast horizon. The $n(n+1)$ elements of $\delta = \left(\delta_0 \left(\text{vec}\Delta\right)'\right)'$ are common parameters that are assumed *a priori* independent of model-specific parameters, with, e.g., the $N\left(0, I_{n(n+1)}\right)$ prior.

2.1 Main MSV and MGARCH models

In MSV specifications for ε_t in (1) it is assumed that ε_t is conditionally Normal (given parameters and latent variables, grouped in θ) with mean vector 0 and covariance matrix Σ_t that depends on latent variables, i.e.

$$\varepsilon_t|\theta \sim N\left(0_{[1\times n]}, \Sigma_t\right).$$

Thus, the corresponding conditional distribution of r_t (given its past and θ) is Normal with mean $\mu_t = \delta_0 + r_{t-1}\Delta$ and covariance matrix Σ_t. Competing n-variate MSV models are defined using different latent processes and different structures of Σ_t (symmetric and positive definite by construction). Main MSV specifications are briefly presented in Table 1.

Table 1: Main MSV structures

Model	Form of the $n \times n$ matrix Σ_t	Number of free parameters [latent variables]* in Σ_t, $(t = 1, 2, \dots, T)$
TSV	Cholesky decomposition: $\Sigma_t = L_t G_t L_t'$, $G_t = \text{diag}\left(\exp\left(q_{11,t}\right), \dots, \exp\left(q_{nn,t}\right)\right)$, L_t with ones on the diagonal and $q_{ij,t}$ below $(i > j)$, $\{q_{kl,t}\}$ are Gaussian AR(1) processes $(k, l = 1, \dots, n;\ k \geq l)$	$\frac{3n(n+1)}{2}$ $\left[\frac{n(n+1)}{2}\right]$
JSV	spectral decomposition: $\Sigma_t = P\Lambda_t P'$, $\Lambda_t = \text{diag}(\exp(q_{11,t}), \dots, \exp(q_{nn,t}))$, $q_{ii,t}$ are Gaussian AR(1) processes $(i = 1, \dots, n)$, P is free orthogonal matrix of eigenvectors of Σ_t	$3n + \frac{n(n-1)}{2}$ $[n]$
BSV	special case of JSV, restriction: $P = I_n$	$3n$ $[n]$
MSF	$\Sigma_t = g_t A$, A is symmetric positive definite (free), $\{\ln g_t\}$ is a Gaussian AR(1) process	$2 + \frac{n(n+1)}{2}$ $[1]$

* without initial conditions, which can be estimated

We describe in some detail the Bayesian analysis of the simplest MSV specification, the Multiplicative Stochastic Factor (MSF, also called Stochastic Discount Factor, SDF) model that is used to construct hybrid MSV-MGARCH structures. The MSF

model relies on just one latent process gt to describe the dynamics of Σ_t (see Jacquier, Polson, Rossi, (1995); g_0 can be fixed by assuming, e.g., $g_0 = 1$): $\varepsilon_t = \zeta_t\sqrt{g_t}$, $\ln g_t = \varphi \ln g_{t-1} + \sigma_g \eta_t$, $\zeta_t \sim iiN(0_{[1\times n]}, A)$, $\eta_t \sim iiN(0,1)$, $\zeta_t \perp \eta_s$ $(t, s \in Z)$.
The conditional covariance matrix of ε_t takes the very simple form $\Sigma_t = g_t A$, which leads to the invariable conditional correlation coefficient $\rho_{ij,t} = \rho_{ij} = \frac{a_{ij}}{\sqrt{a_{ii}a_{jj}}}$ (for $i, j = 1, \ldots, n$); $A = [a_{ij}]$ is a free symmetric positive definite matrix consisting of $\frac{n(n+1)}{2}$ distinct entries.
We can assume independence among parameters and use the same prior distributions as Pajor (2005a,b); for φ: Normal with mean 0 and variance 100, truncated to $(-1, 1)$, for σ_g^{-2}: Exponential with mean 200, for A^{-1}: Wishart with mean I_n.
The n-variate Bayesian VAR(1)-MSF model can be analysed using Gibbs sampling as a tool for simulating samples from the posterior distribution. This is due to the Wishart or Normal forms of the full conditional distributions of A and δ, which make these steps of the Gibbs sampler easy even for large n. Other steps, numerically more demanding, are the same for each n; see Jacquier, Polson, Rossi (1995), and Pajor (2009). Despite its convenience in practical applications, the MSF specification is too restrictive to be useful since it assumes the same dynamics for all entries of Σ_t. This assumption seems too high a price to be paid for the ease of numerical implementation.
In MGARCH specifications for ε_t in (1) it is assumed that the conditional distribution of ε_t (given past, ψ_{t-1}, and parameters) is n-variate Student t with location vector 0, inverse precision matrix H_t and $\nu > 2$ degrees of freedom, i.e. $\varepsilon_t|\theta, \psi_{t-1} \sim St(0_{[1\times n]}, H_t, \nu)$. From the empirical perspective, the use of the Student t distribution instead of conditional Normality is fully justified; see Osiewalski and Pipień (2004), and Osiewalski, Pajor, Pipień (2007).
Particular n-variate MGARCH models are defined by imposing different structures on H_t. They are different functions of the past of ${\varepsilon_t}'\varepsilon_t$ and the past of H_t itself; if they involve only the immediate past (of both, ${\varepsilon_t}'\varepsilon_t$ and H_t), the models are denoted as MGARCH(1,1). The basic forms of H_t are shown in Table 2. For the bivariate case $(n = 2)$, the prior distributions and Bayesian models are presented by Osiewalski, Pajor, Pipień (2007).
Let us focus on the N-SBEKK(1,1) model (corresponding to $\nu \to \infty$, i.e. conditional Normality), which is the simplest n-variate structure allowing for dynamic conditional correlation. The SBEKK(1,1) case can be alternatively parameterised as

$$H_t = (1 - \beta - \gamma) A + \beta \left(\varepsilon'_{t-1}\varepsilon_{t-1}\right) + \gamma H_{t-1} \tag{2}$$

where A is a free symmetric positive definite matrix of order n, with, e.g., an inverted Wishart prior distribution, and β and γ are free scalar parameters, jointly uniformly distributed over the unit simplex. As regards initial conditions for H_t, we can take $H_0 = h_0 I_n$ and treat $h_0 > 0$ as an additional parameter (*a priori* exponentially distributed with mean 1). Note that, under the covariance stationarity restriction $(\beta + \gamma < 1)$, A is the unconditional covariance matrix of ε_t.

Table 2: Basic MGARCH(1,1) structures

Model	Form of the $n \times n$ matrix H_t ($t = 1, 2, \ldots$; H_0 assumed known)	Number of free parameters in H_t
VECH	$\mathrm{vech}(H_t)=\mathrm{vech}(A_0)+B^*\mathrm{vech}(\varepsilon_t{}'\varepsilon_t)+C^*\mathrm{vech}(H_{t-1})$, A_0 is a square matrix of order n (symmetric and positive definite), B^* and C^* free square matrices of order $\frac{n(n+1)}{2}$	$\frac{[1+n(n+1)]n(n+1)}{2}$ $O(n^4)$
diagonal VECH	special case of VECH, restrictions: B^*, C^* - diagonal matrices	$\frac{3n(n+1)}{2}$ $O(n^2)$
BEKK	$H_t = A_0 + B\varepsilon_t{}'\varepsilon_t B' + CH_{t-1}C'$, A_0 is a square matrix of order n (symmetric and positive definite), B and C are free square matrices of order n (with $b_{11} \geq 0$ and $c_{11} \geq 0$ to ensure identification)	$2n^2 + \frac{n(n+1)}{2}$ $O(n^2)$
SBEKK	special case of BEKK, restrictions: $B = \beta^{\frac{1}{2}} I_n$, $C = \gamma^{\frac{1}{2}} I_n$	$2 + \frac{n(n+1)}{2}$
DCC Engle (2002): only $\nu \to \infty$	$H_t = (D_t)^{-1}R_t(D_t)^{-1}$, $D_t = \mathrm{diag}(h_{11,t}^{-\frac{1}{2}}, \ldots, h_{nn,t}^{-\frac{1}{2}})$, $h_{ii,t} = a_{i0} + a_i\varepsilon_{i,t-1}^2 + b_i h_{ii,t-1}$ $(i = 1, \ldots, n)$, $R_t = P_tQ_tP_t$, $P_t = \mathrm{diag}(q_{11,t}^{-\frac{1}{2}}, \ldots, q_{nn,t}^{-\frac{1}{2}})$ $Q_t = (1-\beta-\gamma)S + \beta(1-\frac{2}{\nu})D_t\varepsilon_t'\varepsilon_t D_t + \gamma Q_{t-1}$ S is a square matrix parameter (of order n, with all properties of the correlation matrix); β, γ, a_{i0}, a_i, b_i $(i = 1, \ldots, n)$ are nonnegative scalar parameters $(\beta + \gamma < 1$, $a_{i0} > 0$, $b_i < 1)$	$2 + 3n + \frac{n(n-1)}{2}$ ν not counted $O(n^2)$
CCC	special case of DCC, restriction: $\beta = \gamma = 0$ (thus $R_t = S$)	$3n + \frac{n(n-1)}{2}$

Although the VAR(1)-N-SBEKK(1,1) specification is so simple, its Bayesian analysis cannot rely on a fully automatic Gibbs chain as in the VAR(1)-MSF case. The conditional posterior of the VAR(1) parameters is no longer of the Normal form as δ also appears in H_t through lagged ε_t. The conditional posterior of A is not inverted Wishart as A is not proportional to the conditional covariance matrix; it determines just one term of the sum in (2). The complicated form of the conditional posterior of β and γ is not a problem since they are scalar parameters. But (δ, A) is of high dimension for large n and the use of Metropolis draws within the Gibbs steps (or the Metropolis chain for all the parameters jointly, as in Osiewalski and Pipień (2004) and further works on bivariate GARCH models) can be infeasible. Thus, any practical estimation tool must rely on some crude approximation when the analysed portfolio is very large. We could use OLS for the VAR(1) part and rely on variance targetting in order to fix A; that is, A could be replaced by the empirical covariance matrix of the OLS residuals from the VAR part. In practice, the Bayesian analysis for the scalar parameters (β, γ) and future returns would then be based on the conditional posterior and predictive distributions given the values of the highly dimensional parameters (δ, A).

2.2 MSF-DCC and MSF-SBEKK structures

Osiewalski and Pajor (2007) introduced a new hybrid structure by assuming the following specification for the residual process ε_t in (1):

$$\varepsilon_t = \zeta_t H_t^{\frac{1}{2}} \sqrt{g_t}, \qquad \ln g_t = \varphi \ln g_{t-1} + \sigma_g \eta_t, \qquad (\zeta_t, \eta_t)' \sim iiN\left(0_{[(n+1)\times 1]}, I_{n+1}\right),$$

where H_t has the DCC structure proposed by Engle (2002) and corresponding to the one in Table 2 with $\nu \to \infty$. Now, given the univariate process g_t and the parameters, the n-variate returns r_t in (1) follow the conditional Normal distribution with mean $\mu_t = \delta_0 + r_{t-1}\Delta$ and covariance matrix $\Sigma_t = g_t H_t$ depending on $5 + 3n + \frac{n(n-1)}{2}$ parameters (including h_0). Contrary to the MSF model, the MSF-DCC conditional covariances do not share the same dynamics and the correlation coefficients can vary over time. Empirical evidence presented by Osiewalski and Pajor (2007) and Pajor (2009) shows that the MSF-DCC (or SDF-DCC) model is promising; it keeps the flexible conditional correlation structure of the Engle's DCC specification, but fits data better due to the presence of the latent AR(1) process. However, the Bayesian analysis of the VAR(1) - MSF-DCC model (performed in the above mentioned works only in bivariate and trivariate cases) is numerically too complicated for large n.
Osiewalski (2009) proposes new, extremely simple hybrid structures by assuming the following specification for the residual process ε_t in (1):

$$\varepsilon_t = \zeta_t \Omega_t^{\frac{1}{2}} \sqrt{g_t}, \qquad \ln g_t = \varphi \ln g_{t-1} + \sigma_g \eta_t, \qquad (\zeta_t, \eta_t)' \sim iiN\left(0_{[(n+1)\times 1]}, I_{n+1}\right).$$

That is, ε_t is conditionally (given its past, ψ_{t-1}, the scalar latent process g_t and the parameters) Normal with mean vector 0 and covariance matrix $g_t\Omega_t$, where Ω_t is a time-varying square matrix of order n that preserves the SBEKK structure. Again, g_0 can be fixed as in the MSF model. Two particular forms of Ω_t are considered.
In the type I MSF-SBEKK model $\Omega_t = H_t$, so it follows (2) and does not depend on the latent process. The conditional variances are equal to $g_t h_{ii,t}$, that is they have a more general form than in either the MSF or SBEKK model, but the conditional correlation coefficient does not depend on g_t and thus is of the SBEKK form. Note that this generalised structure does not lead to the MGARCH form of the process $\varepsilon_t^* = \frac{\varepsilon_t}{\sqrt{g_t}}$ as H_t depends on ε_{t-1}, not on ε_{t-1}^*.
In the type II MSF-SBEKK model the MGARCH structure is assumed for ε_t^*, i.e.

$$\Omega_t = H_t^*, \qquad H_t^* = (1 - \beta - \gamma)\, A + \beta\left(\varepsilon_{t-1}^{*\prime}\varepsilon_{t-1}^*\right) + \gamma H_{t-1}^*.$$

Now Ω_t depends on the whole past of the latent process, so do the conditional variances and correlation coefficients of ε_t. Modelling time-varying conditional correlation is no longer as simple as in the pure SBEKK or type I hybrid models.
Note that the "integrated" versions of our hybrid models can be defined by imposing the restriction $\beta + \gamma = 1$; see Pajor (2009) for the MSF-IDCC (or SDF-IDCC) specification and its applications. These models will not be considered in this paper.

The MSF-SBEKK models (as well as MSF-DCC) belong to the MSV class as they describe multivariate volatility using the latent AR(1) process. However, we call them hybrid models in order to stress their difference from traditional "pure" MSV conditional covariance structures that do not depend on past observations.
Both hybrid MSF-SBEKK models seem useful as they combine important properties of their main structural components. The latent AR(1) process in the conditional covariance matrix should help in explaining outlying observations, and the dependence on the past data (through the BEKK structure of Ω_t) prevents the entries of the conditional covariance matrix $g_t\Omega_t$ from sharing the same dynamic pattern. Thus our models have time-varying conditional correlation coefficients without introducing more latent processes. The MSF-SBEKK models describe $\Sigma_t = g_t\Omega_t$ using $5 + \frac{n(n+1)}{2}$ parameters and nest both basic structures. In the limiting case: $\sigma_g \to 0$, $\varphi = 0$ we are back in the SBEKK model, while $\beta = \gamma = 0$ lead to the MSF case.
Assuming that the parameters of our hybrid specifications follow the same priors as in both special cases (MSF, SBEKK), we can write the full Bayesian model as

$$\begin{aligned}
&p\left(r_1,\ldots,r_{T+s},\ln g_1,\ldots,\ln g_{T+s},\delta,A,\beta,\gamma,\varphi,\sigma_g^{-2},h_0\right)\\
&= p\left(r_{T+1},\ldots,r_{T+s}|r_1,\ldots,r_T,\ln g_1,\ldots,\ln g_{T+s},\delta,A,\beta,\gamma,\varphi,\sigma_g^{-2},h_0\right)\\
&\times p\left(\ln g_{T+1},\ldots,\ln g_{T+s}|r_1,\ldots,r_T,\ln g_1,\ldots,\ln g_T,\delta,A,\beta,\gamma,\varphi,\sigma_g^{-2},h_0\right)\\
&\times p\left(r_1,\ldots,r_T,\ln g_1,\ldots,\ln g_T,\delta,A,\beta,\gamma,\varphi,\sigma_g^{-2},h_0\right),
\end{aligned}$$

where

$$\begin{aligned}
&p\left(r_{T+1},\ldots,r_{T+s}|r_1,\ldots,r_T,\ln g_1,\ldots,\ln g_{T+s},\delta,A,\beta,\gamma,\varphi,\sigma_g^{-2},h_0\right)\\
&= \prod_{t=T+1}^{T+s} f_N^n\left(r_t|\mu_t,g_t\Omega_t\right),\\
&p\left(\ln g_{T+1},\ldots,\ln g_{T+s}|r_1,\ldots,r_T,\ln g_1,\ldots,\ln g_T,\delta,A,\beta,\gamma,\varphi,\sigma_g^{-2},h_0\right)\\
&= \prod_{t=T+1}^{T+s} f_N^1\left(\ln g_t|\varphi\ln g_{t-1},\sigma_g^2\right),
\end{aligned}$$

$$\begin{aligned}
&p\left(r_1,\ldots,r_T,\ln g_1,\ldots,\ln g_T,\delta,A,\beta,\gamma,\varphi,\sigma_g^{-2},h_0\right) \qquad (3)\\
&= p\left(\delta\right)p\left(A\right)p\left(h_0\right)p\left(\beta,\gamma\right)p\left(\varphi,\sigma_g^{-2}\right)\prod_{t=1}^{T} f_N^n\left(r_t|\mu_t,g_t\Omega_t\right)f_N^1\left(\ln g_t|\varphi\ln g_{t-1},\sigma_g^2\right)
\end{aligned}$$

The first two densities enable direct Monte Carlo simulation of future g_t and r_t; given all the parameters and $(g_t,\ r_t)$ from the observation period $(t = 1,\ldots,T)$, we successively draw $\ln g_{T+j}$ and r_{T+j} $(j = 1,\ldots,s)$ from their conditional Normal

distributions. The last term, i.e. (3), is the joint density of the observed return rates, the T corresponding latent variables and all the parameters. The posterior density function, proportional to (3), is very complicated and highly dimensional. The only hope to perform any Bayesian analysis is in the application of Gibbs sampling, which is based on the conditional distributions obtained from (3):

$$
\begin{aligned}
&p\left(\delta|r_1,\ldots,r_T,\ln g_1,\ldots,\ln g_T,A,\beta,\gamma,\varphi,\sigma_g^{-2},h_0\right) \propto p\left(\delta\right)\prod_{t=1}^{T} f_N^n\left(r_t|\mu_t,g_t\Omega_t\right),\\
&p\left(A|r_1,\ldots,r_T,\ln g_1,\ldots,\ln g_T,\delta,\beta,\gamma,\varphi,\sigma_g^{-2},h_0\right) \propto p\left(A\right)\prod_{t=1}^{T} f_N^n\left(r_t|\mu_t,g_t\Omega_t\right),\\
&p\left(\beta,\gamma,h_0|r_1,\ldots,r_T,\ln g_1,\ldots,\ln g_T,\delta,A,\varphi,\sigma_g^{-2}\right)\\
&\quad\propto p\left(\beta,\gamma\right)p\left(h_0\right)\prod_{t=1}^{T} f_N^n\left(r_t|\mu_t,g_t\Omega_t\right),\\
&p\left(\varphi,\sigma_g^{-2}|r_1,\ldots,r_T,\ln g_1,\ldots,\ln g_T,\delta,A,\beta,\gamma,h_0\right)\\
&\quad\propto p\left(\varphi,\sigma_g^{-2}\right)\prod_{t=1}^{T} f_N^1\left(\ln g_t|\varphi\ln g_{t-1},\sigma_g^2\right),\\
&p\left(\ln g_t|r_1,\ldots,r_T,\ln g_1,\ldots,\ln g_{t-1},\ln g_{t+1},\ldots,\ln g_T,\delta,A,\beta,\gamma,\varphi,\sigma_g^{-2},h_0\right)\\
&\quad\propto f_N^1\left(\ln g_t|\varphi\ln g_{t-1},\sigma_g^2\right)f_N^1\left(\ln g_{t+1}|\varphi\ln g_t,\sigma_g^2\right)\prod_{j=t}^{T} f_N^n\left(r_j|\mu_j,g_j\Omega_j\right);\\
&\quad t=1,\ldots,T-1;\\
&p\left(\ln g_T|r_1,\ldots,r_T,\ln g_1,\ldots,\ln g_{T-1},\delta,A,\beta,\gamma,\varphi,\sigma_g^{-2},h_0\right)\\
&\quad\propto f_N^1\left(\ln g_T|\varphi\ln g_{T-1},\sigma_g^2\right)f_N^n\left(r_T|\mu_T,g_T\Omega_T\right)
\end{aligned}
$$

The latent variables and parameters related to the SV component of the hybrid structure can be simulated in a similar way as in the pure MSF model, that is relatively easy. Under a Normal-Gamma prior for the pair $\left(\varphi,\sigma_g^{-2}\right)$, its bivariate conditional posterior is also of the Normal-Gamma form. And, most importantly, the univariate conditional posterior densities for $\ln g_t$ do not look too different from the pure MSF case. In fact, it may pay to use the full conditional of g_t^{-1} as it involves a Gamma kernel. In the type I model we obtain for $t=1,\ldots,T-1$

$$
\begin{aligned}
&p\left(g_t^{-1}|r_1,\ldots,r_T,g_1,\ldots,g_{t-1},g_{t+1},\ldots,g_T,\delta,A,\beta,\gamma,\varphi,\sigma_g^{-2},h_0\right)\\
&\quad\propto f_G\left(g_t^{-1}|\frac{n}{2},\frac{d_t}{2}\right)f_N^1\left(\ln g_t|\varphi\ln g_{t-1},\sigma_g^2\right)f_N^1\left(\ln g_{t+1}|\varphi\ln g_t,\sigma_g^2\right);
\end{aligned}
$$

$$p\left(g_T^{-1}|r_1,\ldots,r_T,g_1,\ldots,g_{T-1},\delta,A,\beta,\gamma,\varphi,\sigma_g^{-2},h_0\right)$$
$$\propto f_G\left(g_T^{-1}|\frac{n}{2},\frac{d_T}{2}\right) f_N^1\left(\ln g_T|\varphi\ln g_{T-1},\sigma_g^2\right),$$

where $d_t = (r_t-\mu_t)\,\Omega_t^{-1}\,(r_t-\mu_t)'$. Since the log-Normal kernels can be approximated by some Gamma density, the Metropolis-Hastings steps have very efficient proposal density for g_t^{-1}; see Pajor (2009). In the type II model, the corresponding full conditional densities are more complicated due to the dependence of Ω_t on the past of g_t; we have for $t=1,\ldots,T-1$:

$$p\left(g_t^{-1}|r_1,\ldots,r_T,g_1,\ldots,g_{t-1},g_{t+1},\ldots,g_T,\delta,A,\beta,\gamma,\varphi,\sigma_g^{-2},h_0\right)$$
$$\propto f_G\left(g_t^{-1}|\frac{n}{2},\frac{d_t}{2}\right) f_N^1\left(\ln g_t|\varphi\ln g_{t-1},\sigma_g^2\right) f_N^1\left(\ln g_{t+1}|\varphi\ln g_t,\sigma_g^2\right)\prod_{j=t+1}^{T} f_N^n\left(r_j|\mu_j,\Omega_j\right).$$

The last product, which did not involve g_t (and thus was omitted) in the type I model, is now a complicated function of g_t. So we follow the numerical strategy described in the Appendix in order to efficiently draw (within the Gibbs sampler) the whole vector of $\ln g_t$ for $t=1,\ldots,T$.

The conditionals of the VAR(1) coefficients and the parameters describing the SBEKK part of the structure have non-standard forms. (h_0,β,γ) can be sampled using the Metropolis-Hastings step within the Gibbs sampler. But the full conditional posteriors of δ and A (the parameters of dimension related to n^2) create serious numerical problems that restrict the exact Bayesian analysis to the cases with small or moderate n. These problems can be avoided for large n by fixing δ and A at some preliminary estimates in the way described in section 4.

Before using the MSF-SBEKK models and their approximate Bayesian analysis for large portfolios, we check the empirical adequacy of the proposed hybrid structures in the bivariate case, where exact Bayesian inference and model comparison can be performed even for unparsimonious specifications, like TSV or VECH(1,1).

3 The data and results of formal model comparison in the bivariate case

In order to formally compare the MSF-SBEKK models with other bivariate GARCH and SV specifications we consider almost the same set of basic models and use the same two data sets as Osiewalski, Pajor, Pipień (2007). The first data set consists of the official daily exchange rates of the National Bank of Poland (NBP fixing rates) for the US dollar and German mark in the period 1.02.1996 - 31.12.2001. The length of the modelled time series of their daily growth rates (logarithmic return rates) is 1482. The second data set consists of the daily quotations of the main index of the Warsaw Stock Exchange (WIG) and the S&P500 index of NYSE. We model 1727 logarithmic

returns from the period 8.01.1999-1.02.2006.
Basic descriptive characteristics of the daily growth (or return) rates expressed in percentage points are presented in Table 3. In the case of exchange rates, both series are highly non-Normal and they are quite strongly positively correlated. The other data set shows smaller deviations from Normality and much weaker correlation.
For each data set we use the VAR(1) framework (1). The overall ranking of competing Bayesian models and the decimal logarithms of the Bayes factors in favour of M_1 (i.e. the VAR(1)-TSV model), $\log_{10}(B_{1,j})$ for $j = 1, \ldots, 13$ (calculated for each data set using the Newton and Raftery method), are shown in Table 4. Only the results for the MSF-SBEKK models are new; the remaining ones were obtained by Osiewalski, Pajor, Pipień (2007), who present the Bayesian statistical methodology and particular priors used. Note that $B_{1,j}$ indicates how much better is the fit of M_1 (relative to M_j).

Table 3: Sample characteristics for the two data sets used

Time series (growth rates of:)	Length	average	st. dev.	skewness	kurtosis	correlation
USD/PLN 6.02.1996-31.12.2001	1482	0.03056	0.64173	0.41644	12.8484	0.56675
DEM/PLN 6.02.1996-31.12.2001		0.00290	0.72792	0.68855	12.9013	
WIG 8.01.1999-1.02.2006	1727	0.05589	1.35218	-0.11591	6.0776	0.17396
S&P500 8.01.1999-1.02.2006		0.00058	1.19683	0.09250	4.9096	

Table 4: Decimal logs of Bayes factors in favour of TSV (M_1)

Model VAR(1) + M_i	Number of estimated parameters (and latent variables)	USD/PLN - DEM/PLN		WIG - S&P500	
		Rank	$\log_{10}(B_{\mathrm{TSV},i})$	Rank	$\log_{10}(B_{\mathrm{TSV},i})$
TSV	18 ($+3T$)	1	**0**	1	**0**
JSV	15 ($+2T$)	2	**15**	2	**10.5**
BSV	14 ($+2T$)	12	**124**	3	**16.5**
MSF-SBEKK type I	14 ($+T$)	3	**31**	4	**19**
MSF-DCC	20 ($+T$)	4	**50**	5-6	**20.5**
MSF-SBEKK type II	14 ($+T$)	5	**77.5**	5-6	**20.5**
MSF	12 ($+T$)	6	**92**	7	**26.5**
t-BEKK(1,1)	19	7	**104.5**	11	**35**
t-VECH(1,1)	29	8	**106**	12-13	**36**
t-DiagVECH(1,1)	17	9-10	**119.5**	10	**33.5**
t-SBEKK(1,1)	13	9-10	**120**	12-13	**36**
t-DCC(1,1)	18	11	**122**	8-9	**31.5**
t-CCC(1,1)	15	13	**168.5**	8-9	**31**

The results in Table 4 show that the hybrid structures beat (in terms of the marginal

data density value) both their sub-cases and all MGARCH models considered here. The high ranks of the MSF-SBEKK type I model are particularly promising as it is relatively easy from the numerical perspective. Using just one latent process together with the SBEKK covariance structure seems a powerful modelling strategy. Of course, the use of more latent processes improves fit enormously, but seems infeasible for highly dimensional time series. A detailed discussion of empirical results is presented by Osiewalski, Pajor, Pipień (2007) for all the models, except the new ones. Let us remind here that the conditional Normality assumption for the MGARCH models is strongly rejected. In the case of exchange rates, the model-specific posterior distributions of ν, the degrees of freedom parameter, are concentrated in the interval (4, 5), while for the stock indices they are located close to 11 - 12, with standard deviations clearly excluding Normality.

4 Feasible approximate Bayesian inference in MSF-SBEKK models

For all n-variate volatility structures considered here, except MSF, we are not able to perform exact Bayesian analysis for large n. Now we focus on the MSF-SBEKK models that seem the most promising compromise between generality and parsimony. However, even in these models the conditional posteriors of δ and A (the parameters of dimension related to n^2) are non-standard. Rejection sampling or the Metropolis-Hastings algorithm within Gibbs steps would not be feasible for large n. Since δ is of no particular interest and obtaining its posterior distribution is not important, we can use its value resulting from the application of OLS to the VAR(1) system. We also suggest using the same estimate of A as in the case of the N-SBEKK structure (subsection 2.1); that is, A can be replaced by the empirical covariance matrix of the OLS residuals from the VAR(1) part. However, we can still apply the Gibbs sampler (based on the conditionals presented in subsection 2.2) for the remaining parameters, latent variables and future observations. This is the approximate Bayesian approach we propose for large n.

Formally, one can describe our approach as follows. Let κ and λ group the quantities of interest and nuisance parameters, respectively; in the MSF-SBEKK models $\lambda = (\delta, A)$. We are interested in obtaining the marginal posterior $p(\kappa|data) = \int_A p(\kappa|data, \lambda)\, p(\lambda|data)\, d\lambda$ and we are able to efficiently draw κ from its conditional posterior given λ, but we cannot draw λ. So we replace the marginal posterior of λ, $p(\lambda|data)$, with a very sharp (degenerate) distribution concentrated at some preliminary estimate of λ (say, $\widehat{\lambda}$). This results in using $p\left(\kappa|data, \lambda = \widehat{\lambda}\right)$, as if one conditioned on a data-based value of λ, instead of $p(\kappa|data)$.

In the bivariate cases considered in the previous section it is possible to compare exact and approximate Bayesian results. Thus, in Tables 5 and 6 we present the posterior means and standard deviations of the parameters in both cases. The decimal loga-

rithms of the data density values (both true and "conditional", corresponding to the parameters δ and A fixed at their estimates) are also shown. In this way we check how fixing δ and A at some preliminary estimates changes the Bayesian measure of fit. Big changes can be observed for exchange rates and much smaller ones for the stock data. In the type II model using our estimates leads to higher values of the data density. For financial applications it is important how the exact and approximate

Table 5: Posterior means (and standard deviations) of the parameters of the MSF-SBEKK models and logged data density values for the exchange rates ($T = 1482$)

parameter	MSF-SBEKK type I		MSF-SBEKK type II	
	exact	approximate	exact	approximate
δ_{01}	0.044 (0.009)	*0.030* (0)	0.042 (0.009)	*0.030* (0)
δ_{02}	-0.005 (0.010)	*0.002* (0)	0.003 (0.010)	*0.002* (0)
δ_{11}	-0.020 (0.025)	*0.017* (0)	-0.005 (0.026)	*0.017* (0)
δ_{12}	-0.012 (0.026)	*0.023* (0)	0.001 (0.027)	*0.023* (0)
δ_{21}	-0.012 (0.021)	*-0.081* (0)	-0.010 (0.022)	*-0.081* (0)
δ_{22}	-0.040 (0.025)	*-0.114* (0)	-0.039 (0.026)	*-0.114* (0)
a_{11}	0.153 (0.029)	*0.409* (0)	0.378 (0.140)	*0.409* (0)
a_{12}	-0.053 (0.018)	*0.261* (0)	0.055 (0.056)	*0.261* (0)
a_{22}	0.174 (0.034)	*0.524* (0)	0.480 (0.195)	*0.524* (0)
φ	0.411 (0.086)	**0.332 (0.082)**	0.915 (0.028)	**0.857 (0.032)**
σ^2	0.540 (0.070)	**0.549 (0.062)**	0.173 (0.038)	**0.229 (0.043)**
β	0.084 (0.013)	**0.096 (0.010)**	0.057 (0.015)	**0.034 (0.009)**
γ	0.878 (0.015)	**0.902 (0.011)**	0.925 (0.025)	**0.964 (0.010)**
$\beta + \gamma$	0.962 (0.008)	**0.998 (0.001)**	0.982 (0.012)	**0.999 (0.001)**
h_0	0.053 (0.051)	**0.082 (0.056)**	0.074 (0.076)	**0.062 (0.036)**
$\log_{10}(p(data))$	$d = -852.6$	$d - 20$	$f = -899.2$	$f + 19.3$

Table 6: Posterior means (and standard deviations) of the parameters of the MSF-SBEKK models and logged data density values for the stock indices ($T = 1727$)

parameter	MSF-SBEKK type I		MSF-SBEKK type II	
	exact	approximate	exact	approximate
δ_{01}	0.072 (0.026)	*0.056* (0)	0.071 (0.026)	*0.056* (0)
δ_{02}	0.028 (0.022)	*0.001* (0)	0.028 (0.023)	*0.001* (0)
δ_{11}	0.015 (0.024)	*-0.008* (0)	0.017 (0.024)	*-0.008* (0)
δ_{12}	0.012 (0.020)	*-0.006* (0)	0.015 (0.020)	*-0.006* (0)
δ_{21}	0.302 (0.027)	*0.371* (0)	0.305 (0.027)	*0.371* (0)
δ_{22}	-0.022 (0.026)	*-0.013* (0)	-0.023 (0.025)	*-0.013* (0)
a_{11}	1.127 (0.267)	*1.633* (0)	1.308 (0.254)	*1.633* (0)
a_{12}	0.159 (0.104)	*0.289* (0)	0.240 (0.110)	*0.289* (0)
a_{22}	0.729 (0.176)	*1.432* (0)	1.009 (0.209)	*1.432* (0)
φ	0.872 (0.156)	**0.934 (0.045)**	0.984 (0.007)	**0.987 (0.005)**
σ^2	0.036 (0.041)	**0.020 (0.012)**	0.012 (0.004)	**0.011 (0.003)**
β	0.021 (0.006)	**0.023 (0.005)**	0.020 (0.004)	**0.020 (0.004))**
γ	0.970 (0.007)	**0.971 (0.006)**	0.970 (0.007)	**0.966 (0.009)**
$\beta + \gamma$	0.991 (0.003)	**0.994 (0.003)**	0.990 (0.005)	**0.987 (0.006)**
h_0	2.881 (1.026)	**2.856 (1.008)**	3.606 (0.712)	**3.458 (1.052)**
$\log_{10}(p(data))$	$d = -2278.3$	$d - 0.3$	$f = -2279.8$	$f + 3.3$

posterior results on volatility and conditional correlation may differ. In Figures 1 and 2 we present the results on the time-varying conditional correlation coefficient $\rho_{12,t} = \frac{\omega_{12,t}}{\sqrt{\omega_{11,t}\omega_{22,t}}}$ based on the form of Ω_t, the matrix defining the particular MSF-SBEKK specification. In Figure 3 we show the posterior means of the conditional standard deviation $\sigma_{i,t} = \sqrt{g_t \omega_{ii,t}}$ measuring volatility of the first time series ($i = 1$). In Table 7 basic inferences on conditional correlation and volatility are summarized by time averages of posterior means and standard deviations. The overall impression is that the differences between the model structures and between the exact and approximate approaches are not crucial. In particular, the dynamics of volatility and correlation is described in a very similar way; the empirical correlation coefficient between two sequences of competing estimates is always very high, especially for the stock data. Even though exact and approximate posterior inferences may differ for the basic model parameters (see Table 5, 6), they remain quite close for volatility and conditional correlation.

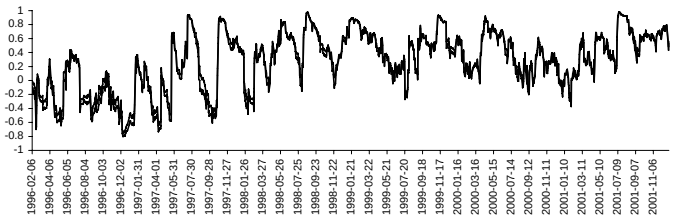
Our examples indicate that there is no substantial difference between the type I and II models, but the numerical demands of the type I model are smaller. We use our own computer code in GAUSS, based on earlier codes written by the second author; see Pajor (2009). The results were obtained using at least 100 000 burnt-in and 500 000 final Gibbs passes, which took about 20 hours of Intel Core2 CPU X6800 (2,93GHz) in the case of the type I model. Generating less draws at the final stage would have saved time at the cost of precision loss. The time needed for carrying out all computations is 5 to 9 times larger in the case of the type II model, where the number of the evaluations of the n-variate Normal density is $O(T^2)$, while it is $O(T)$ for the type I specification. Also, the Gibbs sampler seems to converge faster for the type I model, so less burnt-in Gibbs passes are needed. Hence we recommend to use the simpler type I structure.

Table 7: Conditional correlation and volatility: time averages of the posterior means and standard deviations in the MSF-SBEKK models

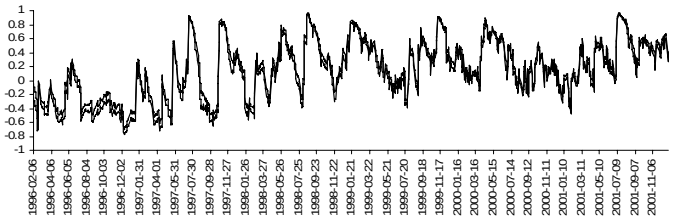
Data set	Average of	type I		type II	
		approximate	exact	approximate	exact
exchange rates	$E(\rho_{12,t}\|data)$	0.2800	0.1458	0.2117	0.1610
	$D(\rho_{12,t}\|data)$	0.0244	0.0378	0.0623	0.0771
	$E(\sigma_{1,t}\|data)$	0.5957	0.5594	0.5320	0.5353
	$D(\sigma_{1,t}\|data)$	0.1932	0.1812	0.1379	0.1280
	$E(\sigma_{2,t}\|data)$	0.6645	0.6191	0.5951	0.6001
	$D(\sigma_{2,t}\|data)$	0.2159	0.2008	0.1544	0.1436
stock indices	$E(\rho_{12,t}\|data)$	0.1899	0.1871	0.1855	0.1900
	$D(\rho_{12,t}\|data)$	0.0154	0.0293	0.0211	0.0324
	$E(\sigma_{1,t}\|data)$	1.2269	1.2314	1.2114	1.2252
	$D(\sigma_{1,t}\|data)$	0.1602	0.1789	0.1353	0.1430
	$E(\sigma_{2,t}\|data)$	1.1429	1.1116	1.1274	1.1169
	$D(\sigma_{2,t}\|data)$	0.1496	0.1623	0.1263	0.1303

Figure 1: Conditional correlation (posterior mean ± standard deviation), exchange rates

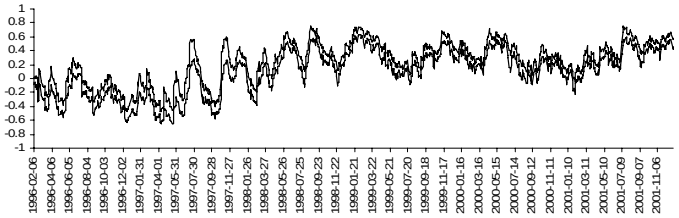
MSF-SBEKK type I, approximate:



MSF-SBEKK type I, exact:



MSF-SBEKK type II, exact:



MSF-SBEKK type II, approximate:

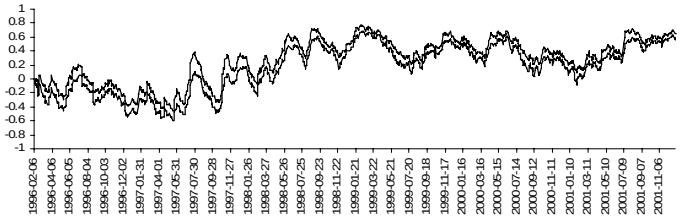
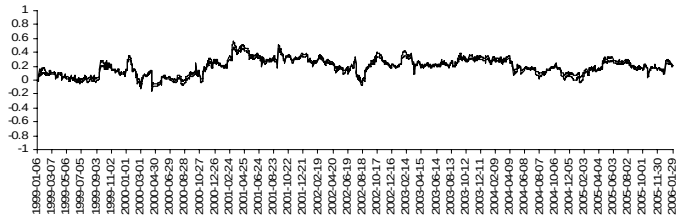
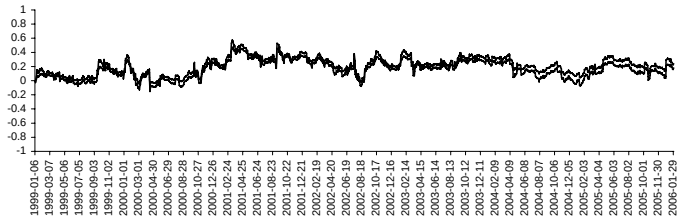


Figure 2: Conditional correlation (posterior mean ± standard deviation); WIG, S&P500

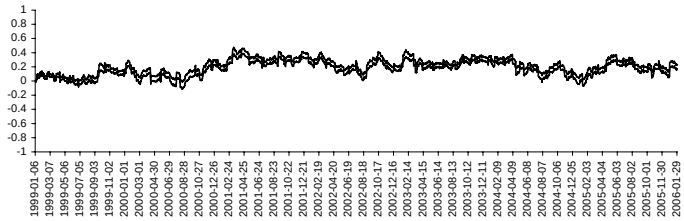
MSF-SBEKK type I, approximate:



MSF-SBEKK type I, exact:



MSF-SBEKK type II, exact:



MSF-SBEKK type II, approximate:

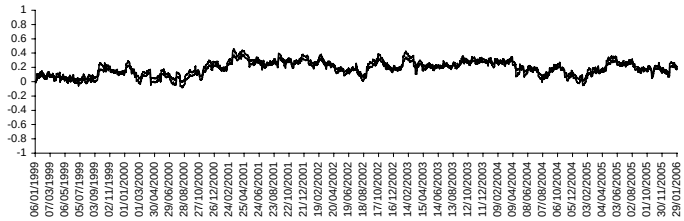
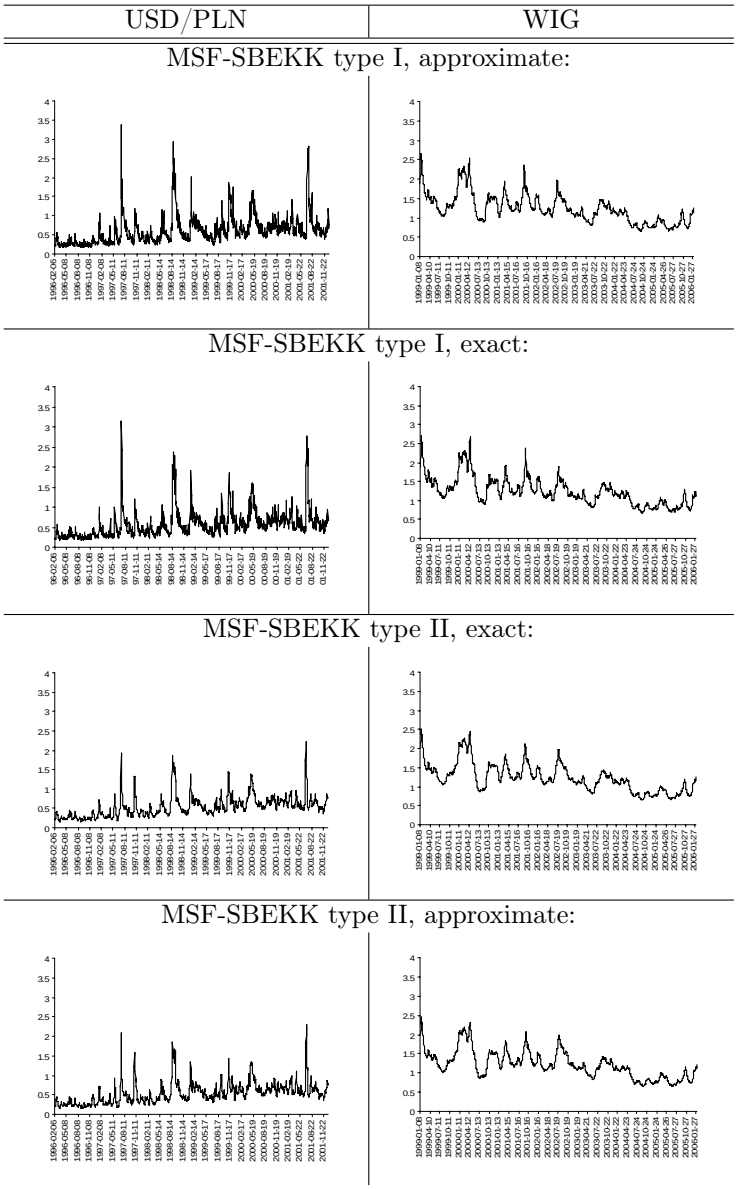


Figure 3: Volatility measure $\sigma_{1,t}$ (posterior means)



5 An example of approximate Bayesian inference for a large number of assets

In order to show how our n-variate MSF-SBEKK models work for large n, we present selected posterior results based on stock data representing 34 companies. Summary statistics for their daily logarithmic returns in the period 30.01.2003-29.08.2007 are shown in Table 8; on August 29, 2007 companies number 1-23 were included in mWIG40 and number 24-34 in WIG20, two important indices of the Warsaw Stock Exchange. The approximate Bayesian approach (using the proposed data-based values of the highly dimensional matrix parameters) is applied. The posterior results on volatility and conditional correlation for the prices of BDX (Budimex) and AGO (Agora) are presented in some detail. Their logarithmic return rates and the posterior means of the conditional standard deviations $\sigma_{i,t}$ are plotted in Figures 4 and 5, while the time averages of the posterior means and standard deviations of $\sigma_{i,t}$ are given (for all companies) in the last columns of Table 8. Note that in both models these averages are similar, with average $D\left(\sigma_{i,t}|data\right)$ always smaller for the type II model; the same applies to the posterior moments for the main MSF-SBEKK parameters (see Table 9). The results on conditional correlation $\rho_{ij,t}$ between BDX and AGO are shown in Figure 6; the time averages of the posterior means (standard deviations) of $\rho_{ij,t}$ for this pair are 0.188 (0.002) and 0.184 (0.006) in the type I and type II model, respectively. In both models the dynamics of volatility and conditional correlation is described in a similar way; the empirical correlation coefficient between two sequences of competing estimates is always above 0.9. For the sequences of $E\left(\sigma_{i,t}|data\right)$ obtained in the models of type I and II, we get 0.992 in the case of BDX and 0.996 in the case of AGO; for the sequences of $E\left(\rho_{2,24;t}|data\right)$ we get 0.916.

The results presented here are obtained using 200 000 burnt-in and 200 000 final Gibbs passes. In the type I model convergence to the stationary distribution seems fast - about 40 000 burnt-in passes would have been sufficient; in terms of computational time, this example (with $n = 34$) is almost 10 times more demanding than the bivariate one. For the type II model, the time needed for carrying out all the computations (with $\lambda = 80$, see Appendix) is 3 times larger than for the type I structure; convergence to the stationary distribution is observed after about 130 000 Gibbs passes.

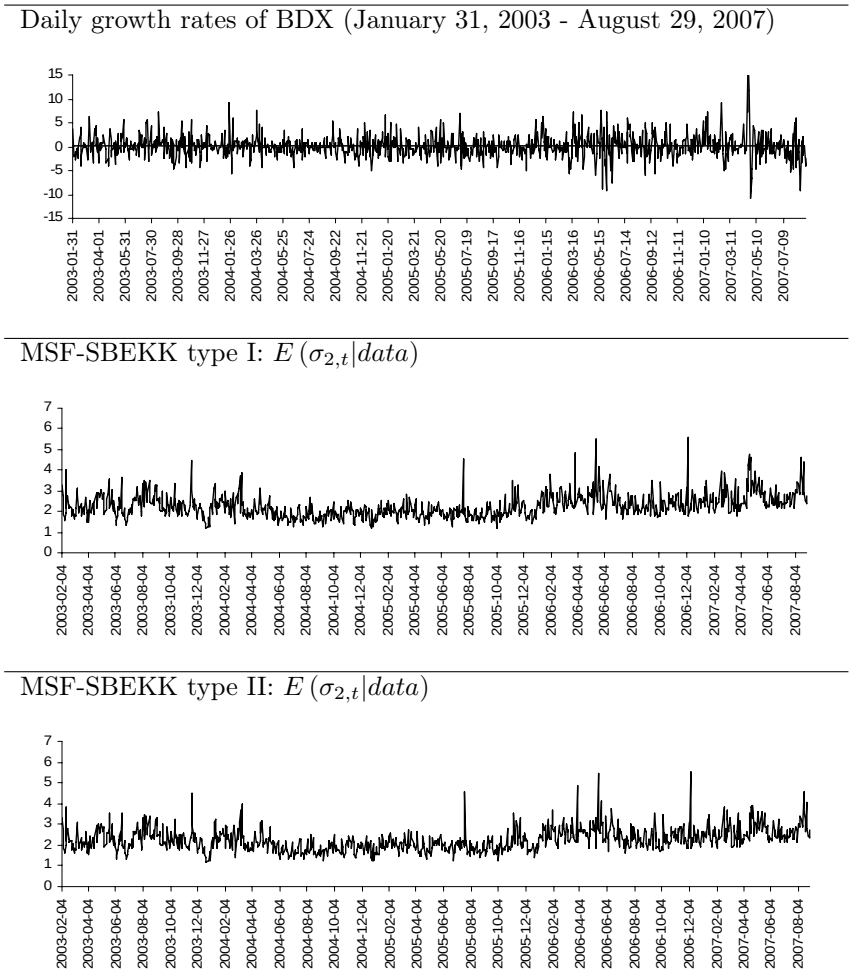
Table 8: Sample characteristics for the data set used ($T = 1149$) and results on volatility

i	company	average	st. dev.	kurtosis	min	max	average $E(\sigma_{i,t}\|data)$ type I	average $D(\sigma_{i,t}\|data)$ type I	average $E(\sigma_{i,t}\|data)$ type II	average $D(\sigma_{i,t}\|data)$ type II
1	BPH	0.112	1.906	5.562	-10.566	9.444	1.819	0.193	1.816	0.184
2	BDX	0.097	2.395	10.843	-10.807	21.035	2.249	0.238	2.241	0.227
3	DUD	0.142	2.731	84.964	-47.505	12.936	2.576	0.273	2.546	0.258
4	ECH	0.203	1.889	6.584	-8.278	8.961	1.790	0.190	1.784	0.181
5	EMP	0.206	2.588	74.084	-15.575	43.621	2.467	0.262	2.443	0.248
6	GRJ	0.188	2.176	10.395	-12.516	15.453	2.081	0.220	2.072	0.210
7	BHW	0.053	1.583	28.478	-20.096	8.734	1.504	0.159	1.495	0.152
8	BSK	0.094	1.373	7.070	-6.432	6.652	1.303	0.138	1.300	0.132
9	KTY	0.113	1.939	5.921	-11.823	9.019	1.845	0.195	1.843	0.187
10	KPX	0.318	3.429	19.599	-15.082	35.398	3.234	0.343	3.207	0.326
11	KRB	0.048	1.850	20.879	-21.472	8.961	1.760	0.186	1.745	0.177
12	MCI	0.373	3.734	11.532	-20.373	33.178	3.553	0.376	3.529	0.358
13	MIL	0.131	2.236	9.139	-12.783	14.458	2.132	0.226	2.121	0.215
14	MSX	0.165	3.716	12.516	-24.381	28.768	3.546	0.376	3.511	0.356
15	MSZ	0.226	4.248	8.241	-25.300	23.974	4.029	0.427	4.015	0.408
16	NET	0.023	1.939	16.115	-20.567	8.444	1.834	0.194	1.821	0.185
17	EMF	0.091	2.999	15.846	-22.012	24.686	2.870	0.304	2.858	0.290
18	ORB	0.122	2.050	7.959	-15.558	10.178	1.952	0.207	1.938	0.197
19	PGF	0.097	2.083	16.043	-10.536	21.767	1.960	0.208	1.950	0.198
20	PRC	0.027	4.948	11.569	-28.768	34.484	4.689	0.497	4.644	0.472
21	STX	0.105	3.762	12.627	-29.523	23.863	3.581	0.379	3.548	0.360
22	STP	0.395	2.742	11.824	-9.237	23.309	2.607	0.276	2.597	0.264
23	VST	0.325	2.758	9.848	-10.536	18.666	2.628	0.278	2.624	0.266
24	**AGO**	**0.009**	**2.070**	**5.495**	**-11.955**	**8.072**	**1.964**	**0.208**	**1.960**	**0.199**
25	**BRE**	**0.167**	**1.895**	**5.010**	**-7.633**	**8.898**	**1.802**	**0.191**	**1.805**	**0.183**
26	**BZW**	**0.108**	**2.066**	**4.092**	**-8.259**	**7.496**	**1.963**	**0.208**	**1.960**	**0.199**
27	**CST**	**0.191**	**1.950**	**9.507**	**-10.488**	**13.262**	**1.844**	**0.195**	**1.835**	**0.186**
28	**GTN**	**0.209**	**3.455**	**35.036**	**-45.392**	**24.613**	**3.312**	**0.351**	**3.275**	**0.334**
29	**KGH**	**0.182**	**2.543**	**5.684**	**-15.590**	**9.093**	**2.426**	**0.257**	**2.421**	**0.245**
30	**PEO**	**0.088**	**1.964**	**4.751**	**-6.579**	**11.919**	**1.869**	**0.198**	**1.870**	**0.190**
31	**PKN**	**0.106**	**1.908**	**3.923**	**-9.298**	**7.746**	**1.815**	**0.192**	**1.818**	**0.184**
32	**PXM**	**0.350**	**2.778**	**7.447**	**-11.725**	**16.252**	**2.630**	**0.279**	**2.619**	**0.266**
33	**PND**	**0.244**	**4.090**	**34.971**	**-53.870**	**28.395**	**3.903**	**0.414**	**3.867**	**0.393**
34	**TPS**	**0.044**	**1.798**	**3.733**	**-8.359**	**5.617**	**1.712**	**0.181**	**1.717**	**0.174**

Table 9: Posterior means (and standard deviations) of the main MSF-SBEKK parameters

	φ	σ^2	β	γ	$\beta+\gamma$
type I model	0.4995 (0.0339)	0.1874 (0.0113)	0.0167 (0.0014)	0.8523 (0.0180)	0.8690 (0.0168)
type II model	0.5572 (0.0304)	0.1775 (0.0102)	0.0152 (0.0014)	0.8824 (0.0163)	0.8976 (0.0152)

Figure 4: Daily growth rates of BDX prices and their volatility estimates

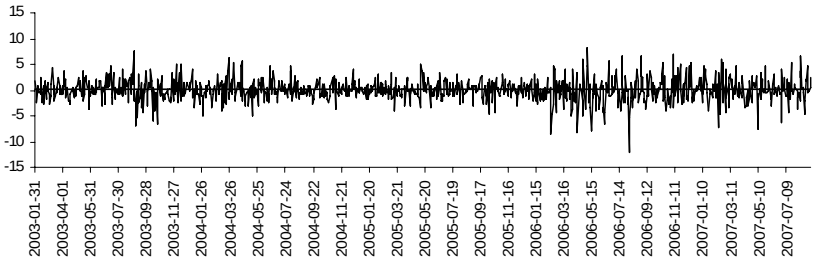


6 Concluding remarks

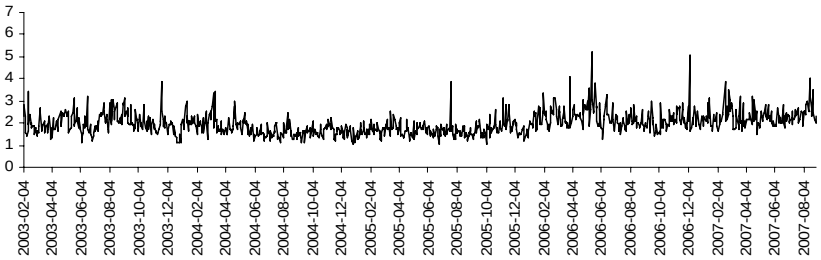
MSV specifications can have much more explanatory power (as measured by Bayes factors) than MGARCH models, but they are more difficult to estimate due to the presence of latent processes. Only the Bayesian approach (equipped with MCMC simulation methods) can simultaneously and efficiently deal with unknown parameters, latent variables and future observables, but not in too complicated MSV models of n-variate time series with large n.

Figure 5: Daily growth rates of AGO prices and their volatility estimates

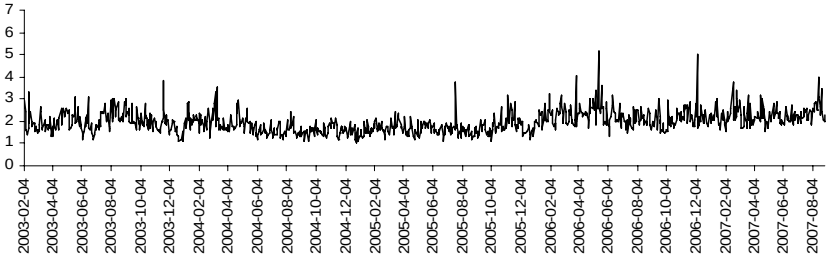
Daily growth rates of AGO (January 31, 2003 - August 29, 2007)



MSF-SBEKK type I: $E\left(\sigma_{24,t}|data\right)$



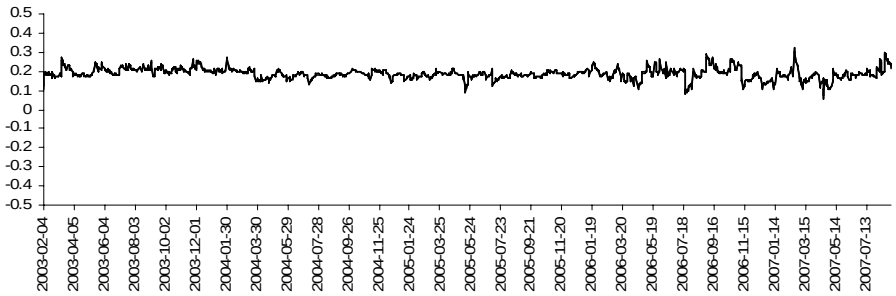
MSF-SBEKK type II: $E\left(\sigma_{24,t}|data\right)$



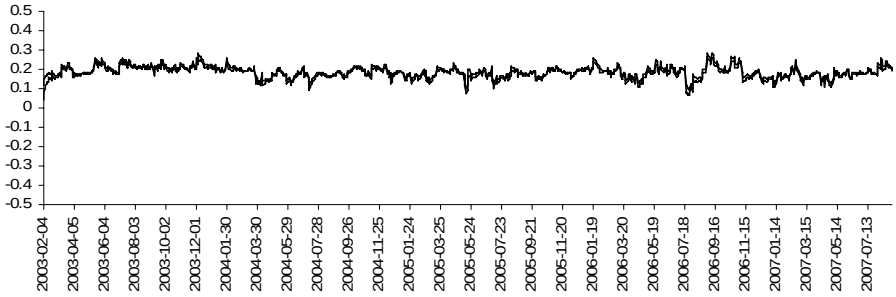
In order to be able to analyse highly multivariate portfolios, we need relatively simple n-variate volatility models that use one latent process and can be easily estimated, but have non-trivial covariance structure. The VAR(1) - MSF-SBEKK type I model meets these conditions. Its Bayesian analysis can rely on preliminary estimates for highly dimensional nuisance parameters and can be performed for large n. Our empirical examples demonstrate both high explanatory power ($n = 2$) and practical

Figure 6: Conditional correlation between BDX and AGO,
$E(\rho_{2,24;t}|data) \pm D(\rho_{2,24;t}|data)$

type I model:



type II model:



usefulness ($n = 2$, $n = 34$) of this new model. In further research one could apply it in the portfolio optimization problems, along the lines presented for MSV models by Pajor (2009). Also, formal Bayesian comparison between our model (that uses past data in the conditional covariance matrix through the SBEKK structure) and the factor models proposed by Chib, Nardari, Shephard (2006) would be interesting.

References

[1] Bauwens L., Laurent S., Rombouts J.V.K., (2006), Multivariate GARCH models: A survey, *Journal of Applied Econometrics* 21, 79-109.

[2] Chib S., Nardari F., Shephard N., (2006), Analysis of high dimensional multivariate stochastic volatility models, *Journal of Econometrics* 134, 341-371.

[3] Chan D., Kohn R., Kirby Ch., (2006), Multivariate Stochastic Volatility models with correlated errors, *Econometric Review* 25, 245-274.

[4] 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.

[5] Fleming J., Kirby Ch., (2003), A closer look at the relation between GARCH and stochastic autoregressive volatility, *Journal of Financial Econometrics* 1, 365-419.

[6] Jacquier E., Polson N., Rossi P., (1995), Models and prior distributions for multivariate stochastic volatility, *technical report*, University of Chicago, Graduate School of Business.

[7] Osiewalski J., (2009), New hybrid models of multivariate volatility (a Bayesian perspective), *Przegląd Statystyczny (Statistical Review)* 56 (no.1), 15-22.

[8] Osiewalski J., Pajor A., (2007), Flexibility and parsimony in multivariate financial modelling: a hybrid bivariate DCC-SV model, [in:] *Financial Markets. Principles of Modeling, Forecasting and Decision-Making* (FindEcon Monograph Series No.3), [ed.:] W. Milo, P. Wdowiński, Łódź University Press, Łódź, 11-26.

[9] Osiewalski J., Pajor A., Pipień M., (2007), Bayesian comparison of bivariate GARCH, SV and hybrid models, [in:] *Proceedings of the 33rd International Conference, MACROMODELS'2006*, [ed.:] W. Welfe, A. Welfe, Łódź, 247-277.

[10] Osiewalski J., Pipień M., (2004), Bayesian comparison of bivariate ARCH-type models for the main exchange rates in Poland, *Journal of Econometrics* 123, 371-391.

[11] Pajor A., (2005a), Bayesian analysis of stochastic volatility model and portfolio allocation, *Acta Universitatis Lodziensis - Folia Oeconomica* 192, 229-249.

[12] Pajor A., (2005b), Bayesian comparison of bivariate SV models for two related time series, *Acta Universitatis Lodziensis - Folia Oeconomica* 190, 177-196.

[13] Pajor A., (2009), Wielowymiarowe procesy wariancji stochastycznej w ekonometrii finansowej. Ujęcie bayesowskie (Multivariate Stochastic Variance Processes in Financial Econometrics. Bayesian approach, in Polish) Cracow University of Economics, Kraków (forthcoming).

[14] Shephard N., Pitt M.K., (1997), Likelihood analysis of non-Gaussian measurement time series, *Biometrika* 84, 653-667.

[15] Smith M., Pitts A., (2006), Foreign exchange intervention by the Bank of Japan: Bayesian analysis using a bivariate Stochastic Volatility model, *Econometric Review* 25, 425-451.

[16] Tsay R.S., (2005), *Analysis of Financial Time Series* (2nd edition), Wiley, New York.

Appendix - Generating latent variables in the MSF-SBEKK model of type II

Let $\ln g = (\ln g_1, \ldots, \ln g_T)'$; our target density is:

$$p\left(\ln g_1, \ldots, \ln g_T | r_1, \ldots, r_T, \delta, A, \beta, \gamma, \varphi, \sigma_g^{-2}, h_0\right) \propto \prod_{t=1}^{T} f_N^n\left(r_t|\mu_t, g_t\Omega_t\right) f_N^1\left(\ln g_t|\varphi \ln g_{t-1}, \sigma_g^2\right).$$

Since it is not practical to sample directly from this density, we use the Metropolis-Hastings (M-H) algorithm. Following Shephard and Pitt (1997) and Smith and Pitts (2006), we propose to partition the elements of $\ln g$ into contiguous blocks and generate one block at a time from a Gaussian distribution. Let $\ln g_{t,a:b} = (\ln g_a, \ldots, \ln g_b)$ be a block of the log volatility vector with $1 \leq a < b \leq T$. Then

$$p\left(\ln g_{t,a:b} | \ln g_{t\backslash a:b}, r_1, \ldots, r_T, \delta, A, \beta, \gamma, \varphi, \sigma_g^{-2}, h_0\right) \propto \prod_{t=a}^{T} f_N^n\left(r_t|\mu_t, g_t\Omega_t\right) \prod_{t=a}^{b+1} f_N^1\left(\ln g_t|\varphi \ln g_{t-1}, \sigma_g^2\right),$$

for $1 \leq a < b < T$, and

$$p\left(\ln g_{t,a:b} | \ln g_{t\backslash a:b}, r_1, \ldots, r_T, \delta, A, \beta, \gamma, \varphi, \sigma_g^{-2}, h_0\right) \propto \prod_{t=a}^{T} f_N^n\left(r_t|\mu_t, g_t\Omega_t\right) \prod_{t=a}^{b} f_N^1\left(\ln g_t|\varphi \ln g_{t-1}, \sigma_g^2\right),$$

for $1 \leq a < b = T$,
where $\ln g_{t\backslash a:b} = \ln g \backslash \ln g_{a:b}$ (i.e. $\ln g_{t\backslash a:b} = (\ln g_1, \ldots, \ln g_{a-1}, lng_{b+1}, \ldots, \ln g_T)'$ for $1 < a < b < T$).
The proposal density is $N\left(\ln g_{t,a:b}^{old}, B\right)$, where $B = \frac{(k+1)LL' - L\mathbf{1}\mathbf{1}'L'}{k+1}\sigma_g^2$ (see Fleming and Kirby 2003), $\mathbf{1}$ is a $k \times 1$ vector of ones, L is a lower-triangular $k \times k$ matrix of ones, $k = b - a + 1$, $\ln g_{t,a:b}^{old}$ is the iterate of $\ln g_{t,a:b}$ from the previous sweep of the sampler. As in Smith and Pitts (2006), the block size (k) is drawn at the beginning of each sweep from a Poisson distribution with mean λ. In our empirical examples $\lambda = 25$ for $n = 2$ and $\lambda = 80$ for $n = 34$.
It is important to stress that Smith and Pitts (2006), and Chan, Kohn and Kirby (2006) use (for the M-H step) the multivariate Gaussian proposal density obtained from a quadratic approximation to

$$l\left(\ln g_{t,a:b}\right) = \ln p\left(\ln g_{t,a:b} | \ln g_{t\backslash a:b}, r_1, \ldots, r_T, \delta, A, \beta, \gamma, \varphi, \sigma_g^{-2}, h_0\right):$$

$$N\left(\widehat{\ln g_{t,a:b}},\left[-\frac{\partial^2 l\left(\widehat{\ln g_{t,a:b}}\right)}{\partial \ln g_{t,a:b}\partial \ln g_{t,a:b}'}\right]^{-1}\right),$$

where $\widehat{\ln g_{t,a:b}}$ is the mode of $l\left(\ln g_{t,a:b}\right)$ that is found numerically. In our model, however, this approach would be computationally too intensive and would make the sampling scheme very slow.