Central European Journal of Economic Modelling and Econometrics

On Sensitivity of Inference in Bayesian MSF-MGARCH Models

Jacek Osiewalski* and Anna Pajor†

Submitted: 16.05.2019, Accepted: 5.09.2019

Abstract

Hybrid MSV-MGARCH models, in particular the MSF-SBEKK specification, proved useful in multivariate modelling of returns on financial and commodity markets. The initial MSF-MGARCH structure, called LN-MSF-MGARCH here, is obtained by multiplying the MGARCH conditional covariance matrix H_t by a scalar random variable g_t such that $\{\ln g_t, t \in Z\}$ is a Gaussian AR(1) latent process with auto-regression parameter φ. Here we also consider an IG-MSF-MGARCH specification, which is a hybrid generalisation of conditionally Student t MGARCH models, since the latent process $\{g_t\}$ is no longer marginally log-normal (LN), but for $\varphi = 0$ it leads to an inverted gamma (IG) distribution for g_t and to the t-MGARCH case. If $\varphi \neq 0$, the latent variables g_t are dependent, so (in comparison to the t-MGARCH specification) we get an additional source of dependence and one more parameter. Due to the existence of latent processes, the Bayesian approach, equipped with MCMC simulation techniques, is a natural and feasible statistical tool to deal with MSF-MGARCH models. In this paper we show how the distributional assumptions for the latent process together with the specification of the prior density for its parameters affect posterior results, in particular the ones related to adequacy of the t-MGARCH model. Our empirical findings demonstrate sensitivity of inference on the latent process and its parameters, but, fortunately, neither on volatility of the returns nor on their conditional correlation. The new IG-MSF-MGARCH specification is based on a more volatile latent process than the older LN-MSF-MGARCH structure, so the new one may lead to lower values of φ – even so low that they can justify the popular t-MGARCH model.

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

JEL Classification: C11, C32, C51

*Cracow University of Economics; e-mail: eeosiewa@cyf-kr.edu.pl; ORCID: 0000-0002-6710-6825
†Cracow University of Economics and Jagiellonian University in Kraków; e-mail: eopajor@cyf-kr.edu.pl; ORCID: 0000-0001-5643-0649

1 Introduction

In modelling of financial time series, hybrid MSV-MGARCH models were introduced in order to use relatively simple model structures that exploit advantages of both model classes: flexibility of the Multivariate Stochastic Volatility (MSV) class, where volatility is modelled by latent stochastic processes, and relative simplicity of the Multivariate GARCH (MGARCH) class; see Osiewalski and Pajor (2007, 2009) and Osiewalski and Osiewalski (2016). In their first attempt, Osiewalski and Pajor (2007) used one latent process and the Dynamic Conditional Correlation (DCC) covariance structure proposed by Engle (2002). However, Osiewalski (2009) and Osiewalski and Pajor (2009) suggested an even simpler model, also based on one latent process, but with the scalar BEKK (Baba, Engle, Kraft, Kroner, 1989) covariance structure. The parsimonious hybrid Multiplicative Stochastic Factor – Scalar BEKK (MSF-SBEKK) specification has been recognized in the literature (see Teräsvirta, 2012; Amado and Teräsvirta, 2013; Carriero, Clark and Marcellino, 2016) and proved useful in multivariate modelling of returns on financial and commodity markets (see Pajor, 2010, 2014; Osiewalski and Pajor, 2010; Pajor and Osiewalski, 2012; Osiewalski and Osiewalski, 2013, 2016; Pajor and Wróblewska, 2017; Wróblewska and Pajor, 2019). Initially proposed MSF-MGARCH models were built using a conditionally normal MGARCH process and multiplying its conditional covariance matrix H_t by g_t such that $\ln g_t$ follows a Gaussian AR(1) process with auto-regression parameter φ. If $\varphi = 0$, then such MSF-MGARCH case reduces to the MGARCH process with the conditional distribution being a continuous mixture of multivariate normal distributions with covariance matrices $g_t H_t$ and g_t log-normally (LN) distributed. In their conference paper, Osiewalski and Pajor (2018) proposed a natural hybrid extension of popular MGARCH models with the Student t conditional distribution. The new model is obtained by multiplying H_t by random variable g_t coming from such latent process (with auto-regression parameter φ) that, for $\varphi = 0$, g_t has an inverted gamma (IG) distribution and leads to the t-MGARCH specification, where the conditional distribution can be represented as a continuous mixture of multivariate normal distributions with covariance matrices $g_t H_t$ and an IG distribution of g_t. If $\varphi \neq 0$, the latent variables g_t are dependent, so (in comparison to the t-MGARCH model) in the new model of the observed time series we get an additional source of dependence and one more parameter. In fact, we could construct as many MSF-MGARCH specifications as there are distributions of latent g_t under $\varphi = 0$. In order to distinguish between the alternative MSF structures, we now apply notation LN-MSF for the one based on the log-normal distribution, and IG-MSF for the new one, based on inverted gamma innovations. Osiewalski and Pajor (2018) used the scalar BEKK (SBEKK) specification as the MGARCH structure and showed how to estimate the IG-MSF-SBEKK model using the Bayesian approach, equipped with Markov Chain Monte Carlo (MCMC) simulation tools. Two empirical examples were also presented in order to illustrate the hybrid extension of the t-SBEKK model and to compare its posterior results to the ones obtained in the LN-MSF-SBEKK case.

However, in Osiewalski and Pajor (2018) some arbitrarily chosen prior structures for the parameters of the latent process in each model were adopted. In this paper we present an empirical example, which shows the impact of both the distributional assumptions for the latent process and the specification of the prior density for its parameters on posterior results. Our empirical findings demonstrate sensitivity of posterior inference on the parameters of the latent process, much less sensitivity of the posterior for the latent process itself, and robustness of posterior inference on volatility and conditional correlation.

In the next section the general form of the MSV-MGARCH model as well as its special LN-MSF-MGARCH and IG-MSF-MGARCH cases are presented. In Section 3 it is briefly shown how to simulate the posterior distribution in the Bayesian IG-MSF-SBEKK model based on a more general prior structure than in Osiewalski and Pajor (2018). In Section 4 alternative prior assumptions for the parameters of the latent process in the LN-MSF-MGARCH and IG-MSF-MGARCH models are considered. In Section 5 our empirical example is shown; it serves not only to illustrate the hybrid extension of the t-SBEKK specification and its validity, but mainly to compare the posterior results obtained in the IG-MSF-MGARCH case to their counterparts in the LN-MSF-SBEKK model and to examine sensitivity of posterior results with respect to the prior assumptions on latent processes in both hybrid models. Concluding remarks are stated in Section 6.

2 Hybrid n-variate volatility specifications

Assume there are n assets. We denote by $r_t = (r_{t1} \dots r_{tn})$ n-variate observations on their logarithmic return rates, and we model them using the basic VAR(1) framework:

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

where δ_0 and Δ are, respectively, $n \times 1$ and $n \times n$ matrix parameters, and T is the length of the observed time series. The hybrid MSV-MGARCH model class for the disturbance term ε_t is defined by the following equality:

$$\varepsilon_t = \zeta_t H_t^{1/2} G_t^{1/2}, \tag{2}$$

where: $\{\zeta_t\}$ is a strict n-variate white noise with unit covariance matrix, $\{\zeta_t\} \sim iiD^{(n)}(0, I_n)$; H_t and G_t are square matrices of order n, symmetric and positive definite for each t; H_t is a non-constant function of the past of ε_t and corresponds to the conditional covariance matrix of some MGARCH specification; G_t is a non-constant function of a (scalar or vector) stochastic latent process $\{g_t\}$, which is non-trivial (i.e., constituted of variables g_t dependent over time); see Osiewalski and Osiewalski (2016). Under (1) and (2), the conditional distribution of r_t (given the past of r_t and the current latent variable g_t) is determined by the distribution of ζ_t; it has mean vector $\mu_t = \delta_0 + r_{t-1}\Delta$ and covariance matrix $\Sigma_t = G_t^{1/2\prime} H_t G_t^{1/2}$, which

depends on both g_t and the past of r_t.
Building upon an idea presented by Osiewalski and Osiewalski (2016), we consider a useful subclass of the MSV-MGARCH class. This subclass corresponds to the Gaussian white noise $\{\zeta_t\}$ and positive-valued scalar latent processes $\{g_t\}$ such that $G_t = g_t I_n$ and

$$\ln g_t = \varphi \ln g_{t-1} + \ln \gamma_t, \quad (3)$$

where $\zeta_t \perp \gamma_s$ for all $t, s \in \{1, \ldots, T\}$, $0 < |\varphi| < 1$ and $\{\gamma_t\}$ is a sequence of independent positive random variables with the same distribution belonging to a specific parametric family. The simplest MSV structure, called Multiplicative Stochastic Factor (MSF) by Osiewalski and Pajor (2009), is based on the assumption that $\{\ln \gamma_t\}$ is a Gaussian white noise with unknown variance σ^2. Since γ_t is log-normal, such model structure is now called LN-MSF-MGARCH. In such model, (3) represents a two-parameter family of stationary and causal Gaussian AR(1) processes and the marginal distribution of g_t is log-normal, so the distribution of r_t given its past is the scale mixture of $N(\mu_t, g_t H_t)$ distributions with log-normal g_t. The mixing distribution depends on φ, and remains log-normal for $\varphi = 0$, the value leading to the MGARCH model with a specific ellipsoidal conditional distribution – the log-normal scale mixture of normal distributions. Osiewalski and Pajor (2018) extend this basic case and consider other latent processes (3), corresponding to different parametric distribution classes of γ_t. In particular, the inverted gamma γ_t in (3) leads to the IG-MSF-MGARCH hybrid model, where $\{\ln \gamma_t\}$ need not have zero mean, so $\{\ln g_t\}$ need not be a white noise process. Assume that γ_t^{-1} is gamma distributed with mean 1 and variance $2/v$, i.e. $\{\gamma_t\} \sim iiIG(v/2, v/2)$, where $v > 2$. For $u_t = \ln \gamma_t$ we have $E(u_t) = \ln(v/2) - \psi_0(v/2)$ and $Var(u_t) = \psi_1(v/2)$, where $\psi_0(.)$ and $\psi_1(.)$ denote the digamma and trigamma function, respectively.
In the LN-MSF-MGARCH and IG-MSF-MGARCH cases, the conditional distribution of r_t (given its past and g_t) is Normal with mean μ_t and covariance matrix $\Sigma_t = g_t H_t$. In the IG-MSF-MGARCH model class we are not able to derive the marginal distribution of g_t, which obviously depends on φ. However, for $\varphi = 0$ (the value excluded in the definition of the hybrid models under consideration) $g_t = \gamma_t$, so the distribution of g_t is known by assumption. In this case ε_t in (2) is, given its past, an IG mixture of n-variate $N(0, g_t H_t)$ distributions – i.e., it has the n-variate Student t distribution with v degrees of freedom, zero non-centrality vector and precision matrix H_t. Thus, $\varphi = 0$ corresponds to the t-MGARCH model, and the IG-MSF-MGARCH structure as a natural hybrid extension of the popular MGARCH specification with the conditional Student t distribution. We focus on a particular, simple form of the MGARCH covariance matrix H_t, namely on the SBEKK form.

3 Bayesian IG-MSF-SBEKK model and MCMC simulation of its posterior

Assume that ε_t in (1) is conditionally Normal (given parameters and latent variables), with mean vector 0 and covariance matrix $g_t H_t$. The SBEKK form of H_t is:

$$H_t = (1-\beta_1-\beta_2)A + \beta_1\left(\varepsilon'_{t-1}\varepsilon_{t-1}\right) + \beta_2 H_{t-1}, \tag{4}$$

where β_1 and β_2 are positive scalar parameters such that $\beta_1 + \beta_2 < 1$, and A is a free symmetric positive definite matrix of order n. The univariate latent process $\{g_t\}$ fulfils (3) with $\{\gamma_t\} \sim iiIG(v/2,\ v/2)$, where $v > 2$.

In order to efficiently estimate the IG-MSF-SBEKK model, which is based on as many latent variables as the number of observations, we use the Bayesian approach equipped with MCMC simulation techniques. The Bayesian statistical model amounts to specifying the joint distribution of all observations, latent variables and "classical" parameters. The assumptions presented so far determine the conditional distribution of the observations and latent variables given the parameters. Thus, it remains to formulate the marginal distribution of the parameters (the prior or *a priori* distribution). We assume independence among groups of parameters and use the same prior distributions as Osiewalski and Pajor (2009) for the same parameters. The $n(n+1)$ elements of $\delta = (\delta_0(\text{vec}\Delta)')'$ are assumed *a priori* independent of other parameters, with the $N(0, I_{n(n+1)})$ prior. Matrix A has an inverted Wishart prior distribution such that A^{-1} has the Wishart prior distribution with mean I_n; the elements of $\beta = (\beta_1, \beta_2)'$ are jointly uniformly distributed over the unit simplex. As regards initial conditions for H_t, we take $H_0 = h_0 I_n$ and treat $h_0 > 0$ as an additional parameter, *a priori* exponentially distributed with mean 1; φ has the uniform distribution over (-1, 1), and for v we assume the gamma distribution with mean λ_a/λ_v, truncated to $(2, +\infty)$. The hyper-parameters λ_a and λ_v serve to check sensitivity of posterior results to prior assumptions on v, the crucial parameter of the latent process in the IG-MSF-MGARCH model.

We can write the full Bayesian model as

$$\begin{aligned} &p\left(r_1,...,r_T,g_1,...,g_T,\delta,A,\beta,\varphi,v,h_0\right) = \\ &= p(\delta)p(A)p(h_0)p(\beta)p(\varphi)p(v)\prod_{t=1}^{T} f_N^n\left(r_t|\mu_t, g_t H_t\right) \times \\ &\times \prod_{t=1}^{T} \frac{\left(\frac{v}{2}g_{t-1}^{\varphi}\right)^{v/2}}{\Gamma\left(\frac{v}{2}\right)}\left(\frac{1}{g_t}\right)^{v/2+1} e^{-(v/2)g_{t-1}^{\varphi}/g_t}. \end{aligned} \tag{5}$$

Throughout the paper we use $f_N^n(\cdot|a,B)$ to denote the density function of the n-variate normal distribution with mean vector a and covariance matrix B, and $f_{IG}(\cdot|c,d)$ to denote the density function of the $IG(c,d)$ distribution. The posterior density function, proportional to (5), is highly dimensional and non-standard. Thus Bayesian analysis is performed on the basis of a MCMC sample from the posterior distribution, which is obtained using Gibbs algorithm, i.e. the sequential sampling from the conditional distributions obtained from (5):

$$p\left(\delta|r_1,\ldots,r_T,g_1,\ldots,g_T,A,\beta,\varphi,v,h_0\right)\propto p(\delta)\prod_{t=1}^{T}f_N^n\left(r_t|\mu_t,g_tH_t\right),$$

$$p\left(A|r_1,\ldots,r_T,g_1,\ldots,g_T,\delta,\beta,\varphi,v,h_0\right)\propto p(A)\prod_{t=1}^{T}f_N^n\left(r_t|\mu_t,g_tH_t\right),$$

$$p\left(\beta_1,\beta_2,h_0|r_1,\ldots,r_T,g_1,\ldots,g_T,\delta,A,\varphi,v\right)\propto p(\beta)p(h_0)\prod_{t=1}^{T}f_N^n\left(r_t|\mu_t,g_tH_t\right),$$

$$p\left(\varphi|r_1,\ldots,r_T,g_1,\ldots,g_T,\delta,A,\beta,v,h_0\right)\propto$$
$$\propto e^{\varphi v/2\sum_{t=1}^{T}\ln g_{t-1}}\times e^{-v/2\sum_{t=1}^{T}g_{t-1}^{\varphi}/g_t}I_{(-1,1)}(\varphi),$$

$$p\left(v|r_1,\ldots,r_T,g_1,\ldots,g_T,\delta,A,\beta,\varphi,h_0\right)\propto\left(\frac{v}{2}\right)^{Tv/2+\lambda_a-1}\Gamma\left(\frac{v}{2}\right)^{-T}e^{-\kappa v},$$

where $\kappa=-\frac{1}{2}\sum_{t=1}^{T}\ln\frac{g_{t-1}^{\varphi}}{g_t}+\frac{1}{2}\sum_{t=1}^{T}\frac{g_{t-1}^{\varphi}}{g_t}+\lambda_v$,

$$p(g_t|r_1,\ldots,r_T,g_1,\ldots,g_{t-1},g_{t+1},\ldots,g_T,\delta,A,\beta,\varphi,v,h_0)\propto$$
$$\propto f_{IG}\left(g_t|\frac{n}{2}+\frac{v}{2}(1-\varphi),\frac{1}{2}(r_t-\mu_t)H_t^{-1}(r_t-\mu_t)'+\frac{v}{2}g_{t-1}^{\varphi}\right)e^{-(v/2)g_t^{\varphi}/g_{t+1}};$$
$$t=1,\ldots,T-1;$$

$$p(g_T|r_1,\ldots,r_T,g_1,\ldots,g_{T-1},\delta,A,\beta,\varphi,v,h_0)\propto$$
$$\propto f_{IG}\left(g_T|\frac{n}{2}+\frac{v}{2}(1-\varphi),\frac{1}{2}(r_T-\mu_T)H_T^{-1}(r_T-\mu_T)'+\frac{v}{2}g_{T-1}^{\varphi}\right).$$

Excluding the full conditional posterior of v, drawing from each conditional distribution above is done through Metropolis-Hastings steps. The full conditional posterior of v is not standard, but the acceptance-rejection method proposed by Geweke (1992) is applied. As regards initial conditions for $\{\ln g_t\}$, we assume $\ln g_0 = 0$.

4 Alternative prior distributions for the latent processes

In Table 1 we show the properties of the prior distributions of the parameters σ^2, v and the marginal variances of latent processes in the LN-MSF-MGARCH and IG-MSF-MGARCH models. Let $\{\ln g_t\}$ be the process given by $\ln g_t = \varphi \ln g_{t-1} + u_t$, where

$\{u_t\}$ is a sequence of independent random variables with finite mean and variance. Since $\ln g_t$ can be expressed as $\ln g_t = \varphi^t \ln g_0 + \sum_{j=0}^{t-1} \varphi^j u_{t-j}$, then (for $\ln g_0$ constant) $E(\ln g_t) = \varphi^t E(\ln g_0) + \sum_{j=0}^{t-1} \varphi^j E(u_{t-j})$ and $Var(\ln g_t) = \sum_{j=0}^{t-1} \varphi^{2j} Var(u_{t-j})$. Note that now in our notation we omit obvious conditioning on the parameters of the process. If $u_t = \ln \gamma_t$ and $\{\gamma_t\} \sim iiIG(v/2, v/2)$, we obtain for $|\varphi| < 1$:

$$E(\ln g_t) = \varphi^t E(\ln g_0) + \frac{1-\varphi^t}{1-\varphi}\left[\ln\frac{v}{2} - \psi_0\left(\frac{v}{2}\right)\right] \overset{t\to+\infty}{\to} \frac{\ln\frac{v}{2} - \psi_0\left(\frac{v}{2}\right)}{1-\varphi},$$

$$Var(\ln g_t) = \sum_{j=0}^{t-1} \varphi^{2j} Var(u_{t-j}) = \frac{1-\varphi^{2t}}{1-\varphi^2}\psi_1\left(\frac{v}{2}\right) \overset{t\to+\infty}{\to} \frac{\psi_1\left(\frac{v}{2}\right)}{1-\varphi^2} = V_G.$$

If $u_t = \eta_t$ and $\eta_t \sim N(0, \sigma^2)$, then $E(\ln g_t) = \varphi^t E(\ln g_0) \overset{t\to+\infty}{\to} 0$; also, for $|\varphi| < 1$, $Var(\ln g_t) = (1-\varphi^{2t})\sigma^2/(1-\varphi^2) \overset{t\to+\infty}{\to} \sigma^2/(1-\varphi^2) = V_N$.

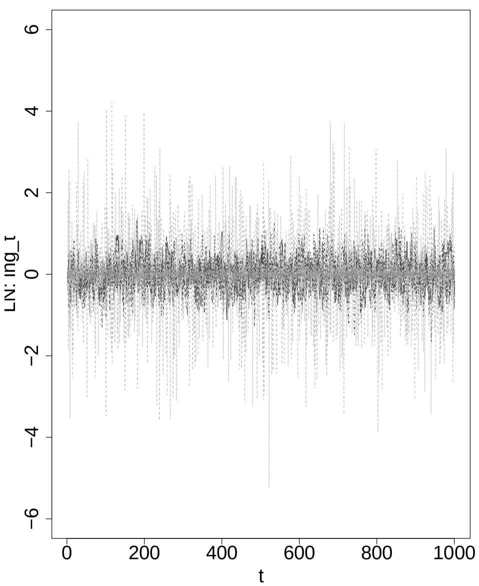
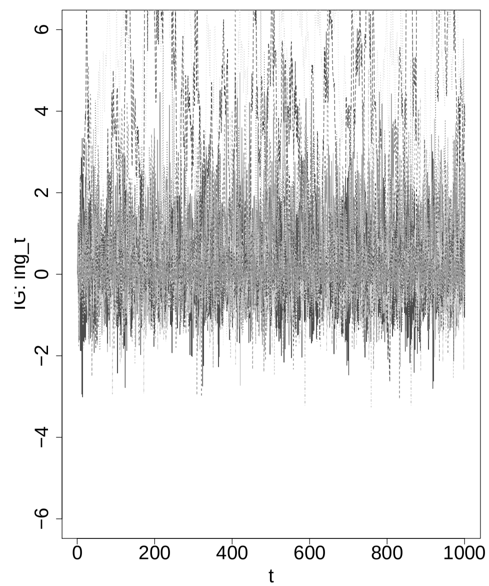
Although the parameters of the latent processes are not comparable, the variances $Var(\ln g_t)$, which are non-linear functions of these parameters, can be compared. In fact, the prior distributions of basic parameters in the LN-MSF-MGARCH and IG-MSF-MGARCH models can be treated as coherent only if they lead to similar prior distributions of $Var(\ln g_t)$. The three inverted gamma priors of σ^2 (presented in Table 1) are very different, and they lead to different priors of V_N. In terms of quantiles, variants III and II yield the most concentrated and the most diffuse distribution of V_N, respectively; the variant II prior of V_N seems also the most distant from zero, if one looks at the quantile of order 0.01. In the case of the three gamma priors of v, variants III and I lead to the most concentrated and the most diffuse distribution of V_G, respectively. Note that the corresponding quantiles of prior distributions of V_N and V_G are quite close in variant III and seem even closer in variant II, while they are completely (qualitatively) different in variant I. When we consider posterior results in two hybrid models, prior variants II and III guarantee that we compare Bayesian models based on similar assumptions about dispersion of the latent process. However, even then the tail behaviour of this process can be different. The simulated paths of the latent processes in two hybrid specifications, plotted in Figures 1–3, illustrate qualitative differences between the two models. While the latent processes based on log-normal and inverted gamma innovations have mostly similar realisations for the variant II priors, the inverted gamma innovations coupled with the variant II prior produces some paths in the far right tail. The variant III priors lead to less concentrated paths in the inverted gamma case than in the log-normal case. As expected, the variant I priors produce paths that are very volatile, but quite different between the two cases, with higher volatility and asymmetry (right skewness) in the inverted gamma case. However, note that the latent process g_t only partly describes volatility of the observed process r_t, because the parameters A and β in matrix H_t also matter and can compensate (to some extent) the tail behaviour of g_t. This can be seen in our example in the next section.

Table 1: Prior properties of σ^2, v, and unconditional variance of latent processes

	Variant I		Variant II		Variant III	
Prior distribution of σ^2 and v	$\sigma^2 \sim IG\left(1; \frac{1}{200}\right)$*	$v \sim Exp\left(\frac{1}{10}\right)$	$\sigma^2 \sim IG\left(\frac{5}{2}; \frac{4}{25}\right)$	$v \sim G\left(\frac{30}{10}; \frac{1}{10}\right)$	$\sigma^2 \sim IG\left(\frac{5}{2}; \frac{1}{40}\right)$**	$v \sim G\left(\frac{121}{10}; \frac{1}{10}\right)$
Mean, standard deviation and mode of prior distribution (non-truncated)	$E(\sigma^2)$ not existing; $Mode(\sigma^2) = \frac{1}{400}$	$E(v) = 10$ $SD(v) = 10$ $Mode(v) = 0$	$E(\sigma^2) \approx 0.1067$ $SD(\sigma^2) \approx 0.1508$ $Mode(\sigma^2) \approx$ 0.0457	$E(v) = 30$ $SD(v) = 17.32$ $Mode(v) = 20$	$E(\sigma^2) \approx 0.0167$ $SD(\sigma^2) \approx 0.0236$ $Mode(\sigma^2) \approx$ 0.0071	$E(v) = 121$ $SD(v) \approx 34.79$ $Mode(v) = 111$
Quantiles*** of prior distribution (truncated). Orders: 0.01,0.25, 0.5, 0.75, 0.99	0.001, 0.004, 0.007, 0.017, 0.502	2.101, 4.866, 8.820, 15.882, 48.263	0.021, 0.048, 0.073, 0.119, 0.579	4.542, 17.293, 26.740, 39.164, 84.245	0.003, 0.008, 0.011, 0.019, 0.090	54.88, 96.08, 117.64, 142.36, 216.03
Quantiles*** of prior distribution of V_N and V_G under assumption that $\varphi = 0$. Orders: 0.01,0.25, 0.5, 0.75, 0.99	0.001, 0.004, 0.007, 0.017, 0.502	0.043, 0.134, 0.251, 0.505, 1.533	0.021, 0.048, 0.073, 0.119, 0.579	0.024, 0.052, 0.078, 0.123, 0.551	0.003, 0.008, 0.011, 0.019, 0.090	0.009, 0.014, 0.017, 0.021, 0.037
Quantiles*** of prior distribution of V_N and V_G. Orders: 0.01,0.25, 0.5, 0.75, 0.99	0.001, 0.006, 0.013, 0.037, 1.907	0.053, 0.203, 0.429, 0.996, 19.174	0.026, 0.069, 0.121, 0.248, 5.443	0.0289, 0.0740, 0.126, 0.254, 5.377	0.004, 0.011, 0.019, 0.039, 0.859	0.011, 0.018, 0.025, 0.042, 0.927

* used in Jacquier, Polson and Rossi (2004), Osiewalski and Pajor (2009, 2018)
** used in Abanto-Valle, Lachos, and Dey (2015), Leão, Abanto-Valle and Chen (2017)
***calculated on the basis of MC simulations (10^6 values were simulated from the prior distribution; in the case of v the Gamma prior was truncated at 2)

Figure 1: Thirty realizations of $\{\ln g_t,\ t = 0, 1, \ldots, 1000\}$. Parameters were generated from the prior distribution

Prior distribution	LN-MSF-SBEKK	IG-MSF-SBEKK
variant I $\sigma^2 \sim IG\left(1; \frac{1}{200}\right)^*$ $v \sim Exp\left(\frac{1}{10}\right)$		

* used in Jacquier, Polson and Rossi (2004), Osiewalski and Pajor (2009, 2018)

Figure 2: Thirty realizations of $\{\ln g_t,\ t = 0, 1, \ldots, 1000\}$. Parameters were generated from the prior distribution

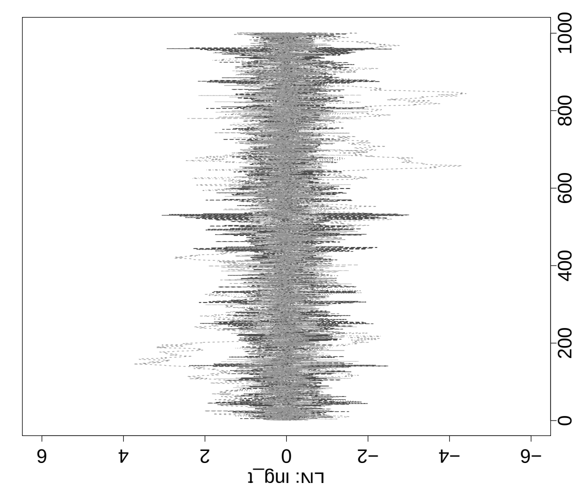
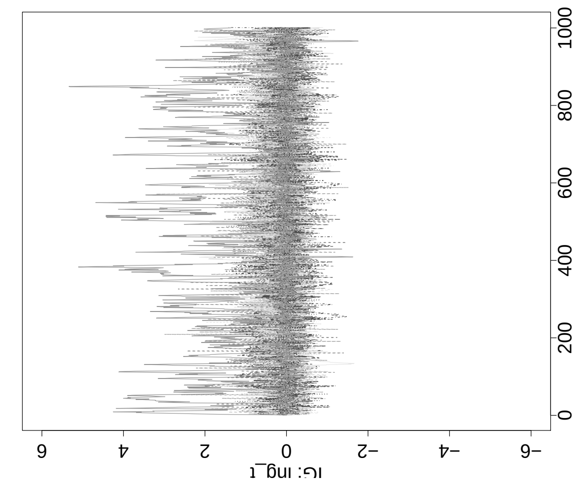
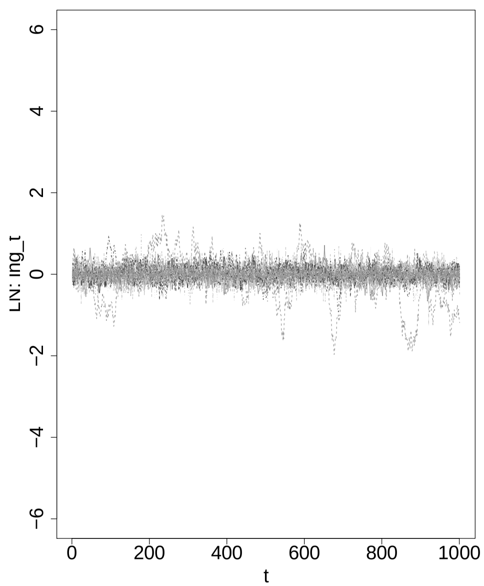
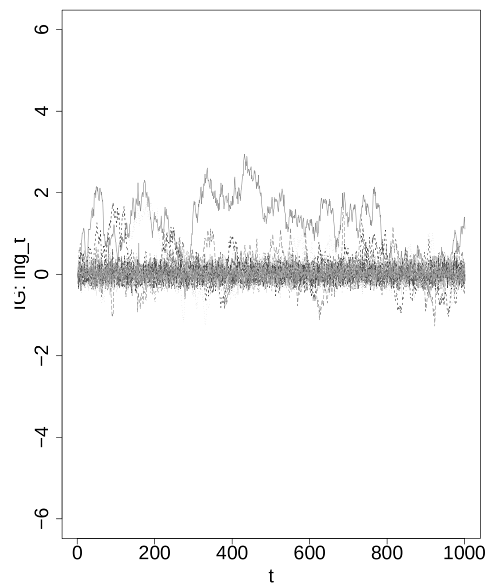
Prior distribution	LN-MSF-SBEKK	IG-MSF-SBEKK
variant II $\sigma^2 \sim IG\left(\frac{5}{2}; \frac{4}{25}\right)$ $v \sim G\left(\frac{30}{10}; \frac{1}{10}\right)$		

Figure 3: Thirty realizations of $\{\ln g_t,\ t = 0, 1, \ldots, 1000\}$. Parameters were generated from the prior distribution

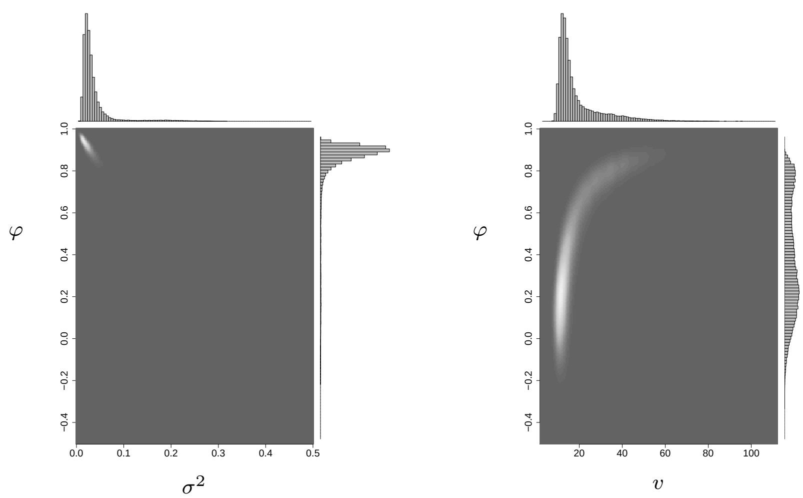
Prior distribution	LN-MSF-SBEKK	IG-MSF-SBEKK
variant III $\sigma^2 \sim IG\left(\frac{5}{2}; \frac{1}{40}\right)$** $v \sim G\left(\frac{121}{10}; \frac{1}{10}\right)$		

** used in Abanto-Valle, Lachos, and Dey (2015), Leão, Abanto-Valle and Chen (2017)

Table 2: Posterior means (and standard deviations) of the parameters of the LN-MSF-SBEKK and IG-MSF-SBEKK models (variant I)

parameter	LN-MSF–SBEKK $\sigma^2 \sim IG\left(1; \frac{1}{200}\right)$		IG-MSF-SBEKK $v \sim Exp\left(\frac{1}{10}\right)$	
δ_{01}	0.072	(0.026)	0.068	(0.026)
δ_{02}	0.027	(0.022)	0.026	(0.023)
δ_{11}	0.015	(0.024)	0.010	(0.024)
δ_{12}	0.012	(0.020)	0.010	(0.020)
δ_{21}	0.302	(0.027)	0.298	(0.026)
δ_{22}	-0.022	(0.025)	-0.024	(0.025)
a_{11}	1.136	(0.268)	0.665	(0.140)
a_{12}	0.162	(0.104)	0.091	(0.070)
a_{22}	0.736	(0.188)	0.465	(0.097)
φ	**0.880**	**(0.148)**	**0.393**	**(0.273)**
σ^2 or v	**0.033**	**(0.038)**	**18.389**	**(10.976)**
β_1	0.021	(0.006)	0.030	(0.007)
β_2	0.970	(0.007)	0.955	(0.009)
$\beta_1 + \beta_2$	0.991	(0.004)	0.986	(0.006)
h_0	2.935	(1.027)	2.261	(0.833)

Figure 4: Histograms of the marginal posterior distributions of (σ^2, φ) and (v, φ) (variant I)



5 An empirical example

We use the same bivariate data set as Osiewalski and Pajor (2009). It 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. They show moderate deviations from normality (empirical excess kurtosis is 3.08 for WIG and 1.91 for S&P500) and weak empirical correlation (0.174) between returns; the skewness coefficient is -0.12 for WIG and 0.09 for S&P500. The empirical part is focused on sensitivity of posterior results with respect to the prior specification of the latent process, so we consider the three variants of the prior distributions, which were defined in Section 4. The results are based on the last 500,000 MCMC states (out of the total 1,530,000 states), treated as a sample from the posterior distribution. We used our own computer codes written in GAUSS. In Table

Figure 5: Histograms of the marginal posterior distributions of σ^2, v, and φ (variant I)

LN-MSF-SBEKK | IG-MSF-SBEKK

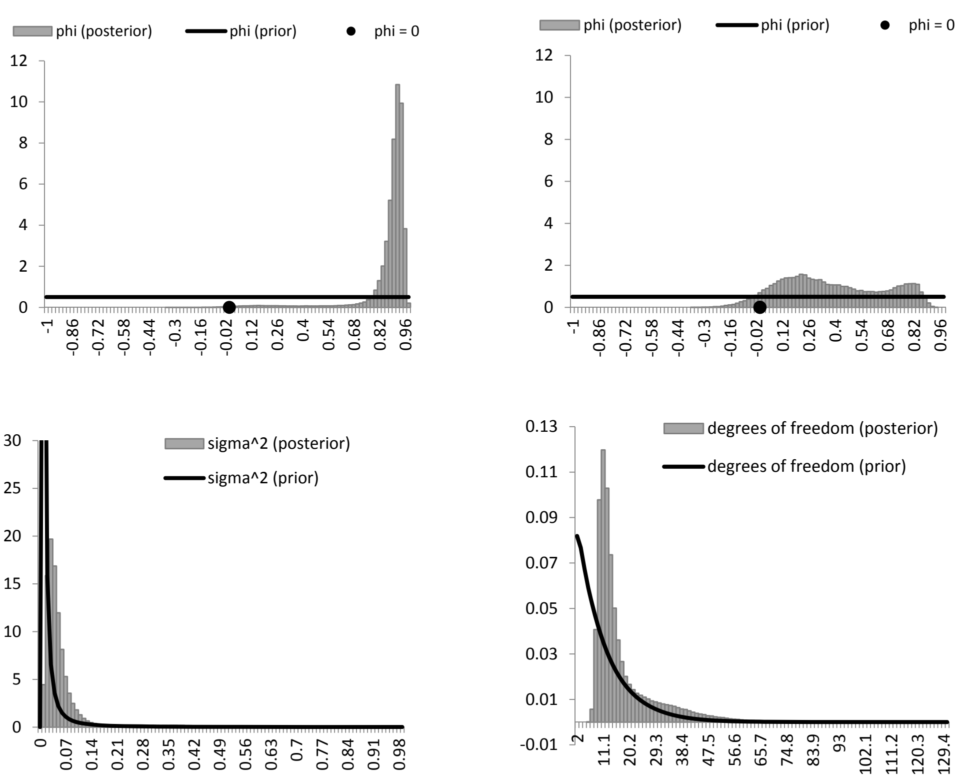


Table 3: Posterior means (and standard deviations) of the parameters of the LN-MSF-SBEKK and IG-MSF-SBEKK models (variant II)

parameter	LN-MSF–SBEKK $\sigma^2 \sim IG\left(\frac{5}{2}; \frac{4}{25}\right)$		IG-MSF-SBEKK $v \sim G\left(\frac{30}{10}; \frac{1}{10}\right)$	
δ_{01}	0.072	(0.026)	0.071	(0.026)
δ_{02}	0.028	(0.023)	0.028	(0.023)
δ_{11}	0.014	(0.024)	0.013	(0.024)
δ_{12}	0.011	(0.021)	0.010	(0.020)
δ_{21}	0.301	(0.027)	0.300	(0.027)
δ_{22}	-0.022	(0.026)	-0.022	(0.026)
a_{11}	1.080	(0.259)	0.694	(0.142)
a_{12}	0.155	(0.113)	0.094	(0.066)
a_{22}	0.704	(0.198)	0.460	(0.099)
φ	**0.774**	**(0.186)**	**0.663**	**(0.253)**
σ^2 or v	**0.065**	**(0.047)**	**34.473**	**(19.113)**
β_1	0.024	(0.006)	0.024	(0.007)
β_2	0.967	(0.008)	0.963	(0.009)
$\beta_1 + \beta_2$	0.992	(0.004)	0.987	(0.005)
h_0	2.530	(0.904)	2.271	(0.829)

Figure 6: Histograms of the marginal posterior distributions of (σ^2, φ) and (v, φ) (variant II)

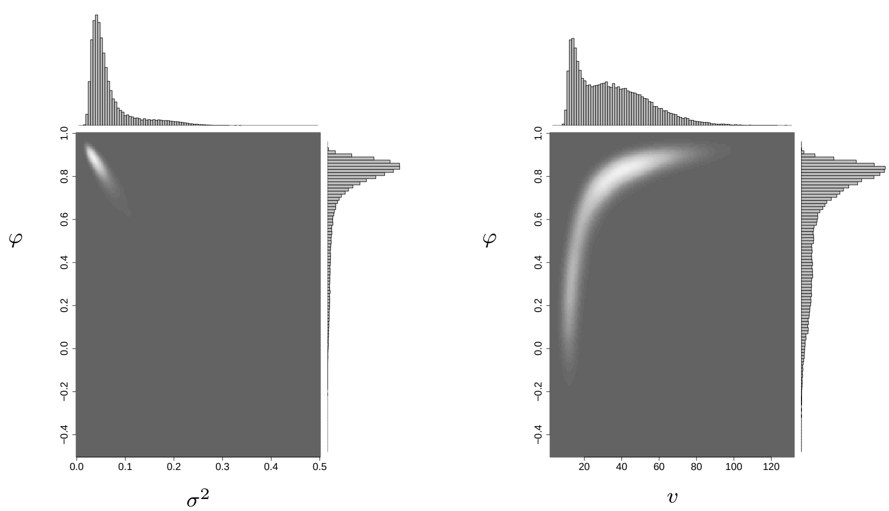
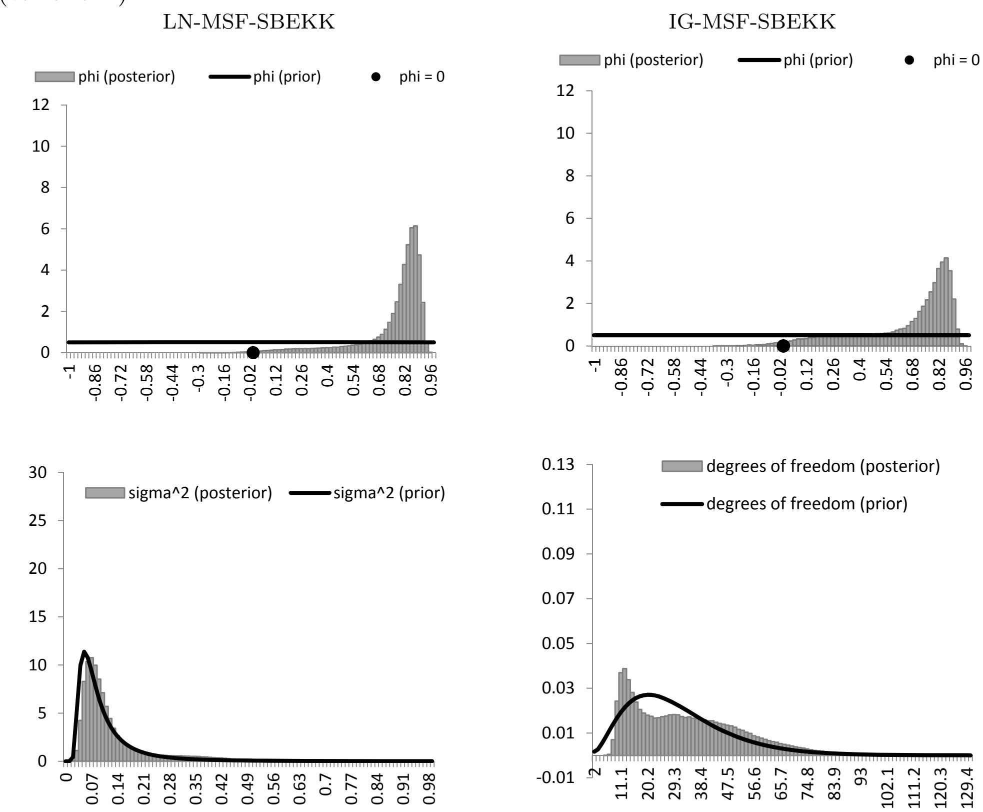


Figure 7: Histograms of the marginal posterior distributions of σ^2, v, and φ (variant II)



2 the posterior means and standard deviations of the LN-MSF-SBEKK and IG-MSF--SBEKK parameters are presented for the prior variant I. It is important to note that the posterior distribution of φ, the latent process auto-regression parameter, is much further from zero in the LN-MSF-SBEKK model. It seems that the LN-MSF-SBEKK model really needs the non-trivial Gaussian AR(1) latent process in order to describe the data, so that the case $\varphi = 0$, i.e. the SBEKK specification with log-normal scale mixture as the conditional distribution, is excluded; see also Figures 4 and 5. The question whether the IG-MSF-SBEKK model can be reduced to the t-SBEKK case is answered based on the results in Table 5. The posterior probability that $\varphi < 0$ is 0.057 and $\varphi = 0$ is included in the highest posterior density (HPD) interval of probability content 0.795. The t-SBEKK model cannot be rejected in the case of prior I, but empirical relevance of the t-SBEKK case is questionable under prior II and even more under prior III (thus it is very sensitive to the prior specification). This

can be seen when looking at the numbers in Tables 3–4 and at the plots in Figures 5, 7 and 9. Also note that the marginal posterior distributions of v and particularly σ^2 are similar to the priors, which means that the data provide little information about the parameters of the error term of the latent process $\{\ln g_t\}$.

The dependence between φ and v in the IG-MSF-SBEKK model, visible in the plots of their bivariate marginal posterior distribution (Figures 4, 6 and 8), is striking. The plots reveal non-linear relations, different for each prior variant, but always positively monotonic in the sense that higher values of one parameter correspond to higher values of the other. Thus we easily see that the posterior distribution concentrating at low values of v tends to indicate at φ closer to zero. Since the bivariate posterior of $(v,\ \varphi)$ is sensitive to the prior, inference on adequacy of the t-SBEKK model is sensitive as well. Assumptions that lead to volatile g_t (like prior II or, even more, prior I) make $\varphi = 0$ quite likely *a posteriori*. In the case of the LN-MSF-SBEKK model we see that the joint posterior of $(\sigma^2,\ \varphi)$ is also characterised by a positive monotonic relation between precision $(1/\sigma^2)$ and autocorrelation (φ). However, the latent process is never so volatile as to support the case $\varphi = 0$.

Table 4: Posterior means (and standard deviations) of the parameters of the LN-MSF-SBEKK and IG-MSF-SBEKK models (variant III)

parameter	LN-MSF–SBEKK $\sigma^2 \sim IG\left(\frac{5}{2}; \frac{1}{40}\right)$		IG-MSF-SBEKK $v \sim G\left(\frac{121}{10}; \frac{1}{10}\right)$	
δ_{01}	0.072	(0.026)	0.072	(0.026)
δ_{02}	0.028	(0.022)	0.027	(0.022)
δ_{11}	0.016	(0.024)	0.017	(0.024)
δ_{12}	0.013	(0.020)	0.013	(0.020)
δ_{21}	0.302	(0.027)	0.302	(0.027)
δ_{22}	−0.022	(0.026)	−0.022	(0.025)
a_{11}	1.170	(0.318)	0.807	(0.162)
a_{12}	0.166	(0.105)	0.116	(0.064)
a_{22}	0.743	(0.188)	0.527	(0.118)
φ	**0.920**	**(0.042)**	**0.925**	**(0.024)**
σ^2 or v	**0.023**	**(0.012)**	**108.81**	**(27.793)**
β_1	0.020	(0.004)	0.018	(0.004)
β_2	0.971	(0.006)	0.971	(0.005)
$\beta_1 + \beta_2$	0.991	(0.003)	0.989	(0.004)
h_0	2.968	(0.987)	2.923	(0.979)

Sensitivity of inferences on the parameters of the latent process does not transform into similar sensitivity of the posterior means of the process itself. In fact, the posterior means of latent variables g_t $(t = 1, 2, \ldots, T)$, characterised in Table 6,

Figure 8: Histograms of the marginal posterior distributions of (σ^2, φ) and (v, φ) (variant III)

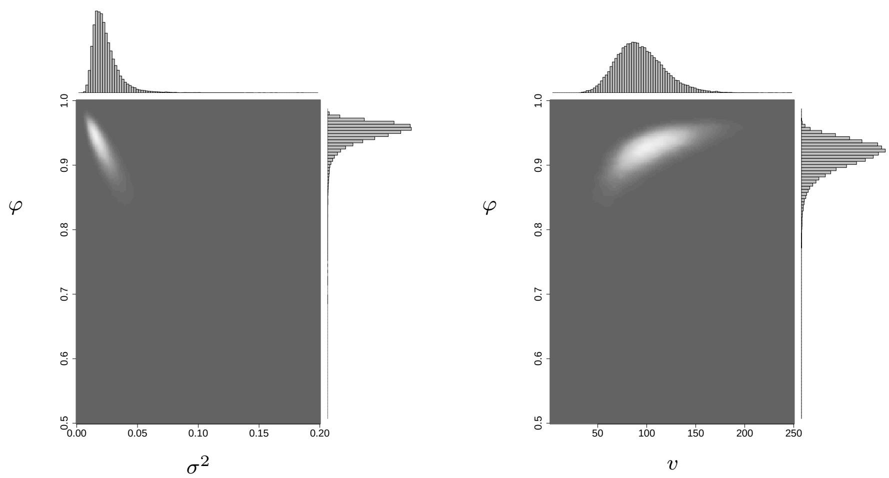


Table 5: Posterior probability that $\varphi < 0$ and probability content of minimum HPD interval including $\varphi = 0$

Variant	Prior distributions	$Pr(\varphi \leq 0 \vert r)$	min HPD
I	LN-MSF-SBEKK: $\sigma^2 \sim IG\left(1; \frac{1}{200}\right)$	0.003	0.9793
	IG-MSF-SBEKK: $v \sim Exp\left(\frac{1}{10}\right)$	0.055	0.7808
II	LN-MSF-SBEKK: $\sigma^2 \sim IG\left(\frac{5}{2}; \frac{4}{25}\right)$	0.004	0.957
	IG-MSF-SBEKK: $v \sim G\left(\frac{30}{10}; \frac{1}{10}\right)$	0.017	0.8905
III	LN-MSF-SBEKK: $\sigma^2 \sim IG\left(\frac{5}{2}; \frac{1}{40}\right)$	0.000	1.000
	IG-MSF-SBEKK: $v \sim G\left(\frac{121}{10}; \frac{1}{10}\right)$	0.000	1.000

indicate quite a bit of robustness. It is important to note the difference in average posterior means of g_t between the LN and IG cases (in the latter case they are about 15% higher). In all six Bayesian models, dispersion of individual posterior means of g_t is in the range 0.28–0.35 and they are highly correlated, which means similar dynamics; the lowest correlation characterises the pairs (IG I, IG III; 0.685), (IG I, LN III; 0.664) and (IG I, LN I; 0.702). All this is illustrated in Figure 10.

While posterior inference on the parameters of the latent process is very sensitive, Tables 2–4 show robustness of inference on the remaining parameters, which enter the SBEKK and VAR elements of our hybrid models. Only posterior results on matrix

Figure 9: Histograms of the marginal posterior distributions of σ^2, v, and φ (variant III)

LN-MSF-SBEKK IG-MSF-SBEKK

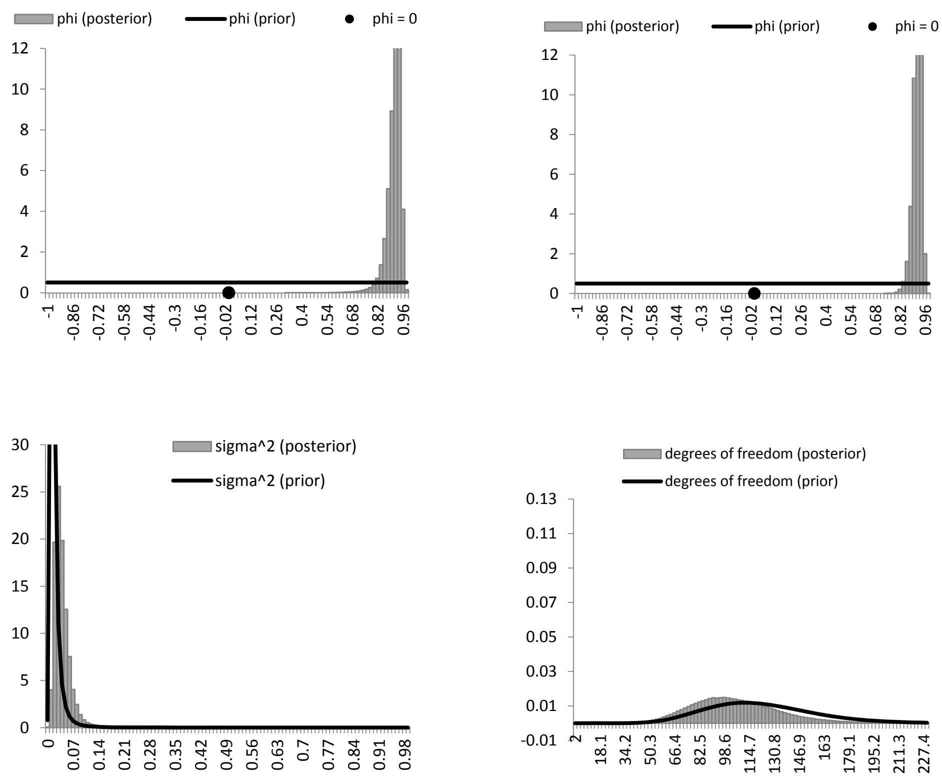


Table 6: Basic characteristics (averages, standard deviations, correlation coefficients) of the posterior means of the latent process g_t in six Bayesian MSF-SBEKK models (two types of latent process innovations: LN or IG, three variants of the prior)

model type	average	st. dev.	correlation coefficients					
			LN I	LN II	LN III	IG I	IG II	IG III
LN I	1.072	0.297	1	0.969	0.998	0.733	0.923	0.996
LN II	1.084	0.279	0.969	1	0.955	0.851	0.985	0.947
LN III	1.068	0.307	0.998	0.955	1	0.697	0.901	0.999
IG I	1.245	0.285	0.733	0.851	0.697	1	0.926	0.685
IG II	1.243	0.303	0.923	0.985	0.901	0.926	1	0.892
IG III	1.210	0.350	0.996	0.947	0.999	0.685	0.892	1

A in the SBEKK part of the conditional covariance matrix of observed returns show a particular pattern of sensitivity: the posterior means are similar for models of the same type (LN or IG), and differ between the two model classes – in the LN case they are about 40-80% higher. This effect is almost perfectly compensated by the systematic difference in average posterior means of g_t between the LN and IG cases and leads, in all Bayesian models, to the same posterior means of the conditional standard deviations of the two observed returns; see Tables 7 and 8.

Table 7: Averages, standard deviations and correlation coefficients of the posterior means of the conditional standard deviation for S&P500 in six Bayesian MSF-SBEKK models (two types of latent process innovations: LN or IG, three variants of the prior)

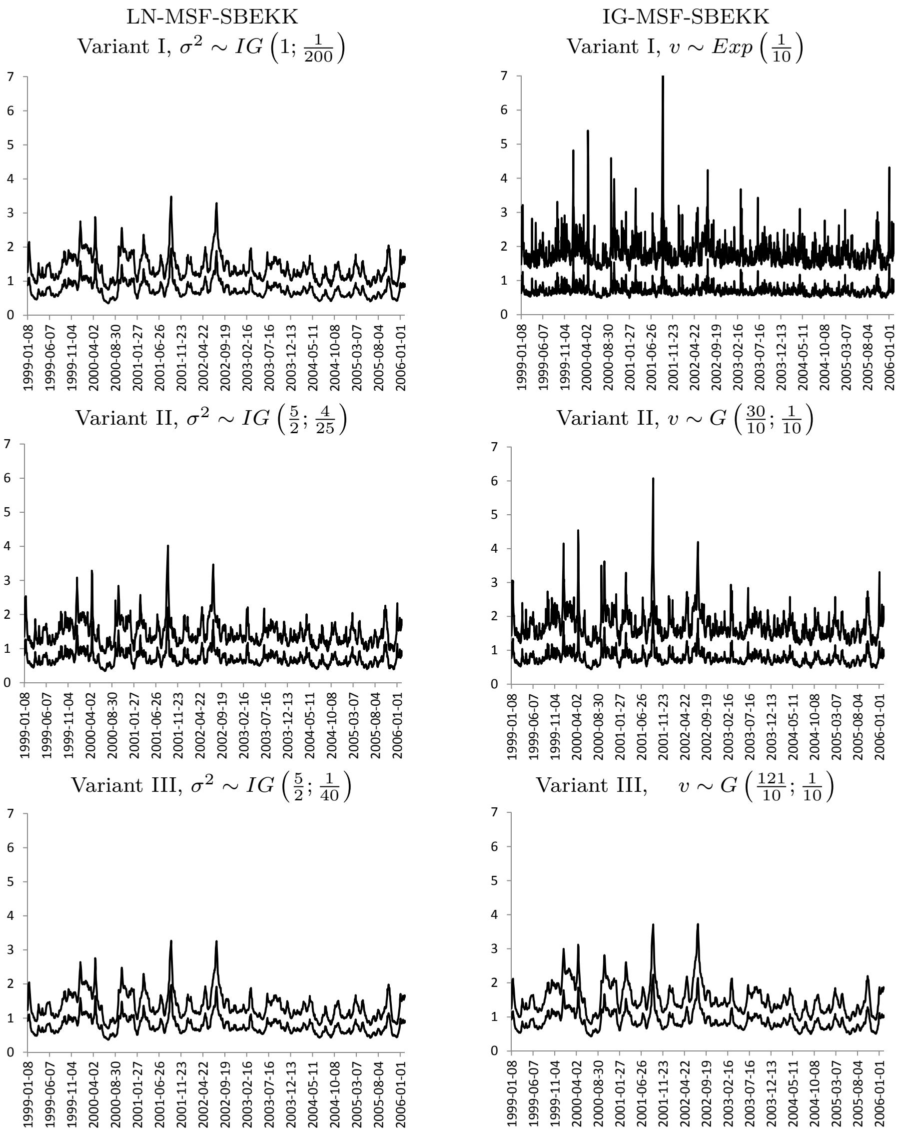
model type	average	st. dev.	correlation coefficients					
			LN I	LN II	LN III	IG I	IG II	IG III
LN I	1.112	0.384	1	0.997	1.000	0.974	0.992	1.000
LN II	1.112	0.385	0.997	1	0.995	0.986	0.998	0.994
LN III	1.112	0.386	1.000	0.995	1	0.969	0.989	1.000
IG I	1.112	0.380	0.974	0.986	0.969	1	0.994	0.968
IG II	1.111	0.381	0.992	0.998	0.989	0.994	1	0.988
IG III	1.113	0.385	1.000	0.994	1.000	0.968	0.988	1

Table 8: Averages, standard deviations and correlation coefficients of the posterior means of the conditional standard deviation for WIG in six Bayesian MSF-SBEKK models (two types of latent process innovations: LN or IG, three variants of the prior)

model type	average	st. dev.	correlation coefficients					
			LN I	LN II	LN III	IG I	IG II	IG III
LN I	1.232	0.367	1	0.995	1.000	0.968	0.989	0.999
LN II	1.230	0.368	0.995	1	0.993	0.983	0.998	0.992
LN III	1.233	0.369	1.000	0.993	1	0.962	0.985	1.000
IG I	1.223	0.365	0.968	0.983	0.962	1	0.992	0.960
IG II	1.228	0.365	0.989	0.998	0.985	0.992	1	0.984
IG III	1.233	0.367	0.999	0.992	1.000	0.960	0.984	1

Robustness of posterior inference on volatility (measured by the sampling conditional standard deviations of n elements of r_t given both g_t and the past of r_t) and on conditional correlation of the analysed stock returns is the most important empirical finding. Tables 7–9 and Figures 11–12 demonstrate that each hybrid structure under any variant of the prior distribution produces almost the same posterior evidence on these unobservable characteristics (of the observed returns), which are of particular interest in econometric analysis of financial (or commodity) markets. Even the most

Figure 10: Latent process g_t (posterior mean ± standard deviation)



different posterior results on volatility (Figure 11) and conditional correlation (Figure 12) are, in fact, very similar.

Table 9: Averages, standard deviations and correlation coefficients of the posterior means of the conditional correlation coefficient for (S&P500, WIG) in six Bayesian MSF-SBEKK models (two types of latent process innovations: LN or IG, three variants of the prior)

model type	average	st. dev.	correlation coefficients					
			LN I	LN II	LN III	IG I	IG II	IG III
LN I	0.187	0.118	1	0.998	1.000	0.979	0.993	1.000
LN II	0.188	0.126	0.998	1	0.997	0.989	0.998	0.997
LN III	0.187	0.116	1.000	0.997	1	0.975	0.991	1.000
IG I	0.187	0.142	0.979	0.989	0.975	1	0.996	0.976
IG II	0.186	0.131	0.993	0.998	0.991	0.996	1	0.991
IG III	0.187	0.115	1.000	0.997	1.000	0.976	0.991	1

Figure 11: Most different results on conditional standard deviation (posterior mean ± standard deviation) for S&P500 (upper panels) and WIG (lower panels)

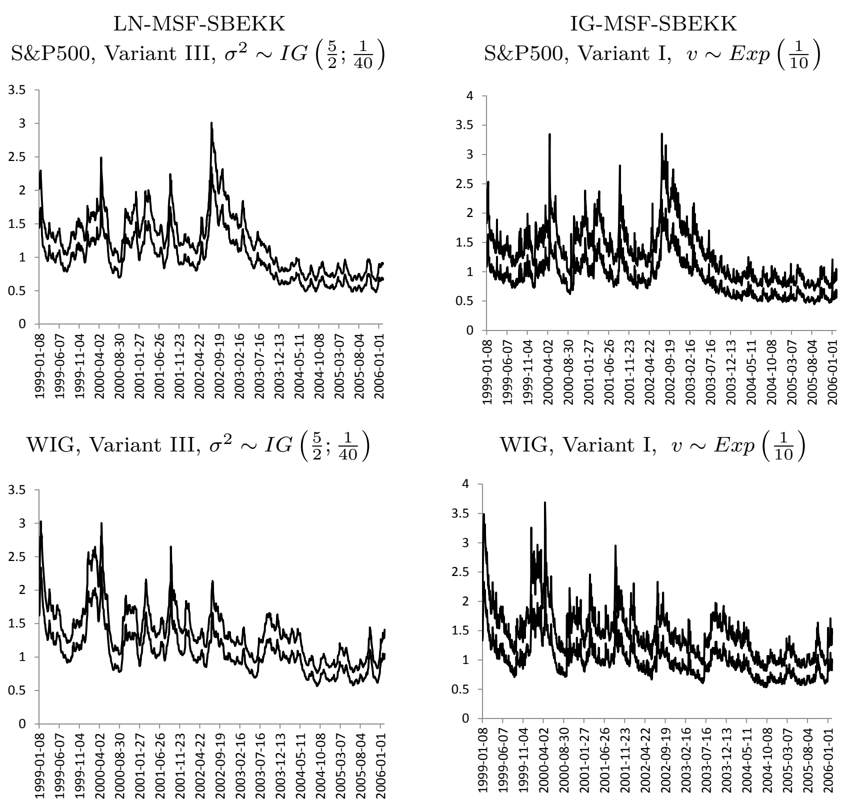
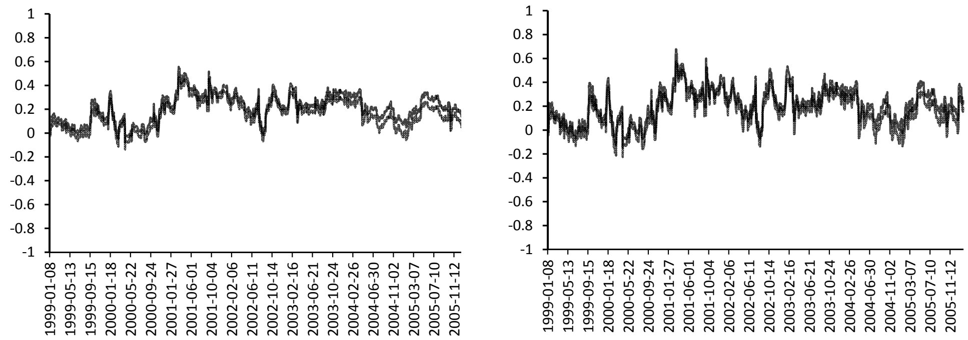


Figure 12: Conditional correlation coefficients (posterior mean ± standard deviation)

LN-MSF-SBEKK
Variant III, $\sigma^2 \sim IG\left(\frac{5}{2}; \frac{1}{40}\right)$

IG-MSF-SBEKK
Variant I, $v \sim Exp\left(\frac{1}{10}\right)$



6 Concluding remarks

We have considered two MSV-MGARCH specifications, the LN-MSF-SBEKK structure, presented by Osiewalski and Pajor (2009), and the IG-MSF-SBEKK model, proposed by Osiewalski and Pajor (2018). Due to the presence of latent variables they are estimated within the Bayesian approach, which numerically relies on MCMC simulations (Gibbs sampling with Metropolis-Hastings steps). Our main task was to study sensitivity of posterior results with respect to the form of the distribution of innovations in the latent process (inverted gamma versus log-normal) and to the prior assumptions about the parameters of the latent process. The empirical example suggests that the IG-MSF-MGARCH specification (that serves to generalise the t-MGARCH model) can relatively easily accommodate heavy tails – through latent process based on inverted gamma disturbances – in comparison to the LN-MSF-MGARCH model, based on log-normal innovations and requiring large values of the latent process auto-regression parameter φ. The posterior results (obtained in six alternative Bayesian models) for the latent process parameters are very sensitive, the posterior results for the latent process itself are much less sensitive (in fact, they are quite robust), and the results for volatilities and conditional correlation (of the analysed bivariate series of returns) are strikingly similar.

Note that we only explore differences and similarities of posterior inferences in our six Bayesian models. Formal Bayesian model comparison (through Bayes factors and posterior odds) is computationally very difficult in the hybrid framework. The crucial issue is that of precisely calculating the numerical value (for the data at hand) of the marginal density of observations $p(r_1, \ldots, r_T)$ in each model, which is the integral of the density (5) with respect to its all other arguments (i.e., latent variables and parameters). In order to approximate $p(r_1, \ldots, r_T)$ within MCMC sampling from the posterior distribution, Osiewalski and Osiewalski (2013, 2016) used the harmonic mean estimator with a specific correction. Such approach does not have so good properties as the corrected arithmetic mean estimator (CAME) proposed by Pajor (2017). However, the use of CAME in dynamic models with latent processes is not numerically feasible yet, due to very high dimensions of Monte Carlo simulation spaces. Thus, in this study we do not calculate the posterior model probabilities for the variants of the LN-MSF-SBEKK and IG-MSF-SBEKK models. Bayesian comparison of alternative specifications (including the t-SBEKK case) with the use of Bayes factors is left for future research. On the other hand, in the empirical example presented in this paper, the main posterior results on volatility and conditional correlation of the observed returns are so similar in all six Bayesian models that formal inference pooling would give almost the same outcome for any posterior distribution over the models. Thus, in this particular case we can take the results from any model, without formal comparison. Whenever both specifications considered in this paper lead to the same posterior inference on quantities of interest, we advocate to use the IG-MSF-MGARCH hybrid, which is a generalisation of the standard t-MGARCH model, so it makes testing this very popular MGARCH specification relatively easy. In our empirical example we have not examined differences and similarities of posterior inferences on risk measures such as Value-at-Risk (VaR) and Expected Shortfall (ES). For hybrid models based on the latent process with lognormal innovations, Bayesian analysis of VaR and ES was presented by Osiewalski and Pajor (2010) and Pajor and Osiewalski (2012). The flexible tail behaviour of the latent process with inverted gamma innovations makes the VaR and ES estimation (based on new hybrid models) very promising.

Acknowledgements

We are grateful to an anonymous referee for useful comments, which helped us to improve the paper. We acknowledge support from a subsidy granted to Cracow University of Economics.

References

[1] Abanto-Valle C.A., Lachos V.H., Dey D.K., (2015), Bayesian estimation of a skew-student-t stochastic volatility model, *Methodology and Computing in Applied Probability* 17(3), 721–738.

[2] Amado C., Teräsvirta T., (2013), Modelling volatility by variance decomposition, *Journal of Econometrics* 175, 142–153.

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

[4] Carriero A., Clark T., Marcellino M., (2016), Common drifting volatility in large Bayesian vars, *Journal of Business and Economic Statistics* 34, 375–390.

[5] 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.

[6] Geweke J., (1992), Priors for macroeconomic time series and their application, *Institute for Empirical Macroeconomics Discussion Paper* No. 64, Federal Reserve Bank of Minneapolis.

[7] Jacquier E., Polson N., Rossi P., (2004), Bayesian Analysis of Stochastic Volatility Models with Fat-Tails and Correlated Errors, *Journal of Econometrics* 122(1), 185–212.

[8] Leão W.L., Abanto-Valle C.A., Chen M.H., (2017), Bayesian analysis of stochastic volatility-in-mean model with leverage and asymmetrically heavy-tailed error using generalized hyperbolic skew Student's t-distribution, *Statistics and its interface* 10, 529–541.

[9] Osiewalski J., (2009), New hybrid models of multivariate volatility (a Bayesian perspective), *Przegląd Statystyczny (Statistical Review)* 56, 15–22.

[10] Osiewalski J., Osiewalski K., (2016), Hybrid MSV-MGARCH models – general remarks and the GMSF-SBEKK specification, *Central European Journal of Economic Modelling and Econometrics* 8, 241–271.

[11] 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.

[12] Osiewalski J., Pajor A., (2009), Bayesian analysis for hybrid MSF–SBEKK models of multivariate volatility, *Central European Journal of Economic Modelling and Econometrics* 1, 179–202.

[13] Osiewalski J., Pajor A., (2010), Bayesian Value-at-Risk for a portfolio: multi– and univariate approaches using MSF–SBEKK models, *Central European Journal of Economic Modelling and Econometrics* 2, 253–277.

[14] Osiewalski J., Pajor A., (2018), A hybrid MSV-MGARCH generalisation of the t-MGARCH model, The 12th Professor Aleksander Zelias International Conference

on Modelling and Forecasting of Socio-Economic Phenomena – *Conference Proceedings* 1, 345–354.

[15] Osiewalski K., Osiewalski J., (2013), A long-run relationship between daily prices on two markets: the Bayesian VAR(2)–MSF-SBEKK model, *Central European Journal of Economic Modelling and Econometrics* 5, 65–83.

[16] Pajor A., (2010), *Wielowymiarowe procesy wariancji stochastycznej w ekonometrii finansowej. Ujęcie bayesowskie*, Cracow University of Economics, Kraków.

[17] Pajor A., (2014), Konstrukcja optymalnego portfela w bayesowskim modelu MSF-SBEKK. Portfele funduszy inwestycyjnych PKO TFI, *Bank i Kredyt* (*Bank & Credit*) 43, 53–77.

[18] Pajor A., (2017), Estimating the marginal likelihood using the arithmetic mean identity, *Bayesian Analysis* 12, 261–287.

[19] Pajor A., Osiewalski J., (2012), Bayesian Value-at-Risk and Expected Shortfall for a large portfolio (multi- and univariate approaches), *Acta Physica Polonica A* 121, 2-B, B-101–B-109.

[20] Pajor A., Wróblewska J., (2017), VEC-MSF models in Bayesian analysis of short- and long-run relationships, *Studies in Nonlinear Dynamics and Econometrics* 21(3), 1–22.

[21] Teräsvirta T., (2012), Nonlinear models for autoregressive conditional heteroscedasticity, [in:] *Handbook of Volatility Models and their Applications*, [ed.:] L. Bauwens, C. Hafner, S. Laurent, Wiley, New York, 49–69.

[22] Wróblewska J., Pajor A., (2019), One-period joint forecasts of Polish inflation, unemployment and interest rate using Bayesian VEC-MSF models, *Central European Journal of Economic Modeling and Econometrics*, 11(1), 23–45.