Bayesian analysis of stochastic volatility-in-mean model with leverage and asymmetrically heavy-tailed error using generalized hyperbolic skew Student's t-distribution*

William L. Leão, Carlos A. Abanto-Valle†, and Ming-Hui Chen

A stochastic volatility-in-mean model with correlated errors using the generalized hyperbolic skew Student-t (GH-ST) distribution provides a robust alternative to the parameter estimation for daily stock returns in the absence of normality. An efficient Markov chain Monte Carlo (MCMC) sampling algorithm is developed for parameter estimation. The deviance information, the Bayesian predictive information and the log-predictive score criterion are used to assess the fit of the proposed model. The proposed method is applied to an analysis of the daily stock return data from the Standard & Poor's 500 index (S&P 500). The empirical results reveal that the stochastic volatility-in-mean model with correlated errors and GH-ST distribution leads to a significant improvement in the goodness-of-fit for the S&P 500 index returns dataset over the usual normal model.

Keywords and phrases: Feedback and leverage effect, GH skew Student-t distribution, Markov chain Monte Carlo, Non-Gaussian and nonlinear state space models, Stochastic volatility-in-mean.

1. INTRODUCTION

Stochastic volatility (SV) models were introduced in the financial literature to describe time-varying volatilities [41, 42, 21]. Although the basic SV model offers a great flexibility in modeling data with time-varying variances, it can suffer from a lack of robustness in the presence of extreme outlying observations [see, e.g., 27, 22, 1, among others] or skewness of the returns. To deal with this problem, Abanto-Valle et al. [5] and Abanto-Valle et al. [2] proposed new stochastic volatility models based on the generalized skew-Student-t and the skew-Student-t distributions for stock returns, which allow a parsimonious, flexible treatment of the

skewness and heavy tails in the conditional distribution of the returns.

However, the volatility of daily stock returns has been estimated with SV models, but the results have relied on an extensive pre-modeling of these series to avoid the problem of simultaneous estimation of the mean and variance. To remedy this problem, [26] introduced the SV in mean (SVM) by incorporating the unobserved volatility as an explanatory variable in the mean equation of the returns and provided an empirical justification that the volatility coefficient in the mean equation is related to the feedback effect, which implies that an increase in the current level of volatility causes agents to increase their forecasts of future volatility and therefore to raise their future required returns. Recently Abanto-Valle et al. [4] extended this class of models by using the scale mixture of normal distributions. It has also long been recognized in stock markets that there is a negative correlation between today's return and tomorrow's volatility. This phenomenon is called "leverage effect" or "asymmetry" [45]. The asymmetric stochastic volatility model is well known to describe these phenomena for stock returns. Markov chain Monte Carlo (MCMC) methods have been used for parameter estimation of SV models with leverage effect. For example, [30] and Omori and Watanabe [31] used an efficient mixture sampler and a block sampler for correlated errors, respectively.

In this article, we propose to enhance the robustness of the specification of the innovation returns in SVM models by introducing scale Generalized Hyperbolic skew Student-t distribution with correlated mean and variance errors. The resulting class of models takes into account the asymmetric effect, heavy-tailedness, the feedback, and leverage effects. We refer to this generalization as the SVML-GH-ST model. The flexibility of the SVML-GH-ST model can also capture time varying features in the mean of the returns and heavy tails simultaneously. The estimation of such intricate models is not straightforward, since volatility now appears in both mean and variance equations with correlated innovation errors. Hence, intensive computational methods are needed. Inference for this new SVML-GH-ST model is performed under the Bayesian paradigm via MCMC methods, which

permits obtaining the posterior distribution of parameters via simulation, starting from reasonable prior assumptions on the parameters. We simulate the log-volatilities and the shape parameters by using the block sampler for correlated errors ([31], [3], [29]) and Metropolis-Hastings algorithms, respectively.

The rest of the paper is organized as follows. Section 2 outlines the SVML-GH-ST model as well as the Bayesian estimation procedure using MCMC methods. Section 3 illustrates our proposed method using simulated data. In Section 4, the proposed class of models is applied to the S&P 500 daily returns and model comparison is provided among the competing SVML models. Finally, we conclude the paper with some concluding remarks and suggestions for future developments in Section 5.

2. THE ASYMMETRIC HEAVY-TAILED STOCHASTIC VOLATILITY-IN-MEAN MODEL WITH LEVERAGE EFFECT

2.1 The SVML-GH-ST model

The basic SV in mean model with leverage effect is defined by

$$
\text{(1a)} \quad y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 e^{h_t} + e^{\frac{h_t}{2}} \epsilon_t,
$$

$$
\text{(1b)} \quad h_{t+1} = \alpha + \phi h_t + \sigma_\eta \eta_t,
$$

$$
\text{(1c)} \quad \begin{pmatrix} \epsilon_t \\ \eta_t \end{pmatrix} \sim \mathcal{N}_2 \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right],
$$

where y_t and h_t are, respectively, the compounded return and the log-volatility at time t. We assume that $|\phi| < 1$, i.e., the log-volatility process is stationary, and the initial value $h_1 \sim \mathcal{N}(\frac{\alpha}{1-\phi}, \frac{(1-\rho^2)\sigma_\eta^2}{1-\phi^2})$. The parameter ρ measures the correlation between ϵ_t and η_t. A negative value of ρ ($\rho < 0$) indicates the so-called leverage effect, i.e., a drop in the return followed by an increase in the volatility. Empirical evidence of the leverage effect can be found in Ghysels et al. [17], Harvey and Shephard [20], Bollerslev and Zhou [8], Omori et al. [30], and Nakajima and Omori [29].

For a joint model of the asymmetric heavy-tailedness and the "feedback" and leverage effects, we replace the normal random variable ϵ_t in (1a) by a random variable from the GH skew Student's t-distribution, denoted by ω_t, which can be written in the form of the normal variance-mean mixture as

$$
\text{(2)} \quad \omega_t = \mu_\omega + \delta z_t + \sqrt{z_t} \epsilon_t,
$$

where $\epsilon_t \sim \mathcal{N}(0,1)$, $z_t \sim \mathcal{IG}\left(\frac{\nu}{2}, \frac{\nu}{2}\right)$, and $\mathcal{IG}(.,.)$ denotes the inverse gamma distribution, respectively. We assume that $\mu_\omega = -\delta \mu_z$, where $\mu_z = E(z_t) = \nu/(\nu - 2)$, to ensure $E(\omega_t) = 0$ and $\nu > 4$ for the finite variance of ω_t.

Using the variance-mean mixture representation of the GH skew Student's t-distribution defined by equation (2), the stochastic volatility-in-mean model with asymmetric

heavy-tailedness and leverage effect can be written hierarchically as

$$
\text{(3a)} \quad y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 e^{h_t} + e^{\frac{h_t}{2}} \{\delta(z_t - \mu_z) + \sqrt{z_t}\epsilon_t\},
$$

$$
\text{(3b)} \quad h_{t+1} = \alpha + \phi h_t + \sigma_\eta \eta_t,
$$

$$
\text{(3c)} \quad \begin{pmatrix} \epsilon_t \\ \eta_t \end{pmatrix} \sim \mathcal{N}_2 \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right],
$$

$$
\text{(3d)} \quad z_t \sim \mathcal{IG}\left(\frac{\nu}{2}, \frac{\nu}{2}\right).
$$

The model defined by equations (3a)–(3d) will be denoted by SVML-GH-ST. In this setup, using equations (3a), (3b), and (3c) with $\delta = 0$ and $z_t = 1 \, \forall t = 1, \ldots, T$, we obtain the SVM model with leverage effect and normal distribution (SVML-N). Equations (3a)–(3d) with $\delta = 0$ define the SVM model with leverage effect and Student-t distribution (SVML-T) [see 3, for details].

Equations (3a)–(3d) can be written in an alternative way as follows

$$
\begin{pmatrix} y_t \\ h_{t+1} \end{pmatrix} \bigg| \boldsymbol{\theta}, z_t, h_t, y_{t-1}
$$

$$
\sim \mathcal{N} \left(\begin{bmatrix} \beta_0 + \beta_1 y_{t-1} + \beta_2 e^{h_t} + e^{\frac{h_t}{2}} \delta(z_t - \mu_z) \\ \alpha + \phi h_t \end{bmatrix}, \right.
$$

$$
\text{(4)} \quad \left. \begin{bmatrix} z_t e^{h_t} & \rho \sigma_\eta z_t^{1/2} e^{h_t/2} \\ \rho \sigma_\eta \sqrt{z_t} e^{h_t/2} & \sigma_\eta^2 (1-\rho^2) \end{bmatrix} \right),
$$

From equation (4), we have that the conditional distribution $y_t | \boldsymbol{\theta}, z_t, h_t, h_{t+1}, y_{t-1}$ follows a normal distribution with mean and variance given by

$$
\mu_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 e^{h_t} + e^{\frac{h_t}{2}} \delta(z_t - \mu_z)
$$

$$
\text{(5)} \quad + \frac{\varphi}{\varphi^2 + \tau^2} \sqrt{z_t} e^{\frac{h_t}{2}} (h_{t+1} - \alpha - \phi h_t),
$$

$$
\text{(6)} \quad V_t = \frac{\tau^2}{\tau^2 + \varphi^2} z_t e^{h_t},
$$

respectively, where $\tau = \sqrt{1 - \rho^2}\sigma_\eta$ and $\varphi = \rho\sigma_\eta$. This conditional distribution will be useful in the development of the block sampler in the subsequent subsections.

2.2 Parameter estimation via MCMC

Let $\boldsymbol{\theta} = (\beta_0, \beta_1, \beta_2, \alpha, \phi, \tau^2, \varphi, \nu)'$ be the vector of parameters for the SVML-GH-ST model, where ν is the parameter of the mixing variables z_t, the degrees of freedom of the GH-ST distribution. We further let $\mathbf{h}_{1:T} = (h_1, h_1, \ldots, h_T)'$, $\mathbf{z}_{1:T} = (z_1, \ldots, z_T)'$ and $\mathbf{y}_{0:T} = (y_0, \ldots, y_T)'$ denote the the vector of the log volatilities, the mixing variables and the information available up to time T, respectively. Using the data augmentation principle, the joint posterior density of the parameters and latent unobservable variables can be written as

$$p(\boldsymbol{\theta}, \mathbf{h}_{1:T}, \mathbf{z}_{1:T} \mid \mathbf{y}_{0:T})$$

$$\propto \left[\prod_{t=1}^{T-1} p(y_t, h_{t+1} \mid z_t, h_t, y_{t-1}, \boldsymbol{\theta}) p(z_t \mid \nu) \right]$$

$$(7) \qquad \times p(y_T \mid z_T, h_T, y_{T-1}, \boldsymbol{\theta}) p(z_T \mid \nu) p(h_1 \mid \boldsymbol{\theta}) p(\boldsymbol{\theta}),$$

where $p(y_t, h_{t+1} \mid z_t, h_t, y_{t-1}, \boldsymbol{\theta})$ is given by equation (4) and $p(\boldsymbol{\theta})$ is the prior distribution. To make Bayesian analysis feasible for parameter estimation in the SVML-GH-ST model, we draw random samples from the posterior distribution of $(\boldsymbol{\theta}, \mathbf{h}_{1:T}, \mathbf{z}_{1:T})$ given $\mathbf{y}_{0:T}$ using MCMC simulation methods. The sampling scheme is described in Algorithm 2.1.

Algorithm 2.1.

1. Set $i = 0$ and get starting values for the parameters $\boldsymbol{\theta}^{(i)}$ and the latent quantities $\mathbf{z}_{1:T}^{(i)}$ and $\mathbf{h}_{1:T}^{(i)}$.
2. Generate $\boldsymbol{\theta}^{(i)}$ in turn from its full conditional distribution, given $\mathbf{y}_{1:T}$, $\mathbf{h}_{1:T}^{(i-1)}$ and $\mathbf{z}_{1:T}^{(i-1)}$.
3. Draw $\mathbf{z}_{1:T}^{(i)} \sim p(\mathbf{z}_{1:T} \mid \boldsymbol{\theta}^{(i)}, \mathbf{h}_{1:T}^{(i-1)}, \mathbf{y}_{0:T})$.
4. Generate $\mathbf{h}_{1:T}$ by blocks as follows:
   i) For $l = 1, \ldots, K$, the knot positions are generated as k_l, the floor of $[T \times \{(l + u_l)/(K+2)\}]$, where the u_l's are independent realizations of the uniform random variable on the interval $(0, 1)$.
   ii) For $l = 1, \ldots, K$, generate $h_{k_{l-1}+1:k_l-1}$ jointly conditional on $\mathbf{y}_{k_{l-1}:k_l-1}$, $\boldsymbol{\theta}^{(i)}$, $\mathbf{z}_{k_{l-1}+1:k_l-1}^{(i)}$, $h_{k_{l-1}}^{(i-1)}$, and $h_{k_l}^{(i-1)}$.
   iii) For $l = 1, \ldots, K$, draw $h_{k_l}^{(i)}$ conditional on $\mathbf{y}_{1:T}$, $\boldsymbol{\theta}^{(i)}$, $h_{k_l-1}^{(i)}$, and $h_{k_l+1}^{(i)}$.
5. Set $i = i + 1$ and return to 2 until convergence is achieved.

The prior distributions of the parameters in the SVML–GH-ST model are specified as follows: $\beta_0 \sim \mathcal{N}(\bar{\beta}_0, \sigma_{\beta_0}^2)$, $\beta_1 \sim \mathcal{N}_{(-1,1)}(\bar{\beta}_1, \sigma_{\beta_1}^2)$, $\beta_2 \sim \mathcal{N}(\bar{\beta}_2, \sigma_{\beta_2}^2)$, $\alpha \mid \tau^2 \sim \mathcal{N}(\alpha_0, \tau^2/q_0)$, $\varphi \mid \tau^2 \sim \mathcal{N}(\varphi_0, \tau^2/p_0)$, $\phi \sim \mathcal{N}_{(-1,1)}(\phi_0, s_\phi^2)$, $\tau^2 \sim \mathcal{IG}(a_\tau/2, S_\tau/2)$, and $\nu \sim \mathcal{G}(a_\nu, b_\nu)$, where a_ν, b_ν, α_0, φ_0, ϕ_0, s_ϕ^2, a_τ, S_τ, p_0, and q_0 are known hyperparameters and $\mathcal{N}_{[.,]}(.,.)$ and $\mathcal{G}(.,.)$ denote the truncated normal and gamma distributions, respectively.

As described in Algorithm 2.1, the Gibbs sampler requires sampling parameters and latent variables from their full conditional distributions. Sampling the log-volatilities $\mathbf{h}_{1:T}$ in Step 4, is the most difficult task due to the nonlinear setup of the observational equation (3a). An efficient strategy is to sample from the conditional posterior distribution of $\mathbf{h}_{1:T}$ by dividing it into several blocks and sampling each block given the other blocks. This idea, called the block sampler or multi-move sampler, was developed by Shephard and Pitt [37] and Watanabe and Omori [44] in the context of state space modeling. They showed that the sampler can produce efficient draws from the target conditional posterior distribution in comparison with a single-move sampler that primitively samples one state, say h_t, at a time given the others,

h_s $(s \neq t)$. For the SV model with leverage, Omori and Watanabe [31] developed the associated multi-move sampler and showed that it produces efficient samples. In the next subsection, we extend their method to sample $\mathbf{h}_{1:T}$ in the SVML-GH-ST model. The full conditional distributions of $\boldsymbol{\theta}$ and the latent variables $\mathbf{z}_{1:T}$ are given in Appendix A. Some of them are easy to simulate from.

2.3 A block sampler algorithm

In order to simulate $\mathbf{h}_{1:T} = (h_1, \ldots, h_T)'$ in the SVML-GH-ST model, we first simulate h_1 conditional on $\mathbf{h}_{2:T}$ and then generate $\mathbf{h}_{2:T}$ conditional on h_1. To sample the vector $\mathbf{h}_{2:T}$, we develop a multi-move block algorithm. In our block sampler, we divide it into $K+1$ blocks, $\mathbf{h}_{k_{l-1}+1:k_l-1} = (h_{k_{l-1}+1}, \ldots, h_{k_l-1})'$ for $l = 1, \ldots, K+1$, with $k_0 = 1$ and $k_{K+1} = T$, where $k_l - 1 - k_{l-1} \geq 2$ is the size of the l−th block. We sample the block of disturbances $\boldsymbol{\eta}_{k_{l-1}:k_l-2} = (\eta_{k_{l-1}}, \ldots, \eta_{k_l-2})'$ given the end conditions $h_{k_{l-1}}$ and h_{k_l} instead of $\mathbf{h}_{k_{l-1}+1:k_l-1} = (h_{k_{l-1}+1}, \ldots, h_{k_l-1})'$. In order to facilitate the exposition, we omit the dependence on $\boldsymbol{\theta}$ and assume that $k_{l-1} = t$ and $k_l = t + k + 1$ for the l−th block such that $t + k < T$. Then $\boldsymbol{\eta}_{t:t+k-1} = (\eta_t, \ldots, \eta_{t+k-1})'$ are sampled at once from their full conditional distribution $f(\boldsymbol{\eta}_{t:t+k-1} | h_t, h_{t+k+1}, \mathbf{y}_{t:t+k}, \mathbf{z}_{t+1:t+k})$[1], which without the constant terms is expressed in log scale as

$$\log f(\boldsymbol{\eta}_{t:t+k-1} | h_t, h_{t+k+1}, \mathbf{y}_{t:t+k}, \mathbf{z}_{t+1:t+k})$$

$$\doteq - \sum_{s=t}^{t+k-1} \frac{\eta_s^2}{2} + \sum_{s=t}^{t+k} l_s$$

$$(8) \qquad - \frac{1}{2\sigma_\eta^2}(h_{t+k+1} - \alpha - \phi h_{t+k})^2 \mathbb{I}(t + k < T),$$

where $\mathbb{I}(t + k < T)$ is an indicator variable. Excluding the constant terms, l_s denotes the conditional distribution of y_s given h_s and h_{s+1} for $s < T$, which is normal with mean μ_s and variance V_s, given by equations (5) and (6), respectively. We define

$$L = \sum_{s=t}^{t+k} l_s - \frac{(h_{t+k+1} - \alpha - \phi h_{t+k})^2}{2\sigma_\eta^2} \mathbb{I}(t + k < T),$$

$\mathbf{d}_{t+1:t+k} = (d_{t+1}, \ldots, d_{t+k})'$, $d_s = \frac{\partial L}{\partial h_s}$ and $\mathbf{Q} = E\left(-\frac{\partial^2 L}{\partial \mathbf{h}_{t+1:t+k} \mathbf{h}'_{t+1:t+k}}\right)$. See equations (B.1) and (B.2) in the Appendix B for details.

Since $-\frac{1}{2} \sum_{s=t}^{t+k-1} \eta_s^2 + L$ in (8) does not have the closed form, we use the Metropolis-Hastings algorithm [10] to sample from this distribution. To obtain the proposal density, we are going to form an approximated linear state space model that mimics (8), from which sampling is easy. Applying a second-order Taylor series expansion to L around the mode $\hat{\boldsymbol{\eta}}_{t:t+k-1}$, we have

---

[1] For the last block, we have $y_T \mid y_{T-1}, h_T \sim \mathcal{N}(\beta_0 + \beta_1 y_{T-1} + \beta_2 e^{h_T} + e^{h_T} \delta(z_T - \mu_z), z_T e^{h_T})$.

$\log f(\boldsymbol{\eta}_{t:t+k-1}|h_t,h_{t+k+1},\mathbf{y}_{t+1:t+k},\mathbf{z}_{t+1:t+k})$

$$\approx \text{const} - \frac{1}{2}\sum_{r=t+1}^{t+k}\eta_r^2 + \hat{L}$$

$$+ \left.\frac{\partial L}{\partial \boldsymbol{\eta}_{t:t+k-1}'}\right|_{\boldsymbol{\eta}_{t:t+k-1}=\hat{\boldsymbol{\eta}}_{t:t+k-1}} (\boldsymbol{\eta}_{t:t+k-1}-\hat{\boldsymbol{\eta}}_{t:t+k-1})$$

$$+ \frac{1}{2}(\boldsymbol{\eta}_{t:t+k-1}-\hat{\boldsymbol{\eta}}_{t:t+k-1})'$$

$$\times E\left(\frac{\partial^2 L}{\partial \boldsymbol{\eta}_{t:t+k-1}\partial \boldsymbol{\eta}_{t:t+k-1}'}\right)\Big|_{\boldsymbol{\eta}_{t:t+k-1}=\hat{\boldsymbol{\eta}}_{t:t+k-1}}$$

$$\times (\boldsymbol{\eta}_{t:t+k-1}-\hat{\boldsymbol{\eta}}_{t:t+k-1})$$

$$= \text{const} - \frac{1}{2}\sum_{r=t+1}^{t+k}\eta_r^2 + \hat{L} + \hat{\mathbf{d}}_{t+1:t+k}'(\mathbf{h}_{t+1:t+k}-\hat{\mathbf{h}}_{t+1:t+k})$$

$$- \frac{1}{2}(\mathbf{h}_{t+1:t+k}-\hat{\mathbf{h}}_{t+1:t+k})'\hat{\mathbf{Q}}(\mathbf{h}_{t+1:t+k}-\hat{\mathbf{h}}_{t+1:t+k})$$

(9)

$$= \text{const} + \log f^*(\boldsymbol{\eta}_{t:t+k-1}|h_t,h_{t+k+1},\boldsymbol{\theta},\mathbf{y}_{t+1:t+k},\mathbf{z}_{t+1:t+k}),$$

where $\hat{\mathbf{d}}_{t+1:t+k}$, $\hat{L}$, and $\hat{\mathbf{Q}}$ denote $\mathbf{d}_{t+1:t+k}$, L, and $\mathbf{Q}$ evaluated at $\mathbf{h}_{t+1:t+k} = \hat{\mathbf{h}}_{t+1:t+k}$. The expectations are taken with respect to y_s's conditional distribution on h_s's. We use an information matrix for $\mathbf{Q}$ because we require that $\mathbf{Q}$ is everywhere strictly positive definite. It can be shown that the proposal density $f^*(\boldsymbol{\eta}_{t:t+k-1}|h_t,h_{t+k+1},\boldsymbol{\theta},\mathbf{y}_{t+1:t+k},\mathbf{z}_{t+1:t+k})$ is the posterior density of $\boldsymbol{\eta}_{t:t+k-1}$ for a linear Gaussian state space model given by equations (10) and (11) below [see 31, 3, for details]. The mode $\hat{\boldsymbol{\eta}}_{t:t+k-1}$ can be found by repeating the following algorithm until convergence.

Algorithm 2.2. 1. Initialize $\hat{\boldsymbol{\eta}}_{t:t+k-1}$ and calculate $\hat{\mathbf{h}}_{t+1:t+k}$ using (3b).

2. Evaluate $\hat{d}_s$, $\hat{M}_s$, and $\hat{N}_s$ using equations (B.1), (B.3) and (B.4) respectively.

3. Compute G_s, J_s, and b_s, for $s = t+2,\ldots,t+k$, recursively, as follows:

$$G_s = \hat{M}_s - \hat{N}_s^2 G_{s-1}^{-1}, \quad G_{t+1} = \hat{M}_{t+1},$$

$$J_s = K_{s-1}^{-1}\hat{N}_s, \quad J_{t+1} = 0, \quad J_{t+k+1} = 0,$$

$$b_s = \hat{d}_s - J_s K_{t-1}^{-1}b_{s-1} \quad b_{t+1} = \hat{d}_{t+1},$$

where $K_s = \sqrt{G_s}$.

4. Define the auxiliary variables $\hat{y}_s = \hat{\gamma}_s + G_s^{-1}b_s$, where

$$\hat{\gamma}_s = \hat{h}_s + K_s^{-1}J_{s+1}\hat{h}_{s+1}, \quad s = t+1,\ldots,t+k.$$

5. Consider the linear Gaussian state-space model

(10) $\hat{y}_s = c_s + Z_s h_s + H_s \boldsymbol{\xi}_s, \quad s = t+1,\ldots,t+k,$

(11) $h_{s+1} = \alpha + \phi h_s + L_s \boldsymbol{\xi}_s, \quad s = t,t+1,\ldots,t+k,$

where $\boldsymbol{\xi}_s \sim \mathcal{N}(\mathbf{0},\mathbf{I}_2)$, $c_s = K_s^{-1}J_{s+1}\alpha$, $Z_s = 1 + K_s^{-1}J_{s+1}\phi$, $H_s = K_s^{-1}[1, J_{s+1}\sigma_\eta]$, and $L_s = [0,\sigma_\eta]$. Ap-

ply the Kalman filter and a disturbance smoother [25] to the linear Gaussian state space model in equations (10) and (11) and obtain the posterior mean of $\boldsymbol{\eta}_{t:t+k-1}$ ($\mathbf{h}_{t+1:t+k}$) and set $\hat{\boldsymbol{\eta}}_{t:t+k-1}$ ($\hat{\mathbf{h}}_{t+1:t+k}$) to this value.

6. Return to Step 2 and repeat the procedure until achieving convergence.

Applying the de Jong and Shephard simulation smoother [11] to the model defined by equations (10) and (11) with the auxiliary variables $\hat{\mathbf{y}}_{t+1:t+k}$ defined in step 4 of Algorithm 2.2 enables us to sample $\boldsymbol{\eta}_{t+1:t+k}$ from the density f^*. Since f is not bounded by f^*, we use the Metropolis-Hastings algorithm to sample from f as recommended by Chib [10].

In the MCMC sampling procedure, we select the expansion block $\hat{\mathbf{h}}_{t+1:t+k}$ in Algorithm 2.2 as follows: the current sample of $\boldsymbol{\eta}_{t:t+k=1}$ ($\mathbf{h}_{t+1:t+k}$) may be taken as an initial value of the $\hat{\boldsymbol{\eta}}_{t:t+k=1}$ ($\hat{\mathbf{h}}_{t+1:t+k}$) in Step 1. Once an initial expansion block $\hat{\mathbf{h}}_{t+1:t+k}$ is selected, we can calculate the auxiliary $\hat{\mathbf{y}}_{t+1:t+k}$ variables in Step 4. Then, applying the Kalman filter and a disturbance smoother to the linear Gaussian state space model consisting of equations (10) and (11) with the artificial $\hat{\mathbf{y}}_{t+1:t+k}$ yields the mean of $\mathbf{h}_{t+1:t+k}$ conditional on $\hat{\mathbf{h}}_{t+1:t+k}$ in the linear Gaussian state space model, which is used as the next $\hat{\mathbf{h}}_{t+1:t+k}$. By repeating the procedure until the smoothed estimates converge, we obtain the posterior mode of $\mathbf{h}_{t+1:t+k}$. This is equivalent to the method of scoring to maximize the logarithm of the conditional posterior density. Although we have just noted that iterating the procedure achieves the mode, this will slow our simulation algorithm if we have to iterate this procedure until full convergence. Instead we suggest using only five iterations of this procedure to provide a reasonably good sequence $\hat{\mathbf{h}}_{t+1:t+k}$ instead of an optimal one.

Finally, we describe the updating procedure of the knot conditions h_{k_l}, for $l = 2,\ldots,K$. As the conditional density $p(h_{k_l} \mid h_{k_l-1},h_{k_l+1})$ does not have a closed form, we use the Metropolis-Hastings algorithm with proposal density $\mathcal{N}(\frac{\alpha(1-\phi)+\phi(h_{k_l-1}+h_{k_l+1})}{1+\phi^2}, \frac{\sigma_\eta^2}{1+\phi^2})$. Let $h_{k_l}^p$ and $h_{k_l}^{(i-1)}$ denote the proposal value and the previous iteration value. Then, the acceptance probability is given by $\alpha_{MH} = \min\{1, \frac{Q(h_{k_l}^p)}{Q(h_{k_l}^{(i-1)})}\}$, where $Q(h_{k_l})$ is the product of the conditional densities $y_{k_l-1} \mid z_{k_l-1},y_{k_l-2},h_{k_l-1},h_{k_l} \sim \mathcal{N}(\mu_{k_l-1},V_{k_l-1})$, and $y_{k_l} \mid z_{k_l},y_{k_l-1},h_{k_l+1},h_{k_l} \sim \mathcal{N}(\mu_{k_l},V_{k_l})$ with μ_s and V_s defined by equations (5) and (6), respectively, for $s = k_l - 1$ and k_l.

3. NUMERICAL ILLUSTRATION WITH A SIMULATED DATASET

In order to assess the performance of the MCMC algorithms described in the previous section, we present results based on a simulated dataset. All the calculations were performed by running a stand alone code developed by the authors using the Scythe statistical library [32], which is available for free download at http://scythe.wustl.edu. We
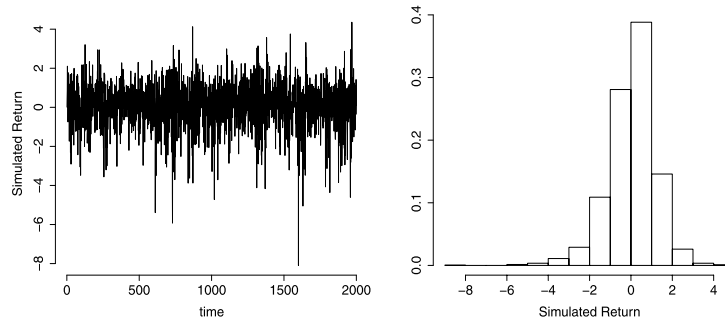
Figure 1. Simulated dataset from the SVML-GH-ST: Time series of returns (left) and the histogram (right).

Table 1. Simulated dataset: summary results for tge SVML-GH-ST model

Parmater	True value	Posterior mean	95% CI	IF	CD
β_0	0.2500	0.2810	$(0.1220, 0.4650)$	6.26	-0.12
β_1	0.0300	0.0260	$(-0.0170, 0.0680)$	1.28	0.95
β_2	-0.2000	-0.2500	$(-0.4450, -0.0700)$	5.38	0.01
α	-0.0080	-0.0160	$(-0.0340, -0.0030)$	10.36	-0.95
ϕ	0.9500	0.9210	$(0.8680, 0.9610)$	10.67	-1.01
σ_η^2	0.0225	0.0330	$(0.0160, 0.0550)$	21.83	0.96
δ	-0.5000	-0.7680	$(-1.6100, -0.3400)$	21.31	-0.32
ρ	0.3500	-0.2350	$(-0.4270, -0.0420)$	7.02	0.52
ν	10.0000	12.4430	$(8.2330, 19.5520)$	20.02	0.15

simulated a dataset of 2000 observations of the SVM-L-GH-SST distribution using $\beta_0 = 0.25$, $\beta_1 = 0.03$, $\beta_3 = -0.2$, $\alpha = -0.008$, $\phi = 0.95$, $\sigma_\eta^2 = 0.0225$, $\rho = -0.35$, and $\nu = 10$, which correspond to typical values found in daily series of returns. Figure 1 shows the raw data and the histograms of the simulated dataset.

We set the prior distributions as follows: $\beta_0 \sim \mathcal{N}(0, 100)$, $\beta_1 \sim \mathcal{N}_{(-1,1)}(0.1, 100)$, $\beta_2 \sim \mathcal{N}(-0.1, 100)$, $\alpha|\tau^2 \sim \mathcal{N}(0, \tau^2/0.002)$, $\phi|\tau^2 \sim \mathcal{N}_{(-1,1)}(0.95, 100)$, $\tau^2 \sim \mathcal{IG}(2.5, 0.025)$, $\varphi|\tau^2 \sim \mathcal{N}(-0.3, \tau^2/0.005)$, $\delta \sim \mathcal{N}(0, 1)$ and $\nu \sim \mathcal{G}(12, 0.5)$. The prior means of β_1 and ϕ are, respectively, 0.0032 and 0.0003 and the corresponding prior variances are 0.3328 and 0.3329. In both cases, the priors are equivalent to the uniform distribution on interval $(-1, 1)$, which gives zero mean and variance of 0.3333. Thus, it is clear that the priors specified for β_1 and ϕ are essentially non-informative.

The number of blocks, K, in the block sampler was set equal to 30 so that each block contained 66 $h_t's$ on average. We conducted the MCMC simulation for 50,000 iterations. The first 10,000 draws were discarded as a "burn-in" period, and then the next 40,000 were recorded. In order to reduce the autocorrelation between successive values of the simulated chain, only every $10th$ values of the chain were stored. With the resulting 4000, we calculated the posterior means, the 95% credible intervals and the convergence diagnostic (CD) statistics proposed by Geweke [16] for all the parameters.
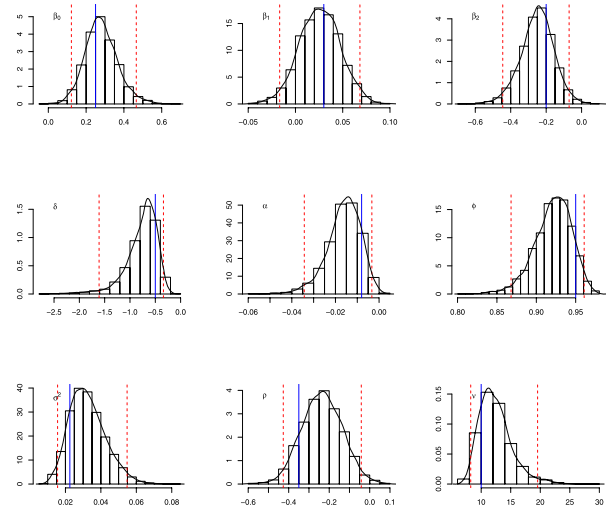

Figure 2. Simulated dataset. Histograms and estimated densities from the MCMC output for the SVML-GH-ST. The solid line indicates the true value and the dotted line the 95% credible interval.

The proposed algorithm is evaluated in terms of how well it estimates the true parameter values. From Table 1 and Figure 2, it can be seen that the estimated results for the parameters appear quite reasonable, because all the 95% credible intervals include true values. According to the CD
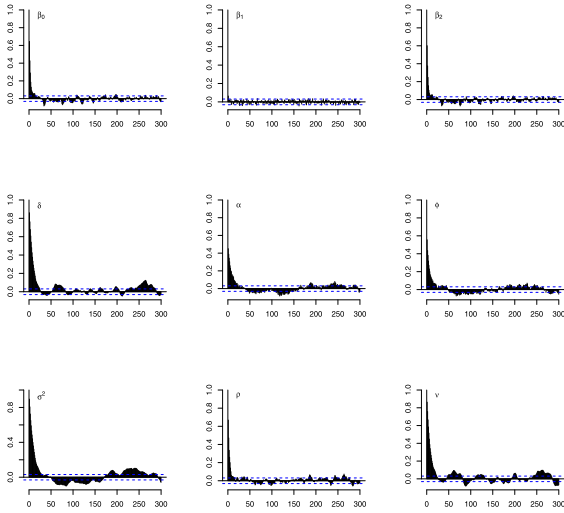
Figure 3. SVML-GH-ST, simulated dataset. Autocorrelation function (acf) for the parameters obtained from the MCMC output.
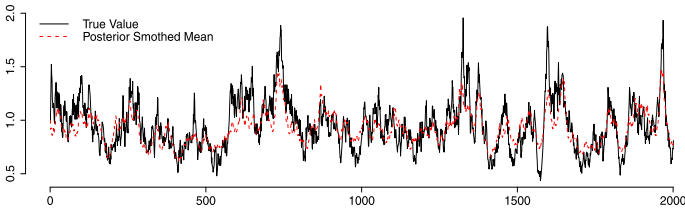

Figure 4. SVML-GH-ST, simulated dataset. True values (solid line) and posterior smoothed mean (dotted line) of $e^{\frac{h_t}{2}}$.

values, the null hypothesis that the sequence of 4000 draws is stationary was accepted at the 5% level for all the parameters in all the models considered here. The inefficiency factor is defined by $1+\sum_{s=1}^{\infty}\rho_s$, where ρ_s is the sample autocorrelation at lag s. It measures how well the MCMC chain mixes [see, e.g, 23]. It is the estimated ratio of the numerical variance of the posterior sample mean to the variance of the sample mean from uncorrelated draws. When the inefficiency factor is equal to m, we need to draw MCMC samples m times as many as the number of uncorrelated samples. From Table 1, we found that our algorithm produces a good mixing of the MCMC chain. This fact is further confirmed in Figure 3, where the the autocorrelation function (acf) of the parameters shows a faster decay.

In Figure 4, the smoothed mean calculated from the MCMC output (dotted line) and true values (solid line) of $e^{\frac{h_t}{2}}$ are shown. They show that the estimated values follow the behavior of the true volatilities.

4. EMPIRICAL APPLICATION

This section analyzes the daily closing prices for the S&P 500 stock market index. The S&P 500 index contains the

Table 2. Summary statistics for the S&P 500 returns

Median	SD	Minimum	Maximum	Skewness	Kurtosis
0.03	1.12	−22.90	10.95	−1.16	29.54

stocks of 500 Large-Cap corporations. Although a majority of those corporations are US based, it also includes other companies having their common stocks within the index. The data set was obtained from the Yahoo finance web site available to download at http://finance.yahoo.com. The period of analysis is January 4, 1980 – December 31, 2015, which yields 9078 observations. Throughout, we work with the compounded return expressed as $y_t = 100(\log P_t - \log P_{t-1})$, where P_t is the closing price on day t.

The compounded S&P 500 returns are plotted in Figure 5 as a time series and also as a histogram. The mean and standard deviation (SD) of returns are 0.03 and 1.12, respectively. As shown in Figure 5, the returns are skewed (-1.16) with heavy tails. From Table 2, we also note that the returns have a large range (minimum, -22.90 and maximum, 10.95). Some extreme observations, explained by some turbulences in financial markets as the stock market crash occurred by October 1987, the Asian financial crises in July 1997, the Russian financial crises in August 1998 and the U. S. market subprime crises in December 2007, contribute to the large kurtosis (29.54) of the S&P 500 returns. As a result, the S&P 500 index returns likely depart from the underlying normality assumption.

We fitted the SVML-N, SVML-T, and SVML-GH-ST models. In all cases, we simulated the h_t's in a multi-move fashion with stochastic knots based on the method described in Section 2.2. We set the prior distributions for the common parameters as follows: $\beta_0 \sim \mathcal{N}(0,100)$, $\beta_1 \sim \mathcal{N}_{(-1,1)}(0.1,100)$, $\beta_2 \sim \mathcal{N}(-0.1,100)$, $\phi \sim \mathcal{N}_{(-1,1)}(0.95,100)$, $\tau^2 \sim \mathcal{IG}(2.5,0.025)$, $\alpha \mid \tau^2 \sim \mathcal{N}(0,\tau^2/0.002)$, and $\varphi \mid \tau^2 \sim \mathcal{N}(-0.3,\tau^2/0.005)$. The prior distribution on the shape parameter was chosen as $\nu \sim \mathcal{G}(12,0.8)$ for the SVML-T and SVML-GH-ST models, respectively. For the SVML-GH-ST, we set $\delta \sim \mathcal{N}(0,100)$. The initial values of the parameters were randomly generated from the prior distributions. We set initial values of all the log-volatilities, h_t, to be zero. Finally the initial $\mathbf{z}_{1:T}$ were generated from the prior $p(z_t \mid \nu)$.

For the block sampler algorithm, we set the number of blocks K to be 180 in such a way that each block contained 50 h'_ts on average. For the SVML-N, SVML-T, and SVML-GH-ST models, we conducted the MCMC simulation for 25,000 iterations. In all the cases, the first 5000 draws were discarded as a burn-in period. As before, in order to reduce the autocorrelation between successive values of the simulated chain, only every $10th$ values of the chain were stored. With the resulting 2000 values, we calculated the posterior means, the 95% credible intervals and the convergence diagnostic (CD) statistics [16]. Table 3 summarizes the results.
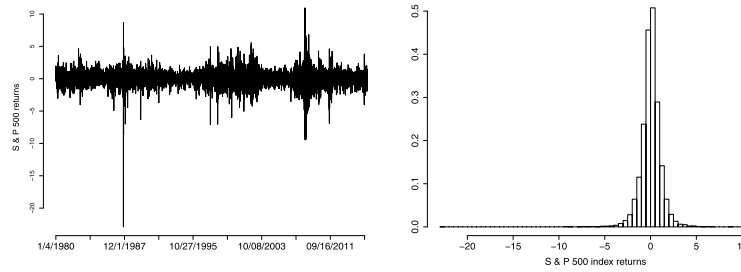
Figure 5. Compounded S & P 500 index returns from January 4, 1980 to December 31, 2015. The left panel shows the plot of the raw series and the right panel the histogram of returns.
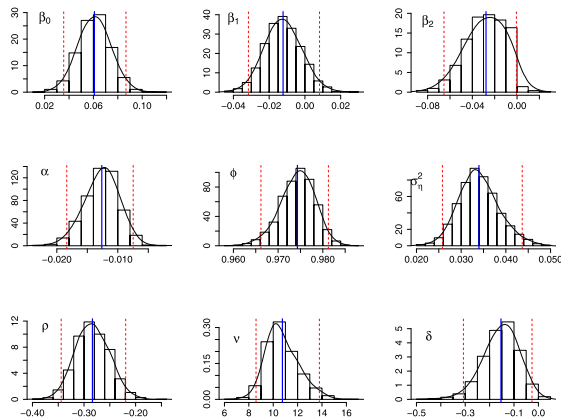

Figure 6. S&P 500 returns dataset. Histograms and estimated densities from the MCMC output for the SVML-GH-ST. The solid line indicates the posterior mean and the dotted line the 95% credible interval.
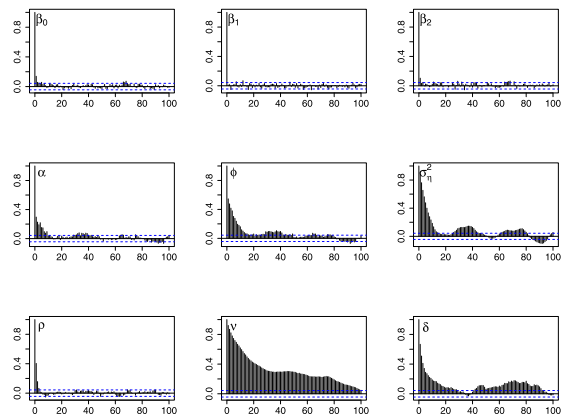

Figure 7. S&P 500 returns dataset. Autocorrelation function (acf) for the parameters obtained from the MCMC output.

According to the CD values, the null hypothesis that the sequence of 2000 draws is stationary was accepted at the 5% level for all the parameters in all the models considered here. From Table 3 and Figure 7, we found that our algorithm yields a good mixing of the MCMC chain.

Table 3 shows that the posterior mean and 95% credible interval of ϕ. For all the models, the posterior means of ϕ are above 0.97, showing higher persistence, as expected. We found that the persistence values of the SVML-T and the SVML-GH-ST are slightly different from the one for the SVML-N. The posterior means of σ_η^2 under the SVML-N and SVML-GH-ST models are greater than the posterior mean under the SVML-T model, indicating that the volatilities of the SVML-T models is less variable than the equivalent SVML-N and SVM-GH-ST models.

The posterior means together with the 95% credible intervals of the three parameters, which govern the mean process for each of the three models, are reported in Table 3. In all cases the posterior mean of β_0 is always positive and statistically significant under each fitted model. The posterior mean of β_1 is positive for the SVML-N and negative for the SVML-T and SVML-GH-ST models and similar to the first-order autocorrelation (not reported here). Since the

95% credible interval contains zero, this coefficient is not significant. The β_2 parameter, which measures both the *ex ante* relationship between returns and volatility and the volatility feedback effect, has a negative posterior mean under all of the fitted models. Although the credible interval of β_2 barely contains zero under the SVML-N and SVML-T the models, its posterior distribution is primarily located in the negative range, as shown in Table 4. The posterior mean of β_2 in the SVML-GH-ST is negative and the 95% credibility interval does not contains zero. This result confirms previous results in the literature and indicates that when investors expect higher persistent levels of volatility in the future, they require compensation for this in the form of higher expected returns.

As expected for all the models considered here, the posterior means of ρ, the correlation coefficient between shocks to return at time t and shocks to volatility at time $t + 1$, are always negative and the 95% credible intervals do not contain zero. This result indicates that this parameter is statistically significant. Hence, we may conclude that there is a strong and significant "leverage effect" for the S & P 500 index returns returns dataset.

We found that the posterior mean of δ is -0.1534, which indicates that the returns are slightly asymmetric. We also found that the 95% credible interval does not contains zero.

Table 3. Estimation results for the S & P 500 index returns. First row: Posterior mean. Second row: 95% credible interval in parentheses. Third row: CD statistics. Fourth row: Inefficiency factors

Parameter	SVML-N	SVML-T	SVML-GH-ST
β_0	0.0614	0.1801	0.0608
	(0.0378,0.0863)	(0.0274,0.0761)	(0.0356,0.0867)
	1.78	-0.70	-0.77
	1.63	1.76	1.78
β_1	0.0017	-0.0081	-0.0122
	(-0.0193,0.0231)	(-0.0271, 0.01226)	(-0.0316,0.0082)
	-0.01	1.08	-0.82
	1.00	1.00	1.00
β_2	-0.0237	-0.0176	-0.0270
	(-0.0535,0.0058)	(-0.0446,0.0017)	(-0.0654, -0.0006)
	0.24	1.78	0.49
	1.54	1.25	1.26
α	-0.0078	-0.0063	-0.0126
	(-0.0125, -0.0037)	(-0.0102, -0.0030)	(-0.0184, -0.0074)
	-1.72	1.82	-1.46
	3.33	7.51	1.26
ϕ	0.9760	0.9787	0.9743
	(0.9679,0.9829)	(0.9643,0.9826)	(0.9661,0.9812)
	-1.84	1.17	0.68
	9.78	11.26	1.26
σ_η^2	0.0332	0.0243	0.0340
	(0.0251,0.0431)	(0.0246,0.0918)	(0.0258,0.0437)
	1.51	-1.12	0.68
	17.30	26.33	13.40
ρ	-0.3622	-0.5038	-0.2837
	(-0.4385, -0.2904)	(-0.5933, -0.4145)	(-0.3443,-0.2192)
	1.78	-1.69	-0.53
	5.25	11.00	2.36
ν	–	9.1988	10.7500
	–	(7.9690,16.9087)	(8.5897,13.7853)
	–	1.32	0.16
	–	26.25	46.83
δ	–	–	-0.1534
	–	–	(-0.3443, -0.2192)
	–	–	-0.53
	–	–	9.22

Table 4. S & P 500 index returndataset: $P(\beta_2 < 0)$ estimated from the MCMC output

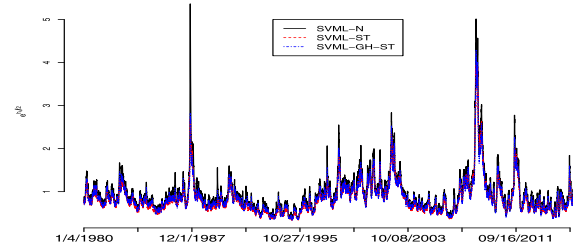|	SVML-N	SVML-T.	SVML-GH-ST
$P(\beta_2 < 0)$	0.9465	0.9710	0.9805

Figure 8. S & P 500 returns dataset. Posterior smoothed mean (dotted line) of $e^{\frac{h_t}{2}}$, SVML-GH-ST (solid line), SVML-T (dotted line), SVML-N (tiny line).

The magnitude of the tail fatness is measured by the shape parameter ν in the SVML-T and SVML-GH-ST models. The posterior means of ν are almost 9.19 and 10.75 under the SVML-T and SVML-GH-ST models, respectively. This difference can be explained by δ, the extra asymmetry parameter, which is considered in the specification of the SVML-GH-ST model. These results seem to indicate that the measurement errors of the stock returns are better explained by heavy-tailed distributions.

Now, we compare the volatility estimates. In Figure 8, we plot the smoothed mean of $e^{\frac{h_t}{2}}$. The posterior smoothed mean of $e^{\frac{h_t}{2}}$ under the SVML-T, SVML-GH-ST models show smoother movements than that under the SVML-N model (solid line). Extreme returns, such as the stock market crash occurred and the U. S. market subprime crises in December 2007, clearly make the differences. The models with heavy tails accommodate possible outliers in a somewhat different way by inflating the variance $e^{\frac{h_t}{2}}$ by $z_t^{\frac{1}{2}} e^{\frac{h_t}{2}}$. This can have a substantial impact, for instance, on the evaluation of derivative instruments and several strategic or tactical asset allocation topics.

To assess the goodness of the estimated models, we calculate the deviance information criteria, DIC [39], Bayesian predictive information criteria, BPIC [6, 7] and the log-predictive score, LPS [19, 18, 12, 5, 2]. The DIC is defined as

$$(12) \quad \text{DIC} = -2E_{\boldsymbol{\theta}|\mathbf{y}_{1:T}}[\log p(\mathbf{y}_{1:T} \mid \boldsymbol{\theta})] + p_D.$$

The second term in (12) measures the complexity of the model by the effective number of parameters, p_D, defined as the difference between the posterior mean of the deviance and the deviance evaluated at the posterior mean of the parameters:

$$(13) \quad p_D = 2[\log p(\mathbf{y}_{1:T} \mid \bar{\boldsymbol{\theta}}) - E_{\boldsymbol{\theta}|\mathbf{y}_{1:T}}[\log p(\mathbf{y}_{1:T} \mid \boldsymbol{\theta})]].$$

To calculate the DIC in the context of SVML-GH-ST model, we use the conditional likelihood $p(\mathbf{y}_{1:T} \mid \alpha, \phi, \sigma_\eta^2, \nu, \delta, \rho, \beta_0, \beta_1, \beta_2, \mathbf{z}_{1:T}, \mathbf{h}_{0:T})$, in this case $\boldsymbol{\theta}$ encompasses $(\alpha, \phi, \sigma_\eta^2, \nu, \delta, \rho, \beta_0, \beta_1, \beta_2)'$, $\mathbf{z}_{1:T}$ and $\mathbf{h}_{1:T}$.

As pointed by Stone [40], Robert and Titterington [34], Celeux et al. [9], and Ando [7], the DIC suffers from some

Table 5. S&P 500 returns dataset. Deviance Information Criteria (DIC), Bayesian predictive information criteria (BPIC) and Log Predictive Score (LPS)

Modelo	DIC	Ranking	BPIC	Ranking	LPS	Ranking
SVML-N	23579.7	3	23988.9	3	1.321	3
SVML-T	23502.3	2	23957.5	2	1.320	2
SVML-GH-ST	23452.7	1	23941.8	1	1.318	1

theoretical aspects. First, in the derivation of DIC, Spiegelhalter et al. [39] assumed that the specified parametric family of probability distributions that generate future observations encompasses the true model. This assumption may not always hold true. Secondly, the observed data are used both to construct the posterior distribution and to compute the posterior mean of the expected log-likelihood. Thus, the bias in the estimate of DIC tends to underestimate the true bias considerably. To overcome these theoretical problems in DIC, recently Ando [7] proposed the Bayesian predictive information criterion (BPIC) as an improved alternative of the DIC. The BPIC criterion is defined as

$$(14) \quad BPIC = -2E_{\boldsymbol{\theta}|\mathbf{y}_{1:T}}[\log\{p(\mathbf{y}_{1:T} \mid \boldsymbol{\theta})\}] + 2T\hat{b},$$

where $\hat{b}$ is given by

$$\hat{b} \approx \frac{1}{T}\left\{ E_{\boldsymbol{\theta}|\mathbf{y}_{1:T}}[\log\{p(\mathbf{y}_{1:T} \mid \boldsymbol{\theta})p(\boldsymbol{\theta})\}] - \log[p(\mathbf{y}_{1:T} \mid \hat{\boldsymbol{\theta}})p(\hat{\boldsymbol{\theta}})] \right.$$
$$(15)$$
$$\left. + \mathrm{tr}\{J_T^{-1}(\hat{\boldsymbol{\theta}})I_T(\hat{\boldsymbol{\theta}})\} + 0.5q \right\}.$$

Here q is the dimension of $\boldsymbol{\theta}$, $E_{\boldsymbol{\theta}|\mathbf{y}_{1:T}}[.]$ denotes the expectation with respect to the posterior distribution, $\hat{\boldsymbol{\theta}}$ is the posterior mode, and

$$I_T(\hat{\boldsymbol{\theta}}) = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{\partial \eta_T(y_t,\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\frac{\partial \eta_T(y_t,\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}\right)\bigg|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}},$$

$$J_T(\hat{\boldsymbol{\theta}}) = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{\partial^2 \eta_T(y_t,\boldsymbol{\theta})}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}'}\right)\bigg|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}},$$

with $\eta_T(y_t,\boldsymbol{\theta}) = \log p(y_t \mid \mathbf{y}_{1:t-1},\boldsymbol{\theta}) + \log p(\boldsymbol{\theta})/T$.

Scoring rules provide summary measures for the evaluation of probabilistic forecast by assigning a numerical score based on the predictive distribution and on the event or value that materializes. The fit of the models studied here will be assessed using log predictive scores [19, 18, 12, 5, 2]. The average log predictive score for the one-step ahead prediction is given by

$$(16) \quad \mathrm{LPS} = -\frac{1}{T}\sum_{t=1}^{T}\log p(y_t \mid \mathbf{y}_{1:t-1},\hat{\boldsymbol{\theta}}),$$

where $\hat{\boldsymbol{\theta}}$ is an estimate of the model parameters and $p(y_t \mid \mathbf{y}_{1:t-1},\hat{\boldsymbol{\theta}})$ is the one-step ahead predictive density. The

smaller the DIC, BPIC and LPS values, the better the model fits the data.

In the SVML class of models, the log-likelihood function, $\log p(\mathbf{y}_{1:T} \mid \boldsymbol{\theta})$ and $p(y_t \mid \mathbf{y}_{1:t-1},\boldsymbol{\theta})$ are estimated using the auxiliary particle filter [see, e.g., 33, 30] with 10,000 particles. Table 5 shows the values of BPIC. According with the DIC, BPIC and LPS criterion, the SVML-GH-ST model fits the data better among all the considered models, suggesting that the S&P 500 index returns return data demonstrate sufficient departure from underlying normality assumptions.

In order to check the distribution assumptions of the SV models, we use an approach similar to Kim, Shephard and Chib [24]. The diagnostics test is based on the probability integral transform of the realizations y_{t+1}^o taken with respect to the one-step-ahead prediction density $p(y_{t+1} \mid \mathbf{y}_{1:t},\boldsymbol{\theta})$. The probability integral transform, ε_{t+1}, is simply the cumulative distribution function corresponding to the prediction density $p(y_{t+1} \mid \mathbf{y}_{1:t},\boldsymbol{\theta})$ evaluated at y_{t+1}^o : $\varepsilon_{t+1} = Pr(y_{t+1} \leq y_{t+1}^o \mid \mathbf{y}_{1:t},\boldsymbol{\theta})$. For $t = 1,\ldots,T$, under the null hypothesis that the true distribution of y_{t+1}^o is $p(y_{t+1} \mid \mathbf{y}_{1:t},\boldsymbol{\theta})$ (or equivalently, the model is correctly specified), the ε_{t+1} converges in distribution to independent and identically distributed uniform random variables on $[0,1]$ [see, 35, 38, 24, 15, 28, among others]. By letting $\varsigma_{t+1} = \Phi^{-1}(\varepsilon_{t+1})$, where $\Phi()$ denotes the standard normal cumulative distribution function, a sequence of independent standard normal random variables ς_{t+1} is obtained, which are the standardized innovations. The probability $Pr(y_{t+1} \leq y_{t+1}^o \mid \mathbf{y}_{1:t},\boldsymbol{\theta})$ can then be approximated by

$$Pr(y_{t+1} \leq y_{t+1}^o \mid \mathbf{y}_{1:t},\boldsymbol{\theta})$$
$$= \frac{1}{N}\sum_{i=1}^{N}Pr(y_{t+1} \leq y_{t+1}^o \mid \mathbf{y}_{1:t},h_{t+1}^{(i)},\boldsymbol{\theta}).$$

The QQ-plots for pseudo residuals the three models fitted, SVML-N, SVML-T and SVML-GH-ST are shown in Figure 9. The qq-plots indicate a lack of fit in the left tail, specially in the SVML-N and SVML-ST models. The indicated mis-specification could be solved by using the SVML with generalized skew-Student-t or skew-Student-t distributions as in Abanto-Valle et al. [5] and Abanto-Valle et al. [2].

5. CONCLUSIONS

This article presented a Bayesian implementation of a robust alternative for estimation in the stochastic volatility-
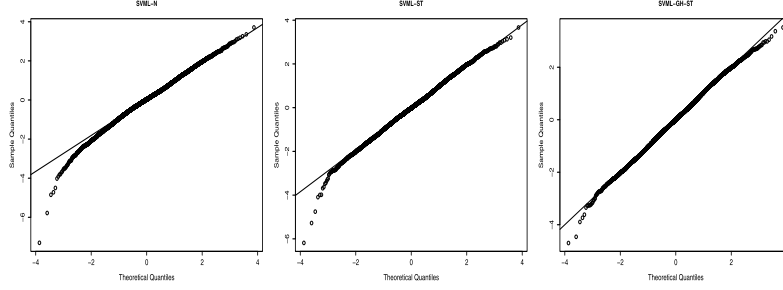
Figure 9. S & P 500 returns dataset. Quantile-Quantile plot of the residuals ς_t. The solid line plots the quantiles of the $\mathcal{N}(0,1)$ against the quantiles of the standard normal and the points were the sorted values of ς_t against the quantiles of the standard normal.

in-mean model with correlated errors, as an extension of the model proposed by [26] and Abanto-Valle et al. [3] via MCMC methods. The SVML model enables us to investigate the dynamic relationship between returns and their time-varying volatility. The Gaussian assumption of the mean innovation was replaced by univariate thick-tailed processes, known as the variance-mean mixture of the normal distribution. Under a Bayesian perspective, we developed an algorithm based on MCMC simulation methods to estimate all the parameters and latent quantities in our proposed SVML-GH-ST model. We illustrated our methods through an empirical application of the S&P 500 returns series, which shows that the SVML-GH-ST model provides a better fit than the SVML-N and SVML-T models in terms of parameter estimates, interpretation, and robustness aspects. The β_2 estimate, which measures both the *ex ante* relationship between returns and volatility and the volatility feedback effect, was found to be negative. These results are in line with those of French et al. [14], who found a similar relationship between unexpected volatility dynamics and returns, and confirm the hypothesis that investors require higher expected returns when unanticipated increases in future volatility are highly persistent. This is consistent with our findings of higher values of ϕ combined with larger negative values for the in-mean parameter. On the other hand, since the posterior mean and 95% credible interval contains only negative values, we can conclude that there is a strong and significant "leverage effect" for the S&P 500 returns dataset.

Our SVML-GH-ST models showed considerable flexibility to accommodate outliers, but their robustness aspects could be seriously affected by the prior of the ν and δ parameters. In this set-up, for example, it would be possible to study different objective priors for the parameters in the GH-ST distributions in the same spirit of the works of Fonseca et al. [13] and Salazar et al. [36] or using a different skew-student-t parameterization as in Abanto-Valle et al. [5] and [2] for example. Nevertheless, an in-depth investigation of this modification is beyond the scope of the present paper, but provides stimulating topics for future research.

APPENDIX A: THE FULL CONDITIONAL DISTRIBUTIONS

In this appendix, we describe the full conditional distributions for the parameters and the mixing latent variables $\mathbf{z}_{1:T}$ of the SVML-GH-ST model.

Full conditional distributions of β_0, β_1, and β_2

Let m_t and V_t be defined by

$$
m_t = \begin{cases} \sqrt{z_t}e^{\frac{h_t}{2}}\frac{\varphi}{\tau^2+\varphi^2}(h_{t+1}-\alpha-\phi h_t), & t < T, \\ 0, & t = T, \end{cases}
$$

$$
V_t = \begin{cases} z_t e^{h_t}\frac{\tau^2}{\tau^2+\varphi^2}, & t < T, \\ z_t e^{h_t}, & t = T. \end{cases}
$$

For parameters β_0, β_1 and β_2, we set the prior distributions as: $\beta_0 \sim \mathcal{N}(\bar{\beta}_0, \sigma^2_{\beta_0})$, $\beta_1 \sim \mathcal{N}_{(-1,1)}(\bar{\beta}_1, \sigma^2_{\beta_1})$, $\beta_2 \sim \mathcal{N}(\bar{\beta}_2, \sigma^2_{\beta_2})$. Then, the full conditionals are given by

(A.1) $\quad \beta_0 \mid \mathbf{y}_{0:T}, \mathbf{h}_{1:T}, \mathbf{z}_{1:T}, \beta_1, \beta_2 \sim \mathcal{N}(\frac{b_{\beta_0}}{a_{\beta_0}}, \frac{1}{a_{\beta_0}})$,

(A.2) $\quad \beta_1 \mid \mathbf{y}_{0:T}, \mathbf{h}_{1:T}, \mathbf{z}_{1:T}, \beta_0, \beta_1 \sim \mathcal{N}(\frac{b_{\beta_1}}{a_{\beta_1}}, \frac{1}{a_{\beta_1}})\mathbb{I}_{|\beta_2|<1}$,

(A.3) $\quad \beta_2 \mid \mathbf{y}_{0:T}, \mathbf{h}_{1:T}, \mathbf{z}_{1:T}, \beta_0, \beta_1 \sim \mathcal{N}(\frac{b_{\beta_2}}{a_{\beta_2}}, \frac{1}{a_{\beta_2}})$,

where $a_{\beta_0} = \sum_{t=1}^{T}\frac{1}{V_t} + \frac{1}{\sigma^2_{\beta_0}}$, $b_{\beta_0} = \sum_{t=1}^{T}\frac{w_t}{V_t} + \frac{\bar{\beta}_0}{\sigma^2_{\beta_0}}$, $a_{\beta_1} = \sum_{t=1}^{T}\frac{y^2_{t-1}}{V_t} + \frac{1}{\sigma^2_{\beta_1}}$, $b_{\beta_1} = \sum_{t=1}^{T}\frac{u_t y_{t-1}}{V_t} + \frac{\bar{\beta}_1}{\sigma^2_{\beta_1}}$, $a_{\beta_2} = \sum_{t=1}^{T}\frac{e^{2h_t}}{V_t} + \frac{1}{\sigma^2_{\beta_2}}$, $b_{\beta_2} = \sum_{t=1}^{T}\frac{r_t e^{h_t}}{V_t} + \frac{\bar{\beta}_2}{\sigma^2_{\beta_2}}$, $w_t = y_t - \beta_1 y_{t-1} - \beta_2 e^{h_t} - e^{\frac{h_t}{2}}\delta(z_t - \mu_z) - m_t$, $u_t = y_t - \beta_0 - \beta_2 e^{h_t} - e^{\frac{h_t}{2}}\delta(z_t - \mu_z) - m_t$, $r_t = y_t - \beta_0 - \beta_1 y_{t-1} - e^{\frac{h_t}{2}}\delta(z_t - \mu_z) - m_t$, and $\mathbb{I}_{|\beta_2|<1}$ is the indicator variable.

Full conditional distributions of α, ϕ, φ, δ, and τ^2

We assume the following prior distributions: $\alpha \mid \tau^2 \sim \mathcal{N}(\alpha_0, \tau^2/q_0)$, $\varphi \mid \tau^2 \sim \mathcal{N}(\varphi_0, \tau^2/p_0)$, $\phi \sim \mathcal{N}_{(-1,1)}(\phi_0, s_\phi^2)$, $\delta \sim \mathcal{N}(\delta_0, s_\delta^2)$, and $\tau^2 \sim \mathcal{GI}(a_\tau/2, S_\tau/2)$, where α_0, φ_0, ϕ_0, s_ϕ^2, δ_0, s_δ^2, a_τ, S_τ, p_0, and q_0 are known hyper parameters.

After some simple but tedious algebra, we obtain

$$(A.4) \qquad \alpha \mid . \sim \mathcal{N}(\frac{B_\alpha}{A_\alpha}, \frac{\tau^2}{A_\alpha}),$$

$$(A.5) \qquad \varphi \mid . \sim \mathcal{N}(\frac{B_\varphi}{A_\varphi}, \frac{\tau^2}{A_\varphi}),$$

$$(A.6) \qquad \delta \mid . \sim \mathcal{N}(\frac{B_\delta}{A_\delta}, \frac{1}{A_\delta}),$$

where $A_\alpha = q_0 + \frac{1+\phi}{1-\phi} + T - 1$, $B_\alpha = \alpha_0 q_0 + (1 + \phi)h_1 + \sum_{t=1}^{T-1} k_t$, $k_t = h_{t+1} - \phi h_t - \varphi g_t z_t^{-\frac{1}{2}} e^{-\frac{h_t}{2}}$, $A_\varphi = p_0 + \sum_{t=1}^{T-1} g_t^2 z_t^{-1} e^{-h_t}$, $B_\varphi = \varphi_0 p_0 + \sum_{t=1}^{T-1} c_t g_t z_t^{-\frac{1}{2}} e^{-\frac{h_t}{2}}$, $A_\delta = -\frac{\varphi}{\tau^2} \sum_{t=1}^{T-1} \frac{1}{\sqrt{z_t}}(z_t - \mu_z)(h_{t+1} - \alpha - \phi h_t) + \left(\frac{\varphi^2 + \tau^2}{\tau^2}\right) \sum_{t=1}^{T-1} \frac{1}{z_t e^{h_t/2}}(z_t - \mu_z)(y_t - \beta_0 - \beta_1 y_{t-1} - \beta_2 e^{h_t})$, $B_\delta = \left(\frac{\varphi^2 + \tau^2}{\tau^2}\right) \sum_{t=1}^{T-1} \left(\frac{1}{z_t}(z_t - \mu_z)^2\right) + \frac{1}{z_n}(z_n - \mu_z)^2 + \frac{1}{s_\delta^2}$, $c_t = h_{t+1} - \alpha - \phi h_t$, and $g_t = y_t - \beta_0 - \beta_1 y_{t-1} - \beta_2 e^{h_t} - e^{\frac{h_t}{2}}\delta(z_t - \mu_z)$. In a similar way, the conditional distribution of ϕ is given by

$$(A.7) \quad p(\phi \mid .) \quad \propto \quad Q(\phi) \exp\left\{ - \frac{A_\phi}{2}\left(\phi - \frac{B_\phi}{A_\phi}\right)^2 \right\},$$

where

$$Q(\phi) = \sqrt{1 - \phi^2} \exp\{-\frac{1 - \phi^2}{2\tau^2}(h_1 - \frac{\alpha}{1 - \phi})^2\},$$

$A_\phi = \frac{1}{s_\phi^2} + \sum_{t=1}^{T-1} \frac{h_t^2}{\tau^2}$, $B_\phi = \frac{\phi_0}{s_\phi^2} + \sum_{t=1}^{T-1} \frac{l_t h_t}{\tau^2}$, $l_t = h_{t+1} - \alpha - \varphi(y_t - \beta_0 - \beta_1 y_{t-1} - \beta_2 e^{h_t})z_t^{-\frac{1}{2}} e^{-\frac{h_t}{2}}$, and $\mathbb{I}_{|\phi|<1}$ is the indicator variable. As $p(\phi \mid \mathbf{h}_{1:T}, \alpha, \sigma_\eta^2)$ in (A.7) does not have a closed form, we sample from it by using the Metropolis-Hastings algorithm with truncated $\mathcal{N}_{(-1,1)}(\frac{b_\phi}{a_\phi}, \frac{\sigma_\eta^2}{a_\phi})$ as the proposal density. The conditional distribution of τ^2 is $\mathcal{IG}(\frac{T_1}{2}, \frac{M_1}{2})$, where $T_1 = a_\tau + T + 2$ and $M_1 = (1 - \phi^2)(h_1 - \frac{\alpha}{1-\phi})^2 + \sum_{t=1}^{T-1}(c_t - \varphi z_t^{-\frac{1}{2}} e^{-\frac{h_t}{2}} g_t)^2 + p_0(\varphi - \varphi_0)^2 + q_0(\alpha - \alpha_0)^2 + S_\tau$. Once τ^2 and φ are sampled, respectively, from their conditional posteriors, we can calculate ρ and σ_η^2 through $\sigma_\eta^2 = \tau^2 + \varphi^2$ and $\rho = \varphi/\sigma_\eta$.

Full conditional distributions of z_t and ν

The full conditional distribution of z_t is given by

$$p(z_t|.) \propto Q(z_t) \left(\frac{\gamma}{\vartheta}\right)^\lambda \frac{z^{\lambda-1}}{2K_\lambda(\gamma, \vartheta)} \exp\left\{ -\frac{1}{2}\left(\vartheta^2 z_t^{-1} + \gamma^2 z_t\right) \right\},$$

where the values of λ, ϑ and γ are the parameters of a distribution $\text{GIG}(\lambda, \vartheta, \gamma)$ whose values are given by

$$\lambda = -\frac{\nu + 1}{2},$$

$$\gamma^2 = \delta^2 \frac{\varphi^2 + \tau^2}{\tau^2},$$

$$\vartheta^2 = \frac{\varphi^2 + \tau^2}{\tau^2} e^{-h_t}$$
$$\times \left(y_t - \beta_0 - \beta_1 y_{t-1} - \beta_2 e^{h_t} + \mu_z e^{h_t/2}\delta\right)^2 + \nu.$$

We sample z_t by the Metropolis-Hastings algorithm. We use $\text{GIG}(\lambda, \vartheta, \gamma)$ as the proposal distribution such that z_t^* and $z_t^{(i-1)}$ are the proposal value and previous iteration value, respectively. Thus, the acceptance probability is given by $\alpha_{MH} = \min\{1, \frac{Q(z_t^*)}{Q(z_t^{(i-1)})}\}$, where

$$
\begin{aligned}
Q(z_t) = {} & \exp\Bigg\{ \frac{\varphi}{\tau^2}\Bigg[z_t^{-1/2} e^{-h_t/2}(h_{t+1} - \alpha - \phi h_t) \\
& \left(y_t - \beta_0 - \beta_1 y_{t-1} - \beta_2 e^{h_t} + \mu_z e^{h_t/2}\delta\right) - \\
& -z_t^{1/2}\delta(h_{t+1} - \alpha - \phi h_t)\Bigg]\mathbb{I}_{\{t<n\}} \Bigg\}.
\end{aligned}
$$

We assume the prior distribution of ν as $\mathcal{G}(a_\nu, b_\nu)\mathbb{I}_{4<\nu\leq40}$. Then, the full conditional distribution of ν is

$$
\begin{aligned}
& p(\nu \mid \mathbf{z}_{1:T}) \\
& \propto \frac{\frac{\nu}{2}^{\frac{T\nu}{2}}}{\Gamma\left(\frac{\nu}{2}\right)^T} \exp\Bigg\{ -\sum_{t=1}^T \frac{1}{2V_t}\Bigg[y_t - \beta_0 - \beta_1 y_{t-1} - \\
& \beta_2 e^{h_t} - e^{\frac{h_t}{2}}\delta\left(z_t - \frac{\nu}{\nu-2}\right) - m_t\Bigg]^2 \\
& -\frac{\nu}{2}\sum_{t=1}^T\left[\frac{1}{z_t} + \log z_t\right]\Bigg\}\nu^{a_\nu-1}\exp\{-b_\nu\nu\}\mathbb{I}_{4<\nu\leq40}.
\end{aligned}
$$

We sample ν by the Metropolis-Hastings algorithm [43, 10]. Let ν^* denote the mode (or approximate mode) of $p(\nu \mid \mathbf{z}_{1:T})$, and let $\ell(\nu) = \log p(\nu \mid \mathbf{z}_{1:T})$. We use the proposal density $\mathcal{N}_{(4,40)}(\mu_\nu, \sigma_\nu^2)$, where $\mu_\nu = \nu^* - \ell'(\nu^*)/\ell''(\nu^*)$ and $\sigma_\nu^2 = -1/\ell''(\nu^*)$. $\ell'(\nu^*)$ and $\ell''(\nu^*)$ are the first and second derivatives of $\ell(\nu)$ evaluated at $\nu = \nu^*$.

APPENDIX B: SOME DERIVATIONS OF THE BLOCK SAMPLER

First, we define

$$
\begin{aligned}
d_s = {} & \frac{\partial L}{\partial h_s} = -\frac{1}{2} + \frac{(y_s - \mu_s)^2}{2V_s} + \frac{(y_s - \mu_s)}{V_s}\frac{\partial \mu_s}{\partial h_s} \\
& + \frac{(y_{s-1} - \mu_{s-1})}{V_{s-1}}\frac{\partial \mu_{s-1}}{\partial h_s}
\end{aligned}
$$

$$\text{(B.1)} \qquad -\phi \frac{(h_{s+1} - \alpha - \phi h_s)}{\sigma_\eta^2} \mathbb{I}(t + k < T)$$

for $\quad s = t + 1, \ldots, t + k$, and

$$\text{(B.2)} \qquad \mathbf{Q} = \begin{pmatrix} M_{t+1} & N_{t+2} & 0 & \ldots & 0 \\ N_{t+2} & M_{t+2} & N_{t+3} & \ldots & 0 \\ 0 & N_{t+3} & M_{t+3} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & N_{t+k} \\ 0 & \ldots & 0 & N_{t+k} & M_{t+k} \end{pmatrix},$$

where

$$M_s = -E\left[\frac{\partial^2 L}{\partial h_s^2}\right] = \frac{1}{2} + \frac{1}{V_s}\left(\frac{\partial \mu_s}{\partial h_s}\right)^2 + \frac{1}{V_{s-1}}\left(\frac{\partial \mu_{s-1}}{\partial h_s}\right)^2$$

$$\text{(B.3)} \qquad + \frac{\phi^2}{\sigma_\eta^2}\mathbb{I}(t + k < T), \qquad s = t+1, \ldots, t+k,$$

(B.4)
$$N_s = -E\left[\frac{\partial^2 L}{\partial h_s \partial h_{s-1}}\right] = \frac{1}{V_{s-1}}\frac{\partial \mu_{s-1}}{\partial h_{s-1}}\frac{\partial \mu_{s-1}}{\partial h_s},$$

where $s = 2, \ldots, T$ and $N_{t+1} = 0$. Next, we define

$$\text{(B.5)} \qquad \frac{\partial \mu_s}{\partial h_s} = \begin{cases} \beta_2 e^{h_s} + \frac{1}{2}e^{\frac{h_s}{2}}\delta(z_s - \mu_z) + \\ \frac{\varphi}{\varphi^2 + \tau^2}\sqrt{z_s}e^{\frac{h_s}{2}}\left[\frac{(h_{s+1} - \alpha - \phi h_s)}{2} - \phi\right] \\ \qquad s = 1, \ldots, T-1, \\ \beta_2 e^{h_s} + \frac{1}{2}e^{\frac{h_s}{2}}\delta(z_s - \mu_z), \\ \qquad s = T, \end{cases}$$

$$\text{(B.6)} \qquad \frac{\partial \mu_{s-1}}{\partial h_s} = \begin{cases} 0, & s = 1, \\ \frac{\varphi}{\varphi^2 + \tau^2}\sqrt{z_{s-1}}e^{\frac{h_{s-1}}{2}}, & s = 2, \ldots, T. \end{cases}$$

ACKNOWLEDGEMENTS

REFERENCES

[1] ABANTO-VALLE, C. A., BANDYOPADHYAY, D., LACHOS, V. H., and ENRIQUEZ, I. [2010], "Robust Bayesian analysis of heavy-tailed stochastic volatility models using scale mixtures of normal distributions," *Computational Statistics & Data Analysis*, 54, 2883–2898. MR2727721

[2] ABANTO-VALLE, C. A., DEY, D. K., and LACHOS, V. H. [2015], "Bayesian estimation of a skew-Student-t stochastic volatility model," *Methodology and Computing in Applied Probability*, 17, 721–738. MR3377857

[3] ABANTO-VALLE, C. A., MIGON, H. S., and LACHOS, V. H. [2011], "Stochastic volatility in mean models with scale mixtures of normal distributions and correlated errors: A Bayesian approach," *Journal of Statistical Planning and Inference*, 54, 1875–1887. MR2763217

[4] ABANTO-VALLE, C. A., MIGON, H. S., and LACHOS, V. H. [2012], "Stochastic volatility in mean models with heavy-tailed distributions," *Brazilian Journal of Probability and Statistics*, 26(4), 402–422. MR2949086

[5] ABANTO-VALLE, C. A., WANG, C., WANG, X., WANG, F.-X., and CHEN, M.-H. [2014], "Bayesian inference for stochastic volatility models using the generalized skew-t distribution with applications to the Shenzhen stock exchange return," *Statistical and Its Interface*, 7, 487–502. MR3302377

[6] ANDO, T. [2006], "Bayesian inference for nonlinear and nongaussian stochastic volatility model with leverge effect," *Journal of Japan Statistical Society*, 36, 173–197. MR2338878

[7] ANDO, T. [2007], "Bayesian predictive information criterion for the evaluation of hierarchical Bayesian and empirical Bayes models," *Biometrika*, 94, 443–458. MR2380571

[8] BOLLERSLEV, T., and ZHOU, H. [2005], "Volatility puzzles: A simple framework for gauging return-volatility regressions," *Journal of Econometrics*, 131, 123–150. MR2275997

[9] CELEUX, G., FORBES, F., ROBERT, C. P., and TITTERINGTON, D. M. [2006], "Deviance information criteria for missing data models," *Bayesian Analysis*, 1, 651–674. MR2282197

[10] CHIB, S. [1995], "Marginal likelihood from the Gibbs output," *Journal of the American Statistical Association*, 90, 1313–1321. MR1379473

[11] DE JONG, P., and SHEPHARD, N. [1995], "The simulation smoother for time series models," *Biometrika*, 82, 339–350. MR1354233

[12] DELATOLA, E.-I., and GRIFFIN, J. E. [2011], "Bayesian nonparametric modelling of the return distribution with stochastic volatility," *Bayesian Analysis*, 6, 901–926. MR2869968

[13] FONSECA, T. C. O., FERREIRA, M. A. R., and MIGON, H. S. [2008], "Objective Bayesian analysis for the Student-t regression model," *Biometrika*, 95, 325–333. MR2521587

[14] FRENCH, K. R., SCHERT, W. G., and STAMBUGH, R. F. [1987], "Expected stock return and volatility," *Journal of Financial Economics*, 19, 3–29.

[15] GERLACH, R., CARTER, C., and KOHN, R. [1999], "Diagnostics for time series analysis," *Journal of Time Series Analysis*, 20, 309–330. MR1693161

[16] GEWEKE, J. [1992], "Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments," in *Bayesian Statistics. Vol. 4.*, eds. J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, Oxford, U.K.: Oxford University Press, pp. 169–193. MR1380276

[17] GHYSELS, E., HARVEY, A. C., and RENAULT, E. [1996], "Stochastic volatility," in *Handbook of Statistics. Vol. 14.*, eds. G. Maddala, and C. R. Rao, Amsterdam: North-Holland, pp. 119–191. MR1602124

[18] Gneiting, T., and Raftery, A. E. [2007], "Strictly proper scoring rules, prediction and estimation," *Journal of the American Statistical Association*, 6, 901–926. MR2345548

[19] GOOD, I. J. [1952], "Rational decisions," *Journal of the Royal Statistical Society, Series B*, 14, 107–114. MR0077033

[20] HARVEY, A. C., and SHEPHARD, N. [1996], "The estimation of an asymmetric stochastic volatility model for asset returns," *Journal of Business and Economic Statistics*, 14, 429–434.

[21] JACQUIER, E., POLSON, N., and ROSSI, P. [1994], "Bayesian analysis of stochastic volatility models," *Journal of Business and Economic Statistics*, 12, 371–418.

[22] JACQUIER, E., POLSON, N., and ROSSI, P. [2004], "Bayesian analysis of stochastic volatility models with fat-tails and correlated errors," *Journal of Econometrics*, 122, 185–212. MR2083256

[23] KIM, S., SHEPARD, N., and CHIB, S. [1998], "Stochastic volatility: Likelihood inference and comparison with ARCH models," *Review of Economic Studies*, 65, 361–393.

[24] Kim, S., Shephard, N., and Chib, S. [1998], "Stochastic volatility: Likelihood inference and comparison with ARCH models," *Review of Economic Studies*, 65, 361–393.

[25] Koopman, S. [1993], "Disturbance smoothers for state space models," *Biometrika*, 80, 117–126. MR1225219

[26] Koopman, S. J., and Uspensky, E. H. [2002], "The stochastic volatility in mean model: Empirical evidence from international tock markets," *Journal of Applied Econometrics*, 17, 667–689.

[27] Liesenfeld, R., and Jung, R. C. [2000], "Stochastic volatility models: Conditional normality versus heavy-tailed distrutions," *Journal of Applied Econometrics*, 15, 137–160.

[28] Liesenfeld, R., and Richard, J.-F. [2003], "Univariate and multivariate stochastic volatility models: Estimation and diagnostics," *Journal of Empirical Finance*, 10, 505–531.

[29] Nakajima, J., and Omori, Y. [2012], "Stochastic volatility model with leverage and asymmetrically heavy-tailed error using GH skew Student's t-distribution," *Computational Statistics & Data Analysis*, 56, 3690–3704. MR2943921

[30] Omori, Y., Chib, S., Shephard, N., and Nakajima, J. [2007], "Stochastic volatility with leverage: Fast likelihood inference," *Journal of Econometrics*, 140, 425–449. MR2408914

[31] Omori, Y., and Watanabe, T. [2008], "Block sampler and posterior mode estimation for asymmetric stochastic volatility models," *Computational Statistics & Data Analysis*, 52, 2892–2910. MR2424767

[32] Pemstein, D., Quinn, K. V., and Martin, A. D. [2011], "The scythe statistical library: An open source C++ library for statistical computation," *Journal of Statistical Software*, 42, 1–26.

[33] Pitt, M., and Shephard, N. [1999], "Filtering via simulation: Auxiliary particle filter," *Journal of the American Statistical Association*, 94, 590–599. MR1702328

[34] Robert, C. P., and Titterington, D. M. [2002], "Discussion on "Bayesian measures of model complexity and fit"," *Biometrical Journal*, 64, 573–590.

[35] Rosenblatt, M. [1952], "Remarks on a multivariate transformation," *Annals of Mathematical Statistics*, 23, 470–472. MR0049525

[36] Salazar, E., Migon, H. S., and Ferreira, M. A. R. [2009], Objective Bayesian analysis for exponential power regression models, Technical report, Federal University of Rio de Janeiro, Departament of Statistics.

[37] Shephard, N., and Pitt, M. [1997], "Likelihood analysis of non-Gaussian measurements time series," *Biometrika*, 84, 653–667. MR1603940

[38] Smith, J. Q. [1985], "Diagnostic checks of non-standard time series models," *Journal of Forecasting*, 4, 283–291.

[39] Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and van der Linde, A. [2002], "Bayesian measures of model complexity and fit," *Journal of the Royal Statistical Society, Series B*, 64, 621–622. MR1979380

[40] Stone, M. [2002], "Discussion on "Bayesian measures of model complexity and fit"," *Journal of the Royal Statistical Society, Series B*, 64, 621.

[41] Taylor, S. [1982], "Financial returns modelled by the product of two stochastic processes-a study of the daily sugar prices 1961–75," in *Time Series Analysis: Theory and Practice, Vol 1*, ed. O. Anderson, Amsterdam: North-Holland, pp. 203–226.

[42] Taylor, S. [1986], *Modeling Financial Time Series*, Chichester: Wiley.

[43] Tierney, L. [1994], "Markov chains for exploring posterior distributions (with discussion)," *Annal of Statistics*, 21, 1701–1762. MR1329166

[44] Watanabe, T., and Omori, Y. [2004], "A multi-move sampler for estimate non-Gaussian time series model: Comments on Shepard and Pitt (1997)," *Biometrika*, 91, 246–248. MR2050475

[45] Yu, J. [2005], "On leverage in stochastic volatility model," *Journal of Econometrics*, 127, 165–178. MR2166061

William L. Leão
Departament of Statistics
Federal University of Rio de Janeiro
Caixa Postal 68530, CEP: 21945-970
Rio de Janeiro
Brazil
E-mail address: caifeng.wang@ustb.edu.cn

Carlos A. Abanto-Valle
Departament of Statistics
Federal University of Rio de Janeiro
Caixa Postal 68530, CEP: 21945-970
Rio de Janeiro
Brazil
E-mail address: cabantovalle@im.ufrj.br

Ming-Hui Chen
Department of Statistics
University of Connecticut
215 Glenbrook Rd, U-4120
Storrs, CT 06269
USA
E-mail address: ming-hui.chen@uconn.edu