Online Appendix for: Bayesian analysis of moving average stochastic volatility models: modelling in-mean effects and leverage for financial time series

Stefanos Dimitrakopoulos[1*] and Michalis Kolossiatis[**]

[*]Economics Division, Leeds University, Leeds, LS6 1AN, UK

[**]Department of Mathematics and Statistics, University of Cyprus, Nicosia, 1678, Cyprus

1 MCMC algorithms

1.1 MCMC algorithm for the MASVM model

The parameter vector for this model is $\boldsymbol{\theta} = \left(\mu, \mu_h, \phi, \sigma_\eta^2, \psi, \boldsymbol{h} = \{h_t\}_{t=1}^T, \lambda\right)$. For convenience, we first write down the likelihood function of the model:

$$f(\boldsymbol{y}|\boldsymbol{\theta}) = (2\pi)^{-T/2} \exp\{-\sum_{t=1}^{T} h_t/2\} \exp\{-(\boldsymbol{y} - \mu\mathbf{1}_T - \lambda\boldsymbol{u})'(H_\psi^{-1})'\Sigma^{-1}H_\psi^{-1}(\boldsymbol{y} - \mu\mathbf{1}_T - \lambda\boldsymbol{u})/2\}$$

where $\boldsymbol{y} = (y_1, y_2, \ldots, y_T)'$ is the data set, $\mathbf{1}_T = (1, 1, \ldots, 1)'$ is a $T \times 1$ vector of ones, $\Sigma = diag(e^{h_1}, e^{h_2}, \ldots, e^{h_T})$, $\boldsymbol{u} = (e^{h_1}, e^{h_2}, \ldots, e^{h_T})'$ and

$$H_\psi = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ \psi & 1 & 0 & \ldots & 0 \\ 0 & \psi & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \psi & 1 \end{pmatrix}.$$

From the volatility equation we also get

$$f(\boldsymbol{h}|\phi, \mu_h, \sigma_\eta^2) = (2\pi\sigma_\eta^2)^{-T/2}\sqrt{1-\phi^2} \exp\{-\frac{1}{2\sigma_\eta^2}\sum_{t=1}^{T-1}(h_{t+1} - \mu_h - \phi(h_t - \mu_h))^2 - \frac{1-\phi^2}{2\sigma_\eta^2}(h_1 - \mu_h)^2\}.$$

The full conditional distributions and related updating steps are as follows:

[1]Correspondence to: Stefanos Dimitrakopoulos, Economics Division, Leeds University Business School, Leeds University, UK, E-mail: s.dimitrakopoulos@leeds.ac.uk.

- $$f((\mu,\lambda)'|\cdots) \propto \exp\left\{-(\boldsymbol{y}-\mu\mathbf{1}_T-\lambda\boldsymbol{u})'(H_\psi^{-1})'\Sigma^{-1}H_\psi^{-1}(\boldsymbol{y}-\mu\mathbf{1}_T-\lambda\boldsymbol{u})/2-\frac{1}{2V_\mu}(\mu-\mu_0)^2-\frac{1}{2V_\lambda}(\lambda-\lambda_0)^2\right\}$$

  $$\Rightarrow (\mu,\lambda)'|\cdots \sim N(D\cdot S, D),$$

  where $S = \left((\mu_0/V_\mu, \lambda_0/V_\lambda)'+X'(H^{-1})'\Sigma^{-1}H^{-1}\boldsymbol{y}\right)$, $D = \left(diag(1/V_\mu, 1/V_\lambda)+X'(H^{-1})'\Sigma^{-1}H^{-1}X\right)^{-1}$ and $X = (\mathbf{1}_T, \boldsymbol{u})$.

  So, we update these jointly by drawing from the above bivariate normal distribution.

- $$f(\phi|\cdots) \propto \sqrt{1-\phi^2}\exp\left\{-\frac{1}{2\sigma_\eta^2}\sum_{t=1}^{T-1}(h_{t+1}-\mu_h-\phi(h_t-\mu_h))^2-\frac{1-\phi^2}{2\sigma_\eta^2}(h_1-\mu_h)^2-\frac{1}{2V_\phi}(\phi-\phi_0)^2\right\}1_{(-1<\phi<1)}$$

  This is therefore not of a known form, so we use a Metropolis-Hastings step with proposal density $N(m_\phi, B_\phi)1_{(-1<\phi<1)}$, with $B_\phi = (1/V_\phi + X_\phi^{'}X_\phi/\sigma_\eta^2)^{-1}$, $m_\phi = B_\phi \cdot (\phi_0/V_\phi + X_\phi^{'}Z_\phi/\sigma_\eta^2)$, where $X_\phi = (h_1-\mu_h, ..., h_{T-1}-\mu_h)^{'}$ and $Z_\phi = (h_2-\mu_h, ..., h_T-\mu_h)^{'}$.

- $$f(\sigma_\eta^2|\cdots) \propto (\sigma_\eta^2)^{-T/2}(\sigma_\eta^2)^{-\nu_h-1}\exp\left\{-\frac{1}{2\sigma_\eta^2}\sum_{t=1}^{T-1}(h_{t+1}-\mu_h-\phi(h_t-\mu_h))^2-\frac{1-\phi^2}{2\sigma_\eta^2}(h_1-\mu_h)^2-\frac{S_h}{2\sigma_\eta^2}\right\}$$

  Therefore, the full conditional distribution for σ_η^2 is

  $$IG\left(\nu_h+T/2, S_h+\sum_{t=1}^{T-1}(h_{t+1}-\mu_h-\phi(h_t-\mu_h))^2/2+(1-\phi^2)(h_1-\mu_h)/2\right),$$

  from which we can directly draw.

- $$f(\mu_h|\cdots) \propto \exp\left\{-\frac{1}{2\sigma_\eta^2}\sum_{t=1}^{T-1}(h_{t+1}-\mu_h-\phi(h_t-\mu_h))^2-\frac{1-\phi^2}{2\sigma_\eta^2}(h_1-\mu_h)^2-\frac{1}{2V_{\mu_h}}(\mu_h-\mu_{h0})^2\right\}$$

  Therefore, the full conditional distribution for μ_h is $N(D_{\mu_h}\cdot S_{\mu_h}, D_{\mu_h})$, where

  $$D_{\mu_h} = \left(1/V_{\mu_h}+\frac{1-\phi^2}{\sigma_\eta^2}+(T-1)(1-\phi)^2/\sigma_\eta^2\right)^{-1}$$

and

$$S_{\mu_h} = \left(\mu_{h0}/V_{\mu_h} + \frac{1-\phi^2}{\sigma_\eta^2}h_1 + \frac{1-\phi}{\sigma_\eta^2}\sum_{t=1}^{T-1}(h_{t+1}-\phi h_t)\right).$$

- $$f(\psi|\cdots) \propto \exp\left\{-(\boldsymbol{y}-\mu\mathbf{1}_T-\lambda\boldsymbol{u})'(H_\psi^{-1})'\Sigma^{-1}H_\psi^{-1}(\boldsymbol{y}-\mu\mathbf{1}_T-\lambda\boldsymbol{u})/2\right\}\exp\left\{-\frac{1}{2V_\psi}(\psi-\psi_0)^2\right\}1_{(-1<\psi<1)}$$

  where ψ appears also in the matrix H_ψ. This is not a standard distribution, so we update it again using a Metropolis-Hastings step as in Chan (2013).

- $$f(\boldsymbol{h}|\cdots) \propto f(\boldsymbol{y}|\boldsymbol{\theta})f(\boldsymbol{h}|\phi,\mu_h,\sigma_\eta^2)$$

  $$\propto \exp\left\{-\sum_{t=1}^{T}h_t/2 - (\boldsymbol{y}-\mu\mathbf{1}_T-\lambda\boldsymbol{u})'(H_\psi^{-1})'\Sigma^{-1}H_\psi^{-1}(\boldsymbol{y}-\mu\mathbf{1}_T-\lambda\boldsymbol{u})/2\right\}$$

  $$\times \exp\left\{-\frac{1}{2\sigma_\eta^2}\sum_{t=1}^{T-1}(h_{t+1}-\mu_h-\phi(h_t-\mu_h))^2 - \frac{1-\phi^2}{2\sigma_\eta^2}(h_1-\mu_h)^2\right\},$$

  where $\boldsymbol{h}$ appears also in $\boldsymbol{u}$.

The above expression is not of known form and quite complicated. To sample from it, we approximate it by a Gaussian distribution, which is then used as a proposal density within the Acceptance-Rejection Metropolis-Hastings (ARMH) algorithm (see, for example, Tierney (1994) and Chib and Greenberg (1995)). Candidate draws from the Gaussian approximation are obtained, using the precision sampler of Chan and Jeliazkov (2009), instead of Kalman filter-based methods.

In particular, the density $f(\boldsymbol{h}|\phi,\mu_h,\sigma_\eta^2)$ is Gaussian, that is, $\boldsymbol{h}|\phi,\mu_h,\sigma_\eta^2 \sim N(H_\phi^{-1}\hat{\boldsymbol{h}},(H_\phi'\hat{\Sigma}^{-1}H_\phi)^{-1})$, where $\hat{\boldsymbol{h}} = (\mu_h,(1-\phi)\mu_h,..,(1-\phi)\mu_h)'$, $\hat{\Sigma} = diag(\sigma_\eta^2/(1-\phi^2),\sigma_\eta^2,...,\sigma_\eta^2)$ and H_ϕ is a lower triangular sparse matrix (with $det(H_\phi)=1$ - hence, it is invertible):

$$H_\phi = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -\phi & 1 & 0 & \cdots & 0 \\ 0 & -\phi & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -\phi & 1 \end{pmatrix}.$$

The logarithm of $f(\boldsymbol{h}|\phi,\mu_h,\sigma_\eta^2)$ can be written as

$$\log p(\boldsymbol{h}|\phi,\mu_h,\sigma_\eta^2) = constant - \frac{1}{2}(\boldsymbol{h}'H_\phi'\hat{\Sigma}^{-1}H_\phi\boldsymbol{h} - 2\boldsymbol{h}'H_\phi'\hat{\Sigma}^{-1}\hat{\boldsymbol{h}}).$$

The density $f(\boldsymbol{y}|\boldsymbol{\theta})$ can also be approximated by a normal density. By taking the second order Taylor expansion of the logarithm of $f(\boldsymbol{y}|\boldsymbol{\theta})$ around $\tilde{\boldsymbol{h}}$, which is the mode of the posterior $\log f(\boldsymbol{h}|\cdots)$ (see below), we have

$$\log f(\boldsymbol{y}|\boldsymbol{\theta}) \approx \log f(\boldsymbol{y}|\boldsymbol{\theta},\tilde{\boldsymbol{h}}) + (\boldsymbol{h}-\tilde{\boldsymbol{h}})'\mathbf{f} - \frac{1}{2}(\boldsymbol{h}-\tilde{\boldsymbol{h}})'G(\boldsymbol{h}-\tilde{\boldsymbol{h}})$$

$$= constant - \frac{1}{2}(\boldsymbol{h}'G\boldsymbol{h} - 2\boldsymbol{h}'(\mathbf{f}+G\tilde{\boldsymbol{h}})),$$

where $\mathbf{f} = (f_1, ..., f_T)'$ is the gradient vector with

$$f_t = \frac{\partial}{\partial h_t}\log f(y_t|\mu,\lambda,h_t,\psi) = -\frac{1}{2}+\frac{1}{2}e^{-h_t}(H_\psi^{-1}(\boldsymbol{y}-\mu\mathbf{1}_T))'_{(t)}(H_\psi^{-1}(\boldsymbol{y}-\mu\mathbf{1}_T))_{(t)}-\frac{1}{2}\lambda^2((H_\psi^{-1})'H_\psi^{-1}\boldsymbol{u})_{(t)},$$

$t = 1, ..., T$, evaluated at $\boldsymbol{h} = \tilde{\boldsymbol{h}}$, and $G = diag(G_1, ..., G_T)$ is the diagonal negative Hessian of the $\log f(\boldsymbol{y}|\boldsymbol{\theta})$, with

$$G_t = -\frac{\partial^2}{\partial h_t^2}\log f(y_t|\mu,\lambda,h_t,\psi) = +\frac{1}{2}e^{-h_t}(H_\psi^{-1}(\boldsymbol{y}-\mu\mathbf{1}_T))'_{(t)}(H_\psi^{-1}(\boldsymbol{y}-\mu\mathbf{1}_T))_{(t)}+\frac{1}{2}\lambda^2((H_\psi^{-1})'H_\psi^{-1}\boldsymbol{u})_{(t)},$$

$t = 1, ..., T$, evaluated at $\boldsymbol{h} = \tilde{\boldsymbol{h}}$, and where $\boldsymbol{v}_{(i)}$ denotes the i-th element of vector $\boldsymbol{v}$.

The logarithm of the posterior distribution of the volatility vector therefore becomes

$$\log f(\boldsymbol{h}|\cdots) \approx constant - \frac{1}{2}(\boldsymbol{h}'\boldsymbol{K_h}\boldsymbol{h} - 2\boldsymbol{h}'\boldsymbol{k_h}) =: \log\ q(\boldsymbol{h}),$$

where $\boldsymbol{K_h} = H_\phi'\hat{\Sigma}^{-1}H_\phi + G$, $\boldsymbol{k_h} = \mathbf{f} + G\tilde{\boldsymbol{h}} + H_\phi'\hat{\Sigma}^{-1}\hat{\boldsymbol{h}}$ and $q(\boldsymbol{h}) \propto N(\hat{\boldsymbol{m}}, \boldsymbol{K_h}^{-1})$, with $\hat{\boldsymbol{m}} = \boldsymbol{K_h}^{-1}\boldsymbol{k_h}$. In other words, the posterior distribution of the volatility vector can be approximated by a normal density with mean $\hat{\boldsymbol{m}}$ and precision matrix $\boldsymbol{K_h}$.

For efficient sampling, the point $\tilde{\boldsymbol{h}}$, around which the second order Taylor expansion is taken, is the mode of the posterior $\log f(\boldsymbol{h}|\cdots)$. The negative Hessian of this posterior distribution evaluated at $\boldsymbol{h} = \tilde{\boldsymbol{h}}$ is $\boldsymbol{K_h}$ and the gradient evaluated at $\boldsymbol{h} = \tilde{\boldsymbol{h}}$ is -$\boldsymbol{K_h}\tilde{\boldsymbol{h}}$ +$\boldsymbol{k_h}$. To find the mode, we apply the Newton-Raphson method as follows: 1) Initialize $\boldsymbol{h} = \tilde{\boldsymbol{h}}^{(1)}$ for some constant vector $\tilde{\boldsymbol{h}}^{(1)}$. 2) Set $\tilde{\boldsymbol{h}} = \tilde{\boldsymbol{h}}^{(l)}$ for $l = 1, 2, ...,$ and compute $\boldsymbol{K_h}$, $\boldsymbol{k_h}$ and $\boldsymbol{h}^{(l+1)} = \boldsymbol{h}^{(l)} + \boldsymbol{K_h}^{-1}(-\boldsymbol{K_h}\boldsymbol{h}^{(l)} + \boldsymbol{k_h}) = \boldsymbol{K_h}^{-1}\boldsymbol{k_h}$. This process is repeated until convergence is achieved.

1.2 MCMC algorithm for the MASVL model

The parameter vector for this model is $\boldsymbol{\theta} = \left(\mu, \mu_h, \phi, \sigma_\eta^2, \psi, \boldsymbol{h} = \{h_t\}_{t=1}^{T+1}, \rho\right)$. An additional complication is that $\boldsymbol{y} = (y_1, y_2, \ldots, y_T)'$ and $\boldsymbol{h}$ are not independent (because of the leverage part), nor are $y_t, y_s,\ t \neq s$ (because of the MA part). It is therefore easier to work with the transformed $\tilde{\boldsymbol{y}} = H_\psi(\boldsymbol{y} - \mu \mathbf{1}_T)$, where H_ψ and $\mathbf{1}_T$ are defined as in the MCMC algorithm for the MASVM model. The joint distribution of $\tilde{\boldsymbol{y}}$ and $\boldsymbol{h}$ is

$$f(\tilde{\boldsymbol{y}}, \boldsymbol{h}|\boldsymbol{\theta}) = f(h_1)\prod_{t=1}^{T} f(\tilde{y}_t, h_{t+1}|\mu_h, \phi, \sigma_\eta^2, \psi, \rho, \{h_j\}_{j<t+1}) = (2\pi)^{T+1}\frac{\sqrt{1-\phi^2}}{\sigma_\eta}(\sigma_\eta^2(1-\rho^2))^{-T/2}\exp\{-\sum_{t=1}^{T} h_t/2\}$$

$$\times \exp\left\{-\frac{1-\phi^2}{2\sigma_\eta^2}h_1 - \left(\begin{pmatrix}\tilde{y}_t \\ h_{t+1}\end{pmatrix}' - \begin{pmatrix}0 \\ \mu_h + \phi(h_t-\mu_h)\end{pmatrix}\right)' * S^{-1} * \left(\begin{pmatrix}\tilde{y}_t \\ h_{t+1}\end{pmatrix}' - \begin{pmatrix}0 \\ \mu_h + \phi(h_t-\mu_h)\end{pmatrix}\right)/2\right\}$$

where $S = \begin{pmatrix} e^{h_t} & \rho\sigma_\eta e^{h_t/2} \\ \rho\sigma_\eta e^{h_t/2} & \sigma_\eta^2 \end{pmatrix}$.

From this, it is straightforward to get the conditional distribution of $\tilde{\boldsymbol{y}}$ given $\boldsymbol{h}$ and $\boldsymbol{\theta}$ which is,

$$f(\tilde{\boldsymbol{y}}|\boldsymbol{h},\boldsymbol{\theta}) \propto \exp\left\{-\frac{1}{2(1-\rho^2)}\sum_{t=1}^{T} e^{-h_t}(\tilde{y}_t - \rho\frac{e^{h_t/2}}{\sigma_\eta}(h_{t+1}-\mu_h-\phi(h_t-\mu_h)))^2\right\}.$$

The full conditional distributions are now as follows:

- $$\mu|\cdots \sim N(D_\mu \cdot S_\mu, D_\mu),$$

  where $S_\mu = (\mu_0/V_\mu + \mathbf{1}_T'(H_\psi^{-1})'\Sigma^{-1}(H_\psi^{-1}\boldsymbol{y} - \frac{\rho}{\sigma_\eta}\exp\{h_t/2\}(\boldsymbol{h}^+ - \mu_h\mathbf{1}_T - \phi(\boldsymbol{h}^- - \mu_h\mathbf{1}_T)))/(1-\rho^2))$,
  $D_\mu = (1/V_\mu + \mathbf{1}_T'(H_\psi^{-1})'\Sigma^{-1}H_\psi^{-1}\mathbf{1}_T/((1-\rho^2)))^{-1}$, $\Sigma = diag(e^{h_1}, e^{h_2}, \ldots, e^{h_T})$,
  $\boldsymbol{h}^+ = (h_2, \ldots, h_{T+1})$ and $\boldsymbol{h}^- = (h_1, \ldots, h_T)$.

- $$f(\phi|\cdots) \propto \exp\left\{-\frac{1}{2(1-\rho^2)}\sum_{t=1}^{T} e^{-h_t}(\tilde{y}_t - \rho\frac{e^{h_t/2}}{\sigma_\eta}(h_{t+1}-\mu_h-\phi(h_t-\mu_h)))^2\right\}\times$$

  $$\sqrt{1-\phi^2}\exp\left\{-\frac{1}{2\sigma_\eta^2}\sum_{t=1}^{T}(h_{t+1}-\mu_h-\phi(h_t-\mu_h))^2 - \frac{1-\phi^2}{2\sigma_\eta^2}(h_1-\mu_h)^2 - \frac{1}{2V_\phi}(\phi-\phi_0)^2\right\}1_{(-1<\phi<1)}$$

  To update ϕ we use a Metropolis-Hastings algorithm with proposal density $N(m_\phi, B_\phi)1_{(-1<\phi<1)}$, where

  $$B_\phi = \left(1/V_\phi + \sum_{t=1}^{T}(h_t-\mu_h)^2/\sigma_\eta^2 + \frac{\rho^2}{\sigma_\eta^2}(\exp\{\frac{\boldsymbol{h}^-}{2}\}'(\boldsymbol{h}^- - \mu_h\mathbf{1}_T)')\Sigma_\rho^{-1}(\exp\{\frac{\boldsymbol{h}^-}{2}\}(\boldsymbol{h}^- - \mu_h\mathbf{1}_T))\right)^{-1},$$

  $$m_\phi = B_\phi\cdot\left(\frac{\phi_0}{V_\phi} + \frac{(\boldsymbol{h}^- - \mu_h\mathbf{1}_T)'(\boldsymbol{h}^+ - \mu_h\mathbf{1}_T)}{\sigma_\eta^2} - \frac{\rho}{\sigma_\eta}(\exp\{\frac{\boldsymbol{h}^-}{2}\}'(\boldsymbol{h}^- - \mu_h\mathbf{1}_T)')\Sigma_\rho^{-1}(\tilde{\boldsymbol{y}} - \frac{\rho}{\sigma_\eta}(\exp\{\frac{\boldsymbol{h}^-}{2}\}(\boldsymbol{h}^+ - \mu_h\mathbf{1}_T)))\right)$$

  and $\Sigma_\rho = diag\left(e^{h_1}(1-\rho^2), e^{h_2}(1-\rho^2), \ldots, e^{h_T}(1-\rho^2)\right)$.

-

$$f(\sigma_\eta^2|\cdots) \propto \exp\left\{-\frac{1}{2(1-\rho^2)}\sum_{t=1}^{T} e^{-h_t}\left(\tilde{y}_t - \rho\frac{e^{h_t/2}}{\sigma_\eta}(h_{t+1}-\mu_h-\phi(h_t-\mu_h))\right)^2\right\} \times$$

$$(\sigma_\eta^2)^{-T/2}(\sigma_\eta^2)^{-\nu_h-1}\exp\left\{-\frac{1}{2\sigma_\eta^2}\sum_{t=1}^{T}(h_{t+1}-\mu_h-\phi(h_t-\mu_h))^2 - \frac{1-\phi^2}{2\sigma_\eta^2}(h_1-\mu_h)^2 - \frac{S_h}{2\sigma_\eta^2}\right\}$$

As opposed to the MASVM model, this distributional form does not correspond to a known distribution. Therefore, to update the value of σ_η^2 we again use Metropolis-Hastings with proposal density $IG(\nu_h+(T+1)/2, S_h+(1-\phi^2)(h_1-\mu_h)^2/2+e_h^{'}e_h/2)$, where $e_h = \boldsymbol{h}^+ - \phi\boldsymbol{h}^- - (1-\phi)\mu_h\mathbf{1}_T$.

-

$$f(\mu_h|\cdots) \propto \exp\left\{-\frac{1}{2(1-\rho^2)}\sum_{t=1}^{T} e^{-h_t}\left(\tilde{y}_t - \rho\frac{e^{h_t/2}}{\sigma_\eta}(h_{t+1}-\mu_h-\phi(h_t-\mu_h))\right)^2\right\} \times$$

$$\exp\left\{-\frac{1}{2\sigma_\eta^2}\sum_{t=1}^{T}(h_{t+1}-\mu_h-\phi(h_t-\mu_h))^2 - \frac{1-\phi^2}{2\sigma_\eta^2}(h_1-\mu_h)^2 - \frac{1}{2V_{\mu_h}}(\mu_h-\mu_{h0})^2\right\}$$

Therefore, the full conditional distribution for μ_h is $N(D_{\mu_h}\cdot S_{\mu_h}, D_{\mu_h})$, where

$D_{\mu_h} = \left(1/V_{\mu_h} + T(1-\phi)^2/\sigma_\eta^2 + (1-\phi^2)/\sigma_\eta^2 + \frac{(1-\phi)^2\rho^2}{\sigma_\eta^2}\exp\{-\frac{\boldsymbol{h}^-}{2}\}'\Sigma_\rho^{-1}\exp[\frac{\boldsymbol{h}^-}{2}]\right)^{-1}$ and

$S_{\mu_h} = \left(\mu_{h0}/V_{\mu_h} + \frac{1-\phi^2}{\sigma_\eta^2}h_1 + \frac{1-\phi}{\sigma_\eta^2}\sum_{t=1}^{T}(h_{t+1}-\phi h_t) - \frac{\rho(1-\phi)}{\sigma_\eta}\exp\{-\frac{\boldsymbol{h}^-}{2}\}'\Sigma_\rho^{-1}(\tilde{\boldsymbol{y}} - \frac{\rho}{\sigma_\eta}\exp\{\frac{\boldsymbol{h}^-}{2}\}(\boldsymbol{h}^+-\phi\boldsymbol{h}^-)\right)$.

-

$$f(\psi|\cdots) \propto \exp\left\{-\frac{1}{2(1-\rho^2)}\sum_{t=1}^{T} e^{-h_t}\left(\tilde{y}_t - \rho\frac{e^{h_t/2}}{\sigma_\eta}(h_{t+1}-\mu_h-\phi(h_t-\mu_h))\right)^2 - \frac{(\psi-\psi_0)^2}{2V_\psi}\right\} 1_{(-1<\psi<1)}$$

where ψ also appears in $\tilde{\boldsymbol{y}} = H_\psi(\boldsymbol{y}-\mu\mathbf{1}_T)$. This is not a standard distribution, so we update it using a Metropolis-Hastings step as in Chan (2013).

-

$$f(\boldsymbol{h}|\cdots) \propto \exp\left\{-\frac{1}{2(1-\rho^2)}\sum_{t=1}^{T} e^{-h_t}(\tilde{y}_t - \rho\frac{e^{h_t/2}}{\sigma_\eta}(h_{t+1}-\mu_h-\phi(h_t-\mu_h)))^2\right\} \times$$

$$\exp\left\{-\sum_{j=1}^{T}h_j/2 - \frac{1}{2\sigma_\eta^2}\sum_{t=1}^{T}(h_{t+1}-\mu_h-\phi(h_t-\mu_h))^2 - \frac{1-\phi^2}{2\sigma_\eta^2}(h_1-\mu_h)^2\right\}.$$

To update the volatility vector we use the same method as the one used in the MASVM model. We only need to modify the algorithm as follows: The gradient vector $\mathbf{f} = (f_1, ..., f_{T+1})'$ and the negative Hessian matrix

$$G = \begin{pmatrix} G_{11} & G_{12} & 0 & \cdots & 0 \\ G_{21} & G_{22} & G_{23} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & G_{T,T-1} & G_{TT} & G_{T,T+1} \\ 0 & \cdots & 0 & G_{T+1,T} & G_{T+1,T+1} \end{pmatrix},$$

are calculated as follows:

The logarithm of the conditional distribution $p(\tilde{y}_t|h_t, h_{t+1}, \mu, \phi, \rho, \psi, \mu_h, \sigma_\eta^2)$ is given by
$\log p(\tilde{y}_t|h_t, h_{t+1}, \mu, \phi, \rho, \psi, \mu_h, \sigma_\eta^2) = -\frac{1}{2(1-\rho^2)} \sum_{t=1}^T e^{-h_t} (\tilde{y}_t - \rho \frac{e^{h_t/2}}{\sigma_\eta} (h_{t+1} - \mu_h - \phi(h_t - \mu_h)))^2$
$- \frac{1}{2} \log \left(2\pi(1-\rho^2)\right) - \frac{h_t}{2}$.

Setting $p_t = p(\tilde{y}_t|h_t, h_{t+1}, \mu, \phi, \rho, \psi, \mu_h, \sigma_\eta^2)$ for notational convenience, we have
$f_1 = \frac{d \log p_t}{dh_t}, \quad f_t = \frac{d(\log p_t + \log p_{t-1})}{dh_t}$,
$G_{11} = -\frac{d^2 \log p_t}{dh_t^2}, \quad G_{tt} = -\frac{d^2(\log p_t + \log p_{t-1})}{dh_t^2}, \quad G_{t-1,t} = -\frac{d^2 \log p_t}{dh_t dh_{t+1}}$,
for $t = 2, ..., T+1$, evaluated at $\boldsymbol{h} = \tilde{\boldsymbol{h}}$, where

$$\begin{aligned}
\frac{d \log p_t}{dh_t} &= -\frac{1}{2} - \frac{1}{2(1-\rho^2)} \Big(-\frac{\tilde{y}_t^2}{\exp(h_t)} - 2\rho^2\phi(h_{t+1} - \phi h_t - \mu_h(1-\phi))/\sigma_\eta^2 \\
&\quad + \frac{\tilde{y}_t \rho}{\exp(h_t/2)\sigma_\eta} (h_{t+1} - \phi h_t - \mu_h(1-\phi) + 2\phi) \Big), \\
\frac{d^2 \log p_t}{dh_t^2} &= -\frac{1}{2(1-\rho^2)} \left(\frac{\tilde{y}_t^2}{\exp(h_t)} + 2\rho^2\phi^2/\sigma_\eta^2 - \frac{\tilde{y}_t \rho}{2\exp(h_t/2)\sigma_\eta} (h_{t+1} - \phi h_t - \mu_h(1-\phi) + 4\phi) \right), \\
\frac{d \log p_t}{dh_{t+1}} &= \frac{1}{(1-\rho^2)} \left(-\rho^2(h_{t+1} - \phi h_t - \mu_h(1-\phi))/\sigma_\eta^2 + \frac{\tilde{y}_t \rho}{\exp(h_t/2)\sigma_\eta} \right), \\
\frac{d^2 \log p_t}{dh_{t+1}^2} &= -\frac{\rho^2}{(1-\rho^2)\sigma_\eta^2}, \\
\frac{d^2 \log p_t}{dh_t h_{t+1}} &= \frac{1}{(1-\rho^2)} \left(\rho^2\phi/\sigma_\eta^2 - \frac{\tilde{y}_t \rho}{2\sigma_\eta \exp(h_t/2)} \right).
\end{aligned}$$

- $$f(\rho|\cdots) \propto \exp\left\{ -\frac{1}{2(1-\rho^2)} \sum_{t=1}^T e^{-h_t} \left(\tilde{y}_t - \rho \frac{e^{h_t/2}}{\sigma_\eta} (h_{t+1} - \mu_h - \phi(h_t - \mu_h)) \right)^2 \right\}$$
$$\times (1-\rho^2)^{-T/2} \exp\{-\frac{1}{2V_\rho}(\rho - \rho_0)^2\} 1_{(-1<\rho<1)}.$$

  This is not a distribution we can directly draw from. However, since ρ is defined in the unit interval, we can use the Griddy-Gibbs method.

1.3 MCMC algorithm for the SVML model

The SV in mean model with leverage (SVML) is given by

$$y_t = \mu + \lambda e^{h_t} + e^{h_t/2}\epsilon_t, \; t = 1, \ldots, T$$

$$h_{t+1} = \mu_h + \phi(h_t - \mu_h) + \eta_t,\ |\phi| < 1$$

where the joint distribution of the errors ϵ_t and η_t is bivariate normal; namely,

$$]\begin{pmatrix} \epsilon_t \\ \eta_t \end{pmatrix} \sim N_2\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} e^{h_t} & \rho\sigma_\eta e^{h_t/2} \\ \rho\sigma_\eta e^{h_t/2} & \sigma_\eta^2 \end{pmatrix}\right).$$

This model was proposed by Abanto-Valle et al. (2011). The MCMC algorithm for the SVML model is straightforward. The step that requires some modifications is the update of the volatility vector, in which the derivatives that appear in the gradient vector and the Hessian matrix take the following form:

$$\begin{aligned}
\frac{d\log p_t}{dh_t} &= -\frac{1}{2} - \frac{1}{2(1-\rho^2)}\Big(-\tfrac{(y_t-\mu-\lambda\exp(h_t))^2}{\exp(h_t)} - 2\lambda(y_t-\mu-\lambda\exp(h_t)) - 2\rho^2\phi(h_{t+1}-\phi h_t-\mu_h(1-\phi))/\sigma_\eta^2 \\
&\quad + \tfrac{(y_t-\mu)\rho}{\exp(h_t/2)\sigma_\eta}(h_{t+1}-\phi h_t-\mu_h(1-\phi)+2\phi) + \tfrac{\lambda\rho\exp(h_t/2)}{\sigma_\eta}(h_{t+1}-\phi h_t-\mu_h(1-\phi)-2\phi)\Big), \\
\frac{d^2\log p_t}{dh_t^2} &= -\frac{1}{2(1-\rho^2)}\Big(\tfrac{(y_t-\mu-\lambda\exp(h_t))^2}{\exp(h_t)} + 2\lambda(y_t-\mu-\lambda\exp(h_t)) + 2\lambda^2\exp(h_t) + 2\rho^2\phi^2/\sigma_\eta^2 \\
&\quad - \tfrac{(y_t-\mu)\rho}{2\exp(h_t/2)\sigma_\eta}(h_{t+1}-\phi h_t-\mu_h(1-\phi)+4\phi) + \tfrac{\lambda\rho\exp(h_t/2)}{2\sigma_\eta}(h_{t+1}-\phi h_t-\mu_h(1-\phi)-4\phi)\Big), \\
\frac{d\log p_t}{dh_{t+1}} &= \frac{1}{(1-\rho^2)}\Big(-\rho^2(h_{t+1}-\phi h_t-\mu_h(1-\phi))/\sigma_\eta^2 + \tfrac{(y_t-\mu-\lambda\exp(h_t))\rho}{\exp(h_t/2)\sigma_\eta}\Big), \\
\frac{d^2\log p_t}{dh_{t+1}^2} &= -\frac{\rho^2}{(1-\rho^2)\sigma_\eta^2}, \\
\frac{d^2\log p_t}{dh_th_{t+1}} &= \frac{1}{(1-\rho^2)}\Big(\rho^2\phi/\sigma_\eta^2 - \tfrac{(y_t-\mu)\rho}{2\sigma_\eta\exp(h_t/2)} - \tfrac{\lambda\rho\exp(h_t/2)}{\sigma_\eta}\Big).
\end{aligned}$$

2 Simulated data

In this section we conduct two simulation experiments to evaluate the accuracy of the estimation methods for the two proposed models. In each experiment we created data from the relevant proposed model and then applied the MCMC algorithm of the corresponding model. For each simulation exercise, we also applied the same data to the nested versions of the corresponding proposed model, in order to assess the contribution of the model components.

In our MCMC simulations we used 50000 iterations, after discarding the initial 30000 runs (burn-in period). For comparison purposes, we calculated the marginal likelihood and the observed-data DIC for the models of interest. For the marginal likelihood the observed-data likelihood is evaluated, using $M_1 = 1000$ and $M_2 = 50$ draws from the importance densities. To compute the observed-data DIC value we run three parallel chains and then we took the average of these DIC estimates.

2.1 Numerical illustration for the MASVM model

We generated T=800 observations from model (1)-(3) of the main paper with the following parameter values: $\mu = 0.2, \lambda = 0.7, \psi = 0.4, \mu_h = 0.2, \phi = 0.98$ and $\sigma_\eta^2 = 0.01$.

We applied these data to three different models: the proposed MASVM model and its special cases, the MASV (*i.e.* for $\lambda = 0$) and SVM (*i.e.* for $\psi = 0$) models. We used the same, quite uninformative priors for most of the common parameters in these models. Specifically, we took

$$\mu \sim N(0, 10), \mu_h \sim N(0, 10), \phi \sim N(0.97, 0.01)1_{(-1<\phi<1)}, \sigma_\eta^2 \sim IG(5, 0.16).$$

For the other parameters, we used

$$\lambda \sim N(0, 100), \psi \sim N(0, 100)1_{(-1<\psi<1)}.$$

Table 1 presents the posterior means and standard deviations of the parameters for the three models.

Table 1: Simulation results for the MASVM model and its nested versions. Standard deviations in parentheses.

Model	MASV	SVM	MASVM
True value	Mean (stand. dev.)	Mean (stand. dev.)	Mean (stand. dev.)
$\mu = 0.2$	1.058 (0.057)	-0.244 (0.542)	0.274 (0.163)
$\mu_h = 0.2$	0.323 (0.200)	0.267 (0.122)	0.274 (0.181)
$\lambda = 0.7$		1.043 (0.596)	0.651 (0.135)
$\psi = 0.4$	0.383 (0.033)		0.357 (0.035)
$\phi = 0.98$	0.962 (0.017)	0.901 (0.051)	0.960 (0.017)
$\sigma_\eta^2 = 0.01$	0.023 (0.007)	0.018 (0.004)	0.021 (0.006)
Log ML	-1289.0	-1320.8	-1280.3
DIC_{obs}	2563.5	2619.4	2559.8

For the proposed MASVM model, based on which the data were generated, we get fairly good estimates for all parameters. However, this is not the case for the rest of the models. In particular, the MASV model fails to satisfactorily estimate the parameters μ and μ_h, while the posterior means of μ and λ are quite far away from their true values in the SVM model.

As expected, the MASVM model has the best fit to the simulated data as it produces the largest log marginal likelihood value and the smallest observed-data DIC value. As far as the MASV and SVM models are concerned, the MASV model has better fit to the data than the SVM model, under both criteria.

As a robustness exercise, we also employed the MASVM model with a beta $B(20, 1.5)$ prior for ϕ, (as in Nakajima and Omori (2012)), as well as the more vague $N(0, 1)1_{(-1<\phi<1)}$ prior for the same parameter. In both cases the simulation results were essentially the same with the ones above. These results are presented in Table 2.

As we mention in the main paper (see section 2.3), higher ϕ could lead to higher λ. To capture this dependence we use the following joint prior on (ϕ, λ):

$$p(\phi, \lambda) = p(\phi)p(\lambda|\varphi) = N(\phi_0, V_\phi)1_{(-1<\phi<1)}\ N\left(\log\left(\frac{1+\phi}{1-\phi}\right), V_\lambda\right). \quad (7)$$

In other words, ϕ is given the same prior as before, whereas the prior of λ is a normal distribution, centered at the value $\log(\frac{1+\phi}{1-\phi})$ (which is an increasing function of ϕ). Alternatively, one could use $tanh(\phi)$ instead of $\log(\frac{1+\phi}{1-\phi})$.

In the MCMC code, the only changes will be the full conditionals of λ and ϕ. For the full conditional of λ, we just need to replace λ_0 with $\log(\frac{1+\phi}{1-\phi})$ (or $tanh(\phi)$), so we can draw (μ, λ) jointly, as we do in the MASVM algorithm, with this substitution.

For ϕ, the posterior expression becomes

$$f(\phi|\cdots) \propto \exp\left\{-\frac{1}{2\sigma_\eta^2}\sum_{t=1}^{T-1}(h_{t+1}-\mu_h-\phi(h_t-\mu_h))^2-\frac{1-\phi^2}{2\sigma_\eta^2}(h_1-\mu_h)^2-\frac{1}{2V_\phi}(\phi-\phi_0)^2-\frac{1}{2V_\lambda}\left(\lambda-\log(\tfrac{1+\phi}{1-\phi})\right)^2\right\}$$
$$\times\sqrt{1-\phi^2}\ 1_{(-1<\phi<1)}.$$

The only difference from above is the last part inside the exponent. This is therefore not of a known form, so we use a Metropolis-Hastings step.The simulation results with this joint prior are presented in the last column of Table 2.

Table 2: Simulation results for the MASVM model with alternative priors. Standard deviations in parentheses.

Model	MASVM	MASVM	MASVM
Prior for ϕ	$B(20, 1.5)$	$N(0, 1)1_{(-1<\phi<1)}$	$p(\lambda, \phi)$
True value	Mean (stand. dev.)	Mean (stand. dev.)	Mean (stand. dev.)
$\mu = 0.2$	0.283 (0.148)	0.260 (0.156)	0.299 (0.209)
$\mu_h = 0.2$	0.146 (0.200)	0.145 (0.163)	0.267 (0.156)
$\lambda = 0.7$	0.624 (0.134)	0.645 (0.144)	0.617 (0.158)
$\psi = 0.4$	0.363 (0.036)	0.387(0.036)	0.355 (0.035)
$\phi = 0.98$	0.968 (0.012)	0.954 (0.020)	0.956 (0.017)
$\sigma_\eta^2 = 0.01$	0.019 (0.005)	0.0203 (0.005)	0.021 (0.008)

2.2 Numerical illustration for the MASVL model

We constructed T=800 data points from model (4)-(6) of the main paper, using the following parameter values: $\mu = 0.2, \rho = -0.5, \psi = 0.6, \mu_h = -2, \phi = 0.94$ and $\sigma_\eta^2 = 0.01$.

We applied these data to three different models: the proposed MASVL model and its special cases, the MASV (*i.e.* for $\rho = 0$) and SVL (*i.e.* for $\psi = 0$) models. We used the same, quite uninformative priors for most of the common parameters in these models. Specifically, we took

$$\mu \sim N(0, 10), \mu_h \sim N(0, 10), \phi \sim N(0.97, 0.01)1_{(-1<\phi<1)}, \sigma_\eta^2 \sim IG(5, 0.16).$$

For the other parameters, we used

$$\psi \sim N(0, 100)1_{(-1<\psi<1)}, \rho \sim N(0, 1)1_{(-1<\rho<1)}.$$

Table 3 presents the posterior means and standard deviations of the parameters for the three models.

Table 3: Simulation results for the MASVL model and its nested versions. Standard deviations in parentheses.

Model	MASV	SVL	MASVL
True value	Mean (stand. dev.)	Mean (stand. dev.)	Mean (stand. dev.)
$\mu = 0.2$	0.163 (0.022)	0.162 (0.015)	0.172 (0.022)
$\mu_h = -2$	-1.925 (0.099)	-1.670 (0.112)	-1.973 (0.100)
$\psi = 0.6$	0.639 (0.028)		0.620 (0.028)
$\phi = 0.94$	0.912 (0.036)	0.929 (0.027)	0.933 (0.031)
$\rho = -0.5$		-0.329 (0.139)	-0.522 (0.167)
$\sigma_\eta^2 = 0.01$	0.030 (0.011)	0.032 (0.013)	0.019 (0.009)
Log ML	-379.9	-504.7	-376.9
DIC_{obs}	742.4	994.6	734.2

The posterior means for all parameters in the proposed MASVL model are closer to the actual values of the parameters than those of the nested models MASV and SVL. For example, the SVL model does not estimate ρ well. Also, both the MASV and SVL models substantially overestimate the error variance σ_η^2.

In terms of the log marginal likelihood and the DIC_{obs}, the MASVL model is performing best, as expected. Regarding the other two models, both criteria significantly favour the MASV model.

As for the case of the MASVM model, we used the beta prior $B(20, 1.5)$ and the vague prior $N(0, 1)1_{(-1<\phi<1)}$ for parameter ϕ of the MASVL model. The results are presented in Table 4, and correspond to no substantial change from the previous priors.

Table 4: Simulation results for the MASVL model with alternative priors. Standard deviations in parentheses.

Model	MASVL	MASVL
Prior for ϕ	$B(20, 1.5)$	$N(0, 1)1_{(-1<\phi<1)}$
True value	Mean (stand. dev.)	Mean (stand. dev.)
$\mu = 0.2$	0.208 (0.022)	0.207 (0.021)
$\mu_h = -2$	-1.987 (0.071)	-1.979 (0.067)
$\psi = 0.6$	0.604 (0.030)	0.603 (0.030)
$\phi = 0.94$	0.951 (0.070)	0.950 (0.083)
$\rho = -0.5$	-0.482 (0.186)	-0.530 (0.276)
$\sigma_\eta^2 = 0.01$	0.026 (0.009)	0.029 (0.012)

In comparing the simulation results for the two proposed models, three conclusions can be drawn. The first one is that omitting any of the components of the proposed SV models leads to biased estimation results. The second conclusion is that, based on the DIC_{obs} and the ML values, the SV model with only the moving average part yields a better model fit than the SV with only in-mean effect (in the first case) or only leverage effect (in the second case). The third conclusion is that data of the second type used, *i.e.* data from a MASVL construction, are better fitted (closer posterior means to the true values) by the corresponding models, compared to MASVM-type data. This means that the inclusion of the in-mean effect creates bigger estimation challenges than the leverage effect.

2.3 Comparison of our method for updating the volatilities with those of Shephard and Pitt (1997) and Omori and Watanabe (2008)

In this simulation study we examine the efficiency of our proposed MCMC methods, based on the Chan (2017) algorithm for the update of volatilities, to those based on the Shephard and Pitt (1997) and Omori and Watanabe (2008) algorithms. To this end, we repeat the simulation exercises that we considered in sections 2.1 and 2.2 of this Online Appendix. Efficiency is measured both in terms of correlations of the posterior draws and computational time. The algorithms are implemented using MATLAB R2017a on a desktop with Intel Core i5-3470 @ 3.20 GHz 3.20 GHz processes with 8 GB RAM.

Regarding the MASVM model, in order to update the latent volatilities, we use the method of hephard and Pitt (1997). In particular, we adopt a straightforward modification of the Abanto-Valle et al. (2012) approach, who used the block sampler of Shephard and Pitt (1997) in the context of a SV in mean model with heavy tails. The results are presented in Table 5.

Table 5: Simulation results for the MASVM model. Standard deviations in parentheses.

Method	Chan (2017)		Shephard and Pitt (1997)	
True value	Mean (stand. dev.)	IF	Mean (stand. dev.)	IF
$\mu = 0.2$	0.274 (0.163)	1.1872	0.271 (0.161)	1.375
$\mu_h = 0.2$	0.274 (0.181)	2.7196	0.279 (0.180)	4.892
$\lambda = 0.7$	0.651 (0.135)	15.018	0.649 (0.141)	35.311
$\psi = 0.4$	0.357 (0.035)	1.3249	0.323 (0.038)	1.8603
$\phi = 0.98$	0.960 (0.017)	30.517	0.959 (0.019)	49.857
$\sigma_\eta^2 = 0.01$	0.021 (0.006)	72.322	0.021 (0.014)	105.332
computational time	1928.18 seconds		3581.48 seconds	

The second column replicates the simulation results in Table 1, using the method of Chan (2017), whereas the fourth column displays the corresponding results when the method of Shephard and Pitt (1997) is used. To monitor the mixing of each algorithmic scheme, we report the Inefficienty Factor (IF) values (lower values correspond to better mixing).

Comparing the two IF columns, it is evident that the mixing of our proposed algorithm yields lower autocorrelations of the posterior draws than that based on Shephard and Pitt (1997). Furthermore, our algorithm, which exploits sparse matrices, speeds up the computations relative to the Shephard and Pitt (1997) method, which is based on Kalman filter techniques. The above conclusions continue to hold for the nested versions of the MASVM model.

For the case of the MASVL model, we compare our proposed algorithm with that based on Omori and Watanabe (2008) for the update of the volatilities. The results are presented in Table 6.

Table 6: Simulation results for the MASVL model. Standard deviations in parentheses.

Method	Chan (2017)		Omori and Watanabe (2008)	
True value	Mean (stand. dev.)	IF	Mean (stand. dev.)	IF
$\mu = 0.2$	0.172 (0.022)	2.656	0.179 (0.035)	3.598
$\mu_h = -2$	-1.973 (0.100)	1.120	-1.971 (0.134)	1.730
$\psi = 0.6$	0.620 (0.028)	2.410	0.619 (0.030)	2.472
$\phi = 0.94$	0.933 (0.031)	117.831	0.945 (0.056)	134.234
$\rho = -0.5$	-0.522 (0.167)	47.623	-0.549 (0.165)	54.293
$\sigma_\eta^2 = 0.01$	0.019 (0.009)	136.442	0.021 (0.042)	143.193
computational time	333.92 seconds		499.84 seconds	

Again, our MCMC algorithm for the MASVL model is faster and produces lowers autocorrelations in the posterior draws of the parameters. Similar results hold also for the nested versions of the MASVL model.

3 Additional empirical results

In this subsection we applied the MASVM model, with the joint prior (7) on (ϕ, λ), on all empirical applications. The results are presented in Table 7. As mentioned in the same paper, the usage of tis prior did not create significant changes in the posterior results.

Table 7: Empirical results for the MASVM model based on the joint prior on (ϕ, λ).

Dataset	Equity Hedge			S&P500			PHP/USD			Energy returns		
	Mean	IF	CD	Mean	IF	CD	Mean	IF	CD	Mean	IF	CD
μ	0.089*	7.461	-0.453	0.057*	1.811	1.722	-0.009	2.336	-0.664	0.314	3.100	0.035
	(0.013)			(0.029)			(0.017)			(0.291)		
λ	-0.472*	6.569	1.100	-0.072	1.489	-0.895	0.009	1.866	1.356	-0.008	2.197	-0.112
	(0.109)			(0.053)			(0.146)			(0.013)		
ψ	0.175*	2.449	-1.947	0.279*	2.183	2.022	0.142*	1.713	0.168	0.214*	1.614	-0.205
	(0.023)			(0.020)			(0.026)			(0.034)		
μ_h	-2.153*	1.223	0.296	-0.641*	1.062	-0.044	-2.216*	1.256	-0.321	3.012*	1.239	1.219
	(0.178)			(0.205)			(0.264)			(0.213)		
ϕ	0.971*	147.65	1.952	0.986*	34.579	-1.338	0.978*	40.743	2.905	0.959*	60.102	0.326
	(0.008)			(0.004)			(0.008)			(0.014)		
σ_η^2	0.039*	361.85	-1.941	0.014*	132.98	1.009	0.023*	108.96	-2.664	0.046*	132.14	-0.328
	(0.008)			(0.002)			(0.005)			(0.013)		

*Significant based on the 95% highest posterior density interval. Standard deviation in parentheses (for the estimated parameters). IF stands for Inefficiency Factor and CD stands for Convergence Diagnostics.

References

Abanto-Valle, C. A., Migon, H. S., 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* 141: 1875-1887.

Abanto-Valle, C. A., Migon, H. S., Lachos, V. H. (2012). Stochastic volatility in mean models with heavy-tailed distributions. *Brazilian Journal of Probability and Statistics* 26: 402-422.

Chan, J. C. C. (2013). Moving average stochastic volatility models with application to inflation forecast. *Journal of Econometrics* 176: 162-172.

Chan, J. C. C. (2017). The stochastic volatility in mean model with time-varying parameters: an application to inflation modeling. *Journal of Business & Economic Statistics* 35: 17-28.

Chan, J. C. C., Jeliazkov, I. (2009). Efficient simulation and integrated likelihood estimation in state space models. *International Journal of Mathematical Modelling and Numerical Optimisation* 1: 101-120.

Chib, S., Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm. *The American Statistician* 49: 327-335.

Nakajima, J., 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.

Omori, Y., Watanabe, T. (2008). Block sampler and posterior mode estimation for asymmetric stochastic volatility models. *Computational Statistics & Data Analysis* 52: 2892-2910.

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

Tierney, L. (1994). Markov Chains for exploring posterior distributions. *The Annals of Statistics* 22: 1701-1728.