APPLIED STOCHASTIC MODELS IN BUSINESS AND INDUSTRY
Appl. Stochastic Models Bus. Ind. 2010; **26**:172–193
Published online 8 June 2009 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/asmb.789

Bayesian modeling of financial returns: A relationship between volatility and trading volume

Carlos A. Abanto-Valle[1,*,†], Helio S. Migon[1] and Hedibert F. Lopes[2]

[1] *Instituto de Matemática, Universidade Federal do Rio de Janeiro, Brazil*
[2] *The University of Chicago Booth School of Business, U.S.A.*

SUMMARY

The modified mixture model with Markov switching volatility specification is introduced to analyze the relationship between stock return volatility and trading volume. We propose to construct an algorithm based on Markov chain Monte Carlo simulation methods to estimate all the parameters in the model using a Bayesian approach. The series of returns and trading volume of the British Petroleum stock will be analyzed. Copyright © 2009 John Wiley & Sons, Ltd.

Received 1 September 2008; Revised 6 April 2009; Accepted 6 April 2009

KEY WORDS: stochastic volatility; nonlinear and non-Gaussian state space models; Markov process of first order; Markov chain Monte Carlo

1. INTRODUCTION

The dynamics of the relationship between stock return volatility and trading volume has a long history in the finance literature. Karpoff [1] provides a good survey of this literature, discussing the return–volume relation in various financial markets. In considering this problem, Clark [2] started the discussion by presenting the intuitively appealing mixture of distributions hypothesis (MDH). According to the MDH, return and trading volume are driven by the same underlying latent information flow variable, i.e. price movements and the trading volume changes are caused primarily by the arrival of new information and the volatility process that incorporates this information into market prices. Although much of the empirical research documents a positive correlation between trading volume and return volatility, the evidence on whether the observed relation can be reconciled with the predictions of market microstructure theory is mixed (see, for example, [3–5]).

*Correspondence to: Carlos A. Abanto-Valle, Instituto de Matemática, Universidade Federal do Rio de Janeiro, Caixa Postal 68530, CEP: 21945-970, Rio de Janeiro-RJ, Brazil.
†E-mail: cabantovalle@im.ufrj.br

Contract/grant sponsor: FAPERJ; contract/grant number: E-26/171.092/2006

There are several variants of the MDH in the literature. These include the models of Clark [2], Tauchen and Pitts [3], Harris [6] and Andersen [7]. A first approach to merge the insights into the MDH with those of the market microstructure theory is the empirical model of daily return–volume relationship developed by Andersen [7]. He combines several important features of these models—for instance an asymmetric information structure and the presence of liquidity or noise traders—with the MDH and the related concept of stochastic volatility (SV). The resulting model, called the modified mixture model (MMM), is estimated with a dynamic AR(1) stochastic volatility process for the latent rate of information arrival, as proposed by Andersen [8], by using the generalized method of moments. Mahieu and Bauer [9] and Watanabe [10] implemented the MMM from a Bayesian viewpoint using simulation techniques based on Markov chain Monte Carlo (MCMC) methods to estimate the parameters and the latent process.

As pointed out by Lamoureux and Lastrapes [11] and Hamilton and Susmel [12], the high persistence in volatility may, however, be caused by a structural change in volatility. Some researchers such as So *et al.* [13] and Kalimipalli and Susumel [14] combine the SV model with the Markov switching (MS) model proposed by Hamilton [15] to accommodate the shift in the mean of log-volatility.

In this article we propose to expand the log-volatility specification used in Mahieu and Bauer [9] by introducing the MS volatility specification proposed by So *et al.* [13], which allows taking into consideration different volatility regimes.

The remainder of the article is organized as follows: Section 2 presents the relation between stock return volatility and trading volume with the extended specification for the log-volatility. Section 3 shows the Bayesian estimation procedure using MCMC methods. Section 4 shows an application using an artificial data set. Section 5 presents an empirical application on the return and trading volume series for the British Petroleum (BP) stock. Finally, Section 6 concludes.

2. THE MODEL

Andersen [7] develops an empirical return volatility–trading volume model using the theoretical framework of Glosten and Milgrom [16]. In his specification, the trading volume has two components that are directly related to informed and uninformed traders. The uninformed component is governed by a time invariant Poisson process with constant intensity m_0, while the informed volume has a Poisson distribution with parameter, which is a function of the information flow, that is, $m_1 e^{h_t}$. An empirical version of the MMM of Andersen [7], which was formulated by Mahieu and Bauer [9], leads to the following specification:

$$y_t \mid h_t \sim \mathcal{N}(0, e^{h_t}) \tag{1}$$

$$v_t \mid h_t \sim \mathcal{P}(m_0 + m_1 e^{h_t}), \quad m_0, m_1 > 0 \tag{2}$$

$$h_t = \alpha + \phi h_{t-1} + \eta_t, \quad \eta_t \sim \mathcal{N}(0, \sigma_\eta^2) \tag{3}$$

where y_t, v_t and h_t are, respectively, the compounded return, the trading volume and the log-volatility on day t. $\mathcal{N}(\cdot,\cdot)$ and $\mathcal{P}(\cdot)$ indicate the normal and Poisson distributions respectively. In Equation (2), m_0 reflects the uninformed component of trading volume and is related to liquidity traders. The remaining part of trading volume that is induced by new information is represented by $m_1 e^{h_t}$. The MMM defined by Equations (1)–(3) will be denoted as SV-VOL. Note that the

DOI: 10.1002/asmb

univariate SV model used extensively in the financial literature (see [17–20] among others) is specified by Equations (1) and (3).

In this article we modify the SV-VOL model by allowing the log-volatility specification to incorporate regime-switching properties; that is, the parameter determining the level of the log-volatility is allowed occasional discrete shifts:

$$h_t = \alpha_{S_t} + \phi h_{t-1} + \eta_t, \quad \eta_t \sim \mathcal{N}(0, \sigma_\eta^2) \tag{4}$$

In (4), the switching dynamic is governed by a k-state first-order Markov process with transition probabilities $p_{ij} = \mathrm{P}(S_t = j \mid S_{t-1} = i)$, where $i, j = 1, \ldots, k$ and $\sum_{j=1}^{k} p_{ij} = 1$. In this new specification S_t is the indicator variable showing the mean level of the state at time t, that is, the state indicator, S_t defines a particular regime for the parameter values. In this article, we consider only two regimes, $S_t = 0$ and $S_t = 1$, which indicate high and low volatility regimes, respectively. In order to avoid identifiability problems, we assume that $\alpha_0 = \gamma_0$ and $\alpha_1 = \gamma_0 + \gamma_1 S_t$ with the restriction $\gamma_1 < 0$ enforced for identification of each regime. The resulting MMM with MS defined by Equations (1), (2) and (4) will be denoted as MSSV-VOL.

It is well known in the finance literature that the asset returns follows a heavy tail distribution. We modify the normality specification of the returns in order to capture heavy tailed features in the marginal distribution of random errors using the mixture representation of the Student t distribution given by

$$y_t \mid h_t \sim \mathcal{N}(0, \lambda_t^{-1} e^{h_t}) \tag{5}$$

$$\lambda_t \sim \mathcal{G}\left(\frac{\nu}{2}, \frac{\nu}{2}\right) \tag{6}$$

where $\mathcal{G}(.,.)$ indicates the gamma distribution. Model defined by Equations (2), (4)–(6) will be denoted as MSSVt-VOL. Note that, (5) and (6) give us the SV model with Student t errors (SVt) [21–23]. Finally, Equations (4)–(6) define the MSSVt model.

3. MSSV-VOL MODEL ESTIMATION USING MCMC

A Bayesian approach to parameter estimation in the MSSV-VOL and MSSVt-VOL models relies on MCMC techniques. We propose to construct an algorithm based on MCMC simulation methods to make the Bayesian analysis of all the parameters feasible.

The MSSV model, defined by Equations (1) and (4), has been studied from a Bayesian viewpoint using MCMC methods. Multi-move samplers have been used to update the log-volatilities. For example, So *et al.* [13] use the mixture sampler Gibbs sampling [18, 24] the block sampler Gibbs sampling [25, 26].

Let $\boldsymbol{\theta} = (\gamma_0, \gamma_1, \phi, \sigma_\eta^2, m_0, m_1, p_{00}, p_{11}, \pi_0, \nu)'$ be the vector of parameters of the MSSVt-VOL model, $\mathbf{h}_{0:T} = (h_0, h_1, \ldots, h_T)'$ be the vector of the log-volatilities and $\mathbf{S}_{0:T} = (S_0, S_1, \ldots, S_T)'$ the states of the first-order Markov process, $\boldsymbol{\lambda}_{1:T} = (\lambda_1, \ldots, \lambda_T)'$ the mixing variables and $\pi_0 = P(S_0 = 0)$. Let $\mathbf{y}_{1:T} = (y_1, \ldots, y_T)'$ and $\mathbf{v}_{1:T} = (v_1, \ldots, v_T)'$ represent the information available up to time T.

The Bayesian approach for estimating the MSSVt-VOL model uses the data augmentation principle, which considers $\mathbf{h}_{0:T}$, $\boldsymbol{\lambda}_{1:T}$ and $\mathbf{S}_{0:T}$ as latent parameters. By using Bayes' theorem, the

DOI: 10.1002/asmb

joint posterior density of parameter and latent variables has the following decomposition:

$$p(\mathbf{h}_{0:T}, \boldsymbol{\lambda}_{1:T}, \mathbf{S}_{0:T}, \boldsymbol{\theta} \mid \mathbf{y}_{1:T}, \mathbf{v}_{1:T}) \propto p(\boldsymbol{\theta}) p(\mathbf{y}_{1:T} \mid \boldsymbol{\lambda}_{1:T}, \mathbf{h}_{0:T}) p(\mathbf{v}_{1:T} \mid \mathbf{h}_{0:T}, \boldsymbol{\theta}) \times p(\boldsymbol{\lambda}_{1:T} \mid \boldsymbol{\theta}) p(\mathbf{h}_{0:T} \mid \mathbf{S}_{0:T}, \boldsymbol{\theta}) p(\mathbf{S}_{0:T} \mid \boldsymbol{\theta}) \tag{7}$$

where

$$p(\mathbf{y}_{1:T} \mid \boldsymbol{\lambda}_{1:T}, \mathbf{h}_{0:T}) \propto \prod_{t=1}^{T} \lambda_t^{1/2} e^{-(h_t + \lambda_t y_t^2 e^{-h_t})/2} \tag{8}$$

$$p(\mathbf{v}_{1:T} \mid \mathbf{h}_{0:T}, \boldsymbol{\theta}) \propto \prod_{t=1}^{T} [m_0 + m_1 e^{h_t}]^{v_t} [e^{-m_0 - m_1 e^{h_t}}] \tag{9}$$

$$p(\boldsymbol{\lambda}_{1:T} \mid \boldsymbol{\theta}) \propto \prod_{t=1}^{T} \lambda_t^{(v/2)-1} e^{-v\lambda_t/2} \tag{10}$$

$$p(\mathbf{h}_{0:T} \mid \mathbf{S}_{0:T}, \boldsymbol{\theta}) \propto e^{-\frac{1-\phi^2}{2\sigma_\eta^2}\left(h_0 - \frac{\alpha_{S_0}}{1-\phi}\right)^2} \prod_{t=1}^{T} e^{-\frac{1}{2\sigma_\eta^2}(h_t - \alpha_{S_t} - \phi h_{t-1})^2} \tag{11}$$

$$p(\mathbf{S}_{0:T} \mid \boldsymbol{\theta}) \propto \pi_{S_0} \prod_{t=1}^{T} p_{S_{t-1} S_t} \tag{12}$$

$$\pi_i = P(S_0 = i), \quad i = 0, 1 \tag{13}$$

For the unknown parameters in the MSSVt-VOL model, the prior distributions are set as: $\phi \sim \mathcal{N}_{(-1,1)}(\bar{\phi}, \sigma_\phi^2)$, $\sigma_\eta^2 \sim \mathcal{IG}(T_0/2, M_0/2)$, $p_{ii} \sim \mathcal{Be}(\lambda_{i0}, \lambda_{i1}), i = 0, 1$, $\pi_0 \sim \mathcal{Be}(l_0, l_1)$, $\boldsymbol{\gamma} = (\gamma_0, \gamma_1)' \sim N_{2(\gamma_1 < 0)}(\bar{\boldsymbol{\gamma}}, \bar{\mathbf{B}})$, $m_0 \sim \mathcal{G}(a_0, b_0)$, $m_1 \sim \mathcal{G}(a_1, b_1)$, $v \sim \mathcal{G}(a_v, b_v)$, where as usual $\mathcal{N}_{(.,.)}(.,.)$, $\mathcal{IG}(.,.)$, $\mathcal{Be}(.,.)$, $\mathcal{N}_{p(.)}(.,.)$ represent the truncated normal, inverse gamma, beta and the p-multivariate truncated normal distribution, respectively.

Since the posterior density $p(\mathbf{h}_{0:T}, \boldsymbol{\lambda}_{1:T}, \mathbf{S}_{0:T}, \boldsymbol{\theta} \mid \mathbf{y}_{1:T}, \mathbf{v}_{1:T})$ does not have closed form, first we sample the parameters $\boldsymbol{\theta}$ and next the latent variables $\mathbf{S}_{0:T}$, $\boldsymbol{\lambda}_{1:T}$ and $\mathbf{h}_{0:T}$ using Gibbs sampling. The sampling scheme is described by Algorithm 3.1.

Algorithm 3.1

1. Set $i = 0$ and get starting values for the parameters $\boldsymbol{\theta}^{(i)}$, the states $\mathbf{S}_{0:T}^{(i)}$, $\boldsymbol{\lambda}_{1:T}^{(i)}$ and $\mathbf{h}_{0:T}^{(i)}$
2. Draw $\boldsymbol{\theta}^{(i+1)} \sim p(\boldsymbol{\theta} \mid \mathbf{h}_{0:T}^{(i)}, \boldsymbol{\lambda}_{1:T}^{(i)}, \mathbf{S}_{0:T}^{(i)}, \mathbf{y}_{1:T}, \mathbf{v}_{1:T})$
3. Draw $\mathbf{S}_{0:T}^{(i+1)} \sim p(\mathbf{S}_{0:T} \mid \boldsymbol{\theta}^{(i+1)}, \boldsymbol{\lambda}_{1:T}^{(i)}, \mathbf{h}_{0:T}^{(i)}, \mathbf{y}_{1:T}, \mathbf{v}_{1:T})$
4. Draw $\boldsymbol{\lambda}_{0:T}^{(i+1)} \sim p(\boldsymbol{\lambda}_{0:T} \mid \boldsymbol{\theta}^{(i+1)}, \mathbf{S}_{0:T}^{(i+1)}, \mathbf{h}_{0:T}^{(i)}, \mathbf{y}_{1:T}, \mathbf{v}_{1:T})$
5. Draw $\mathbf{h}_{0:T}^{(i+1)} \sim p(\mathbf{h}_{0:T} \mid \boldsymbol{\theta}^{(i+1)}, \mathbf{S}_{0:T}^{(i+1)}, \mathbf{y}_{1:T}, \mathbf{v}_{1:T})$
6. Set $i = i + 1$ and return to 2 until achieving convergence.

As described by Algorithm 3.1, the Gibbs sampler requires sampling parameters and latent variables from their full conditionals. Sampling the log-volatilities $\mathbf{h}_{0:T}$ in Step 5 is the most difficult task due to the nonlinear setup in Equations (1) and (2). In order to avoid the higher correlations due

DOI: 10.1002/asmb

to the Markovian structure of the h_t's, we develop a multi-move sampler [25–27] to sample the $\mathbf{h}_{0:T}$ by blocks. Details on the full conditionals of $\boldsymbol{\theta}$ and the latent variable $\mathbf{S}_{0:T}$ are given in the Appendix.

Let us consider the MSSVt-VOL model. In order to simulate $\mathbf{h}_{0:T}$, we break the problem into two steps: first, we simulate h_0 conditional on $\mathbf{h}_{1:T}$ and next $\mathbf{h}_{1:T}$ conditional on h_0. In our block sampler, we divide $\mathbf{h}_{1:T}$ into $K+1$ blocks, $\mathbf{h}_{k_{i-1}+1:k_i-1}=(h_{k_{i-1}+1},\ldots,h_{k_i-1})'$ for $i=1,\ldots,K+1$, with $k_0=0$ and $k_{K+1}=T$, where $k_i-1-k_{i-1}\geqslant 2$ is the size of the ith block. Following Shephard and Pitt [25] and Omori and Watanabe [27], the K knots $(k_1,\ldots,k_K)$ are generated randomly using

$$k_i=\text{int}[T\times\{(i+u_i)/(K+2)\}],\quad i=1,\ldots,K \tag{14}$$

where the u_i's are independent realizations of the uniform random variable on the interval (0,1) and int$[x]$ represents the floor of x. We control the single tuning parameter K to obtain the efficient sampler. If K is too large, the sampler will be slow because of rejections; if K is too small, it will be correlated because of the structure of the model.

We sample the ith block of disturbances $\boldsymbol{\eta}_{k_{i-1}+1:k_i-1}=(\eta_{k_{i-1}+1},\ldots,\eta_{k_i-1})'$ instead of $\mathbf{h}_{k_{i-1}+1:k_i-1}=(h_{k_{i-1}+1},\ldots,h_{k_i-1})'$, exploiting the fact that the innovations η_t are i.i.d. with $\mathcal{N}(0,\sigma^2_\eta)$. Suppose that the conditional knots of the ith block are $k_{i-1}=t$ and $k_i=t+k+1$, such that $t+k<T$. The full conditional distribution of $\boldsymbol{\eta}_{t+1:t+k}=(\eta_{t+1},\ldots,\eta_{t+k})'$ given h_t,h_{t+k+1} and the other parameters, $f(\boldsymbol{\eta}_{t+1:t+k}|h_t,h_{t+k+1},\boldsymbol{\theta},\mathbf{y}_{t+1:t+k},\mathbf{v}_{t+1:t+k},\boldsymbol{\lambda}_{t+1:t+k},\mathbf{S}_{t+1:t+k+1})$, which is expressed in the log scale as

$$\begin{aligned}&\log f(\boldsymbol{\eta}_{t+1:t+k}|h_t,h_{t+k+1},\boldsymbol{\theta},\mathbf{y}_{t+1:t+k},\mathbf{v}_{t+1:t+k},\boldsymbol{\lambda}_{t+1:t+k},\mathbf{S}_{t+1:t+k+1})\\&\quad=\text{const}-\frac{1}{2\sigma^2_\eta}\sum_{r=t+1}^{t+k}\eta_r^2+\sum_{r=t+1}^{t+k}\log f(y_r,v_r\,|\,\lambda_r,h_r)-\frac{1}{2\sigma^2_\eta}(h_{t+k+1}-\alpha_{S_{t+k+1}}-\phi h_{t+k})^2\end{aligned} \tag{15}$$

As (15) does not have closed form, we use the Metropolis–Hastings acceptance–rejection algorithm [28, 29] to sampling from. To obtain the proposal density, we are going to form an approximated linear state space model that mimics (15), from which sampling is easy. Then, for the MSSVt-VOL model, the log of the joint distribution of y_r and v_r is given by

$$\log p(y_r,v_r\,|\,\lambda_r,h_r)=\text{const}-\frac{h_r}{2}-\frac{\lambda_t}{2}y_r^2e^{-h_r}-(m_0+m_1e^{h_r})+v_r\log(m_0+m_1e^{h_r}) \tag{16}$$

We denote the log of $p(y_r,v_r\,|\,\lambda_r,h_r)$ by $l(h_r)$ and write the first and second derivatives of $l(h_r)$ with respect to h_r by l' and l'', respectively. Applying a second-order Taylor series expansion to $\sum_{r=t+1}^{t+k}l(h_r)$ in Equation (15) around some preliminary estimate of $\boldsymbol{\eta}_{t+1:t+k}$, denoted by $\hat{\boldsymbol{\eta}}_{t+1:t+k}$, we have

$$\begin{aligned}&\log f(\boldsymbol{\eta}_{t+1:t+k}|h_t,h_{t+k+1},\boldsymbol{\theta},\mathbf{y}_{t+1:t+k},\mathbf{v}_{t+1:t+k},\boldsymbol{\lambda}_{t+1:t+k},\mathbf{S}_{t+1:t+k+1})\\&\quad\approx\text{const}-\frac{1}{2\sigma^2_\eta}\sum_{r=t+1}^{t+k}\eta_r^2-\frac{1}{2\sigma^2_\eta}(h_{t+k+1}-\alpha_{S_{t+k+1}}-\phi h_{t+k})^2\\&\qquad+\sum_{r=t+1}^{t+k}\left\{l(\hat{h}_r)+(h_r-\hat{h}_r)l'(\hat{h}_r)+\frac{1}{2}(h_r-\hat{h}_r)^2l''_F(\hat{h}_r)\right\}\end{aligned} \tag{17}$$

where $\hat{\mathbf{h}}_{t+1:t+k}$ is the estimate of $\mathbf{h}_{t+1:t+k}$ corresponding to $\hat{\boldsymbol{\eta}}_{t+1:t+k}$. As $l(h_r)$ is not concave, we propose using $l''_F(h_r)$ in place of $l''(h_r)$, which can be positive for some values of h_r. To ensure that $l_F''(h_r)$ is everywhere strictly negative,[‡] it is defined as

$$l''_F(h_r) = \mathbb{E}[l''(h_r)] = -\frac{1}{2} - \frac{m_1^2 e^{2h_r}}{m_0 + m_1 e^{h_r}} \tag{18}$$

The expectation in (18) is taken with respect to the joint density of y_r and v_r conditional on h_r and λ_r.

After some simple but tedious algebra in (17), we have the approximating normal density g as follows:

$$\begin{aligned}
&\log f(\boldsymbol{\eta}_{t+1:t+k} | h_t, h_{t+k+1}, \boldsymbol{\theta}, \mathbf{y}_{t+1:t+k}, \mathbf{v}_{t+1:t+k}, \boldsymbol{\lambda}_{t+1:t+k}, \mathbf{S}_{t+1:t+k+1}) \\
&= \text{const} - \frac{1}{2\sigma_\eta^2} \sum_{r=t+1}^{t+k} \eta_r^2 + \frac{1}{2} \sum_{r=t+1}^{t+k-1} l''_F(\hat{h}_r) \left(\hat{h}_r - \frac{l'(\hat{h}_r)}{l''_F(\hat{h}_r)} - h_r \right)^2 - \frac{\phi^2 - l''_F(\hat{h}_{t+k})\sigma_\eta^2}{2\sigma_\eta^2} \\
&\times \left\{ \frac{\sigma_\eta^2}{\phi^2 - l''_F(\hat{h}_{t+k})} \left(l'(\hat{h}_{t+k}) - l''_F(\hat{h}_{t+k})\hat{h}_{t+k} + \frac{\phi - \alpha_{S_{t+k+1}}}{\sigma_\eta^2} h_{t+k+1} \right) - h_{t+k} \right\}^2 = \log g
\end{aligned} \tag{19}$$

From (19), we define auxiliary variables d_r and $\hat{y}_r$ for $r = t+1, \ldots, t+k-1$ as follows:

$$\begin{aligned}
d_r &= -\frac{1}{l''_F(\hat{h}_r)} \\
\hat{y}_r &= \hat{h}_r + d_r l'(\hat{h}_r)
\end{aligned} \tag{20}$$

For $r = t+k < T$,

$$d_r = \frac{\sigma_\eta^2}{\phi - \sigma_\eta^2 l''_F(\hat{h}_{t+k})}$$

and

$$\hat{y}_r = d_r \left[l'(\hat{h}_r) - l''_F(\hat{h}_r)\hat{h}_r + \frac{(\phi - \alpha_{s_{r+1}})}{\sigma_\eta^2} h_{r+1} \right] \tag{21}$$

When $r = t+k = T$, we use (20) to define the auxiliary variables.

The resulting normalized density in (19), g, is a k-dimensional normal density, which is the exact density of $\boldsymbol{\eta}_{t+1:t+k}$ conditional on $\hat{\mathbf{y}}_{t+1:t+k}$ in the linear Gaussian state space model:

$$\hat{y}_r = h_r + \varepsilon_r, \quad \varepsilon_r \sim N(0, d_r) \tag{22}$$

$$h_r = \alpha_{s_r} + \phi h_{r-1} + \eta_r, \quad \eta_r \sim N(0, \sigma_\eta^2) \tag{23}$$

---

‡In the context of the SV-VOL model, Watanabe [10] uses an alternative expression, $\min\{l''(h_r), -0.001\}$, to ensure the strictly negative condition on $l_F''(h_r)$.

DOI: 10.1002/asmb

Applying the de Jong and Shepard [30] simulation smoother to this model with the artificial $\hat{\mathbf{y}}_{t+1:t+k}$ enables us to sample $\boldsymbol{\eta}_{t+1:t+k}$ from the density g. Since g does not bound f, we use the Metropolis–Hastings acceptance–rejection algorithm [28, 29] to sample from f. In the MSSV-VOL case, we use the same procedure with $\lambda_t = 1$ for $t = 1, \ldots, T$.

We select the expansion block $\hat{\mathbf{h}}_{t+1:t+k}$ as follows. Once an initial expansion block $\hat{\mathbf{h}}_{t+1:t+k}$ is selected, we can calculate the artificial $\hat{\mathbf{y}}_{t+1:t+k}$. In the MCMC implementation, the previous sample of $\mathbf{h}_{t+1:t+k}$ may be taken as an initial value of the $\hat{\mathbf{h}}_{t+1:t+k}$. Then, applying the Kalman filter and a disturbance smoother to the linear Gaussian state space model consisting of Equations (22) and (23) with the artificial $\hat{\mathbf{h}}_{t+1:t+k}$ yields the mean of $\hat{\mathbf{h}}_{t+1:t+k}$ conditional on $\hat{\mathbf{y}}_{t+1:t+k}$ 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 to use only five iterations of this procedure to provide reasonably good sequence of $\hat{\mathbf{h}}_{t+1:t+k}$ instead of an optimal one. The procedure is summarized in Algorithm 3.2.

Algorithm 3.2

1. Initialize $\hat{\mathbf{h}}_{t+1:t+k}$.
2. Evaluate recursively $l'(\hat{h}_r)$ and $l''_F(\hat{h}_r)$ for $r = t+1, \ldots, t+k$.
3. Conditional on the current values of γ_0, γ_1, ϕ and σ_η^2, $\boldsymbol{\lambda}_{t+1:t+k}$, $\mathbf{S}_{t+1:t+k+1}$ define the auxiliary variables $\hat{y}_r$ and d_r using Equations (20) or (21) for $r = t+1, \ldots, t+k$.
4. Consider the linear Gaussian state space model in (22) and (23). Apply the Kalman Filter and a disturbance smoother [31] and obtain the posterior mean of $\boldsymbol{\eta}_{t:t+k}$ ($\mathbf{h}_{t:t+k}$) and set $\hat{\boldsymbol{\eta}}_{t:t+k}$ ($\hat{\mathbf{h}}_{t:t+k}$) to this value.
5. Return to step 2 and repeat the procedure until achieve convergence.

4. NUMERICAL ILLUSTRATION WITH ARTIFICIAL DATA SET

In order to asses the performance of the MCMC algorithms described in the previous section, we present results based on a simulated data set. All the calculations were performed running stand alone code developed by the authors using the Scythe statistical library§ [32]. We simulated a data set of 1500 observations using $\gamma_0 = -0.5$, $\gamma_1 = -0.25$, $\phi = 0.7$, $\sigma_\eta^2 = 0.2$, $m_0 = 0.85$, $m_1 = 0.15$, $p_{00} = 0.98$ and $p_{11} = 0.98$, which correspond to typical values found in daily series of returns and trading volume.

We set the prior distributions as: $\phi \sim \mathcal{N}_{(-1,1)}(0.95, 10)$, $\sigma_\eta^2 \sim \mathcal{IG}(\frac{5}{2}, \frac{0.1}{2})$, $p_{00} \sim \mathcal{B}e(50, 1.5)$, $p_{11} \sim \mathcal{B}e(1.5, 50)$, $\pi_0 \sim \mathcal{B}e(0.5, 0.5)$, $\boldsymbol{\gamma} \sim N_{2(\gamma_1<0)}(\bar{\boldsymbol{\gamma}}, \bar{\mathbf{B}})$, $m_0 \sim \mathcal{G}(0.08, 0.1)$, $m_1 \sim \mathcal{G}(1, 10)$, where $\bar{\boldsymbol{\gamma}} = (-0.5, -0.25)'$ and $\bar{\mathbf{B}} = \text{diag}(4.0, 1.0)$. The initial values of the parameters are randomly generated from the prior distributions. We set all the log-volatilities, h_t, to zero. Finally, the initial $\mathbf{S}_{0:T}$ are generated conditional on the initial values of $\boldsymbol{\theta}$ and $\mathbf{h}_{0:T}$.

§The Scythe statistical library is available for free at the web site http://scythe.wustl.edu. It is an open source C++ library for statistical computation. It includes a suite of matrix manipulation functions, a suite of random number generators and a suite of numerical optimizers.

Table I. Simulated data set: posterior mean, standard error of the posterior mean, 95% interval and convergence diagnostic(CD) for the MSSV-VOL.

Parameter	True	Mean	MC error	95% interval	CD
γ_0	−0.50	−0.5367	0.0105	(−0.9160, −0.2868)	−0.16
γ_1	−0.25	−0.2401	0.0069	(−0.5492, −0.0201)	−0.16
ϕ	0.70	0.6861	0.0055	(0.4450, 0.8202)	−0.72
σ_η^2	0.20	0.2677	0.0033	(0.1728, 0.3965)	−0.22
p_{00}	0.98	0.9698	0.0004	(0.9175, 0.9966)	−0.03
p_{11}	0.98	0.9717	0.0007	(0.9224, 0.9954)	−1.32
π_0	0.50	0.4925	0.0024	(0.0588, 0.9344)	−0.15
m_0	0.85	0.8332	0.0003	(0.7779, 0.8855)	−0.54
m_1	0.15	0.0880	0.0017	(0.0023, 0.3105)	0.98

The number of blocks, K, in the block sampler was set equal to 60, so that each block contained 25 h_t's on average. We conducted the MCMC simulation for 40 000 iterations. The first 10 000 draws were discarded as a burn-in period, and then the next 30 000 were recorded. Table I reports the posterior means, the Monte Carlo (MC) error of the posterior means, the 95% intervals and the convergence diagnostic (CD) statistics proposed by Geweke [33] for all the parameters. The 95% credibility intervals are estimated using the 2.5th and the 97.5th percentiles of the posterior samples.

The proposed algorithm is evaluated in terms of how well it estimates the true parameter values. It can be seen that the estimated results for the parameters appear quite reasonable. In Figure 1 all the 95% credibility intervals include true values. All parameters passed the CD test of Geweke [33] and also Heidelbelger and Welch [34], although the last one is not reported.

In Figure 2, the smoothed mean (dotted line) and the posterior 95% credibility interval (solid line) of h_t calculated from MCMC output are shown. They are compared with the true log-volatilities (points) showing that the estimated values follow the behavior of the true volatilities.

Figure 3 shows the probability that S_t are in the high volatility period (top), and the true values of state indicator variable S_t versus the estimated values using the MSSV-VOL model (bottom). In the majority of cases, the state indicators are well estimated.[¶]

5. EMPIRICAL APPLICATION

This section analyzes the daily closing prices and trading volume corrected by dividends and stock splits for the BP Company stock series listed on the London Stock Exchange (LSE).[‖] The analyzed period starts January 5, 1999, and ends July 17, 2008, yielding 2398 observations. Throughout we work with the mean corrected returns computed as

$$y_t = 100\left\{(\log P_t - \log P_{t-1}) - \frac{1}{T}\sum_{j=1}^{T}(\log P_j - \log P_{j-1})\right\}$$

[¶]The state indicators and $P(S_t = 0)$ obtained with the MSSV model give similar results, but they are not reported.
[‖]The data set was obtained from the Yahoo finance web site at http://finance.yahoo.com.

DOI: 10.1002/asmb

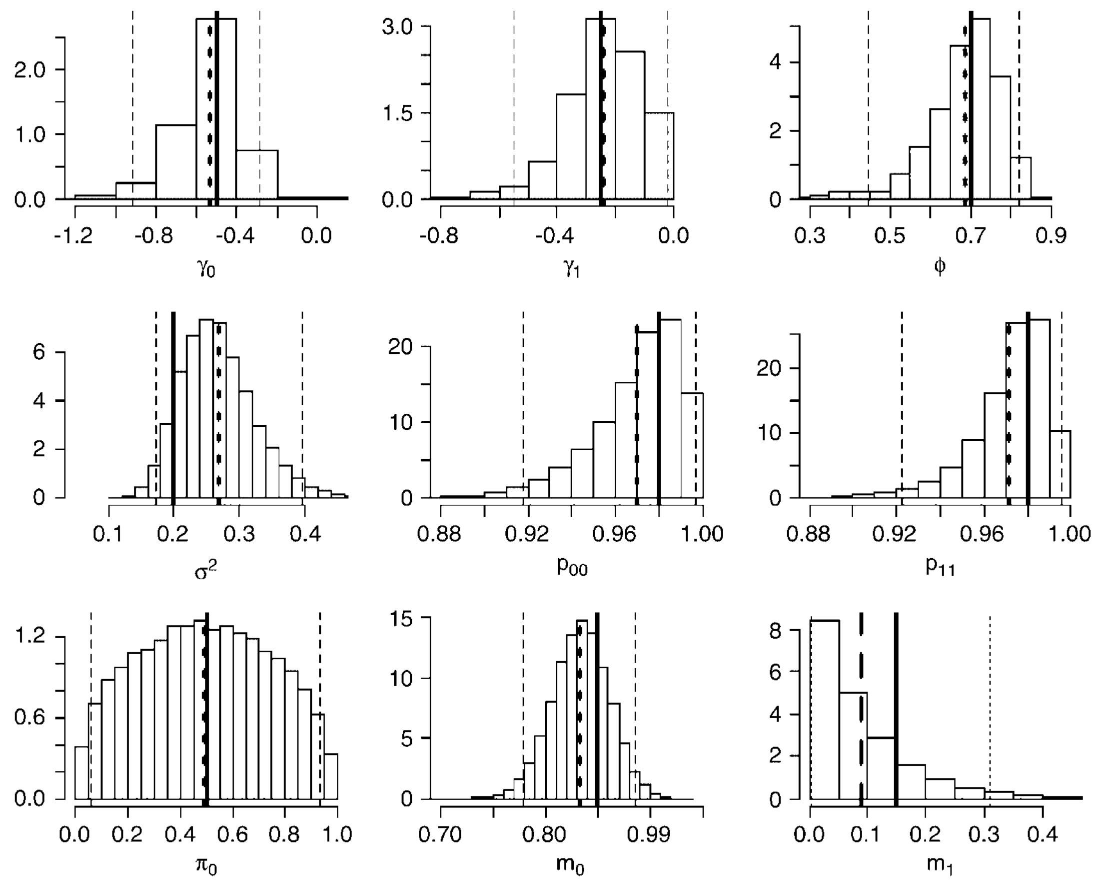

Figure 1. Simulated data: histograms of the parameters from MCMC output for the MSSV-VOL model. The tiny dotted and the dotted lines indicate the 2.5 and 97.5th percentiles and posterior mean, respectively. The solid line indicates the true value.

where P_t is the closing price on day t. To make the volume series stationary, the volume data are adjusted by regressing the log of the trading volume on a constant and on time $t=1,2,\ldots,T$. The exponential function of the residuals of this regression is then linearly transformed so that the raw data and the detrended data have the same mean and variance. For the following results, the detrended series is multiplied by 10^{-6}.

Table II summarizes descriptive statistics for the corrected compounded return and the detrended trading volume; the time series plots are shown in Figure 4. For the return series, the basic statistics, the mean, standard deviation, skewness and kurtosis are given as 0.00, 1.57, −0.12 and 4.98, respectively. Note that the kurtosis of the returns is above three, so that daily BP stock returns are not likely to follow a normal distribution.

We fitted the SV, SVt, MSSV, MSSVt, SV-VOL, SVt-VOL, MSSV-VOL and the MSSVt-VOL models. In all cases, we simulated the h_t's in a multi-move manner with stochastic knots based on the method described in Section 3. For the SV, SVt, SV-VOL and SVt-VOL models we set

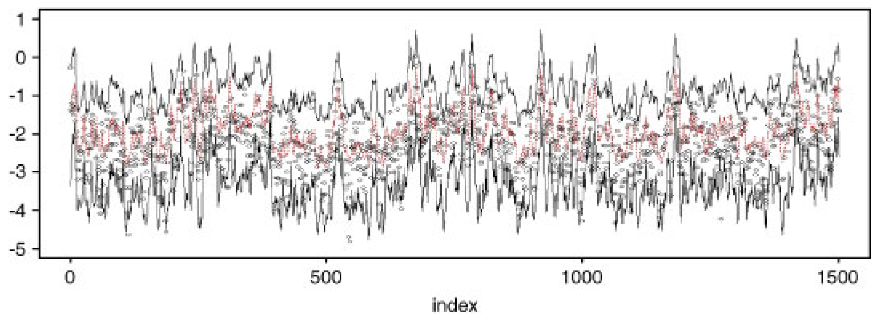

Figure 2. MSSV-VOL model, simulated data. True log-volatilities (dotted line), estimated posterior smoothed mean of h_t (points), 2.5 and 97.5th percentiles (solid lines).

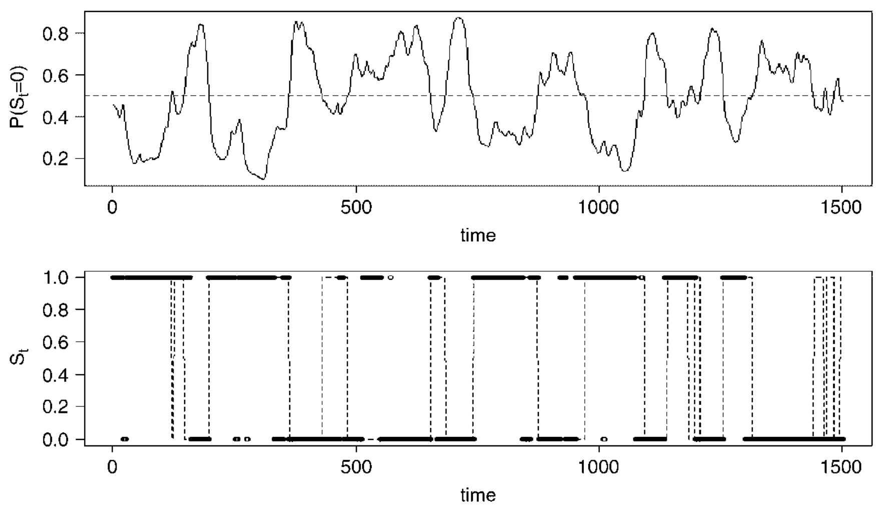

Figure 3. Simulated data set. Top: $P(S_t=0)$. Bottom: True states S_t(points) vs estimated S_t by the MSSV-VOL model (dotted line).

Table II. Summary statistics for BP stock series.

	Mean	S.D.	Max	Min	Skewness	Kurtosis
Returns	0.00	1.57	8.33	−7.92	−0.12	4.98
Volume	3.33	1.43	17.41	0.63	2.17	12.77

the prior distributions as: $\alpha \sim \mathcal{N}(0, 10)$, $\phi \sim \mathcal{N}_{(-1,1)}(0.95, 10)$, $\sigma_\eta^2 \sim \mathcal{IG}(\frac{5}{2}, \frac{0.05}{2})$, $v \sim \mathcal{G}(12.0, 0.8)$, $m_0 \sim \mathcal{G}(0.08, 0.1)$ and $m_1 \sim \mathcal{G}(1, 10)$. For the other models, we adopt the same priors as in Section 4. We set K, the number of blocks, as 60, in a such way that each block contained 40 h_t's on average.

The initial values on the MCMC simulation of the parameters are randomly generated from the prior distributions. We set all the log-volatilities, h_t, to zero, and the λ_t are sampled from the prior $\mathcal{G}(\frac{v}{2}, \frac{v}{2})$. Finally, the initial $\mathbf{S}_{0:T}$ are generated conditional on the initial values of $\boldsymbol{\theta}$ and $\mathbf{h}_{0:T}$.

DOI: 10.1002/asmb

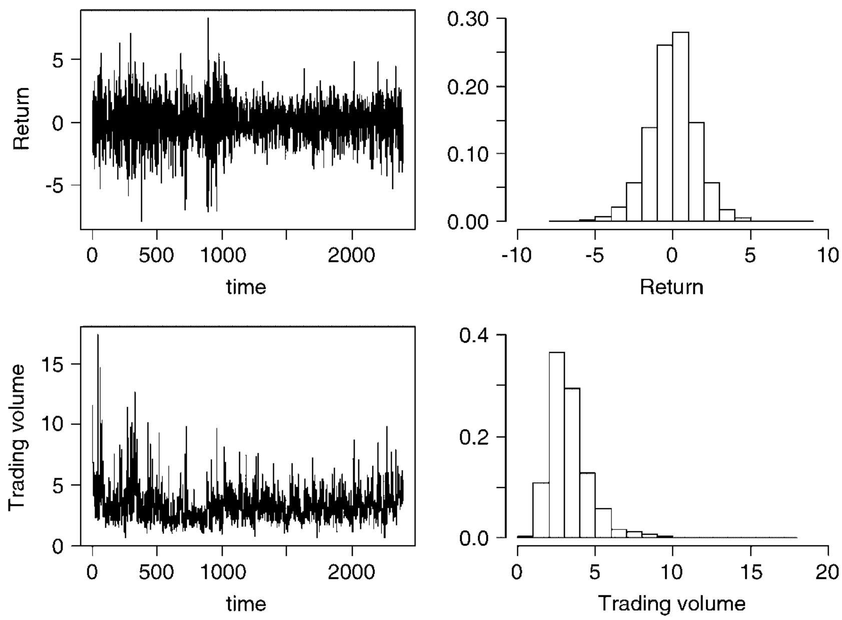

Figure 4. BP data with sample period from January 5, 1999 to July 17, 2007. Top: raw series and histogram of corrected returns. Bottom: raw series and histogram of trading volume.

Next, we run a chain by 60 000 iterations. The first 10 000 draws were discarded as a burn-in period, and then the next 50 000 were recorded. Using these 50 000 draws we calculated the posterior means, the MC error of the posterior means, the 95% intervals and the CD. Tables III and IV summarize the results. According to the CD values, the null hypothesis that the sequence of 50 000 draws is stationary is accepted at the 5% level for all the parameters in all the models considered here.

Table III reports the estimation results of SV model. The posterior mean and 95% interval of ϕ are 0.9372 and (0.9044, 0.9629), respectively, exhibiting high persistence in return volatility. Similar results were found in the SVt model. From Table III, we have that the persistence parameter ϕ dropping from 0.9372 to 0.7506 (MSSV model) and 0.7571 (MSSVt model), which agrees with previous results in MSSV models [13, 24, 35]. On the other hand, the posterior mean and the 95% interval of σ_η^2 are 0.1456 and (0.1010, 0.2080), which are higher than the 0.0588 and (0.0353, 0.0940) in the SV model and the 0.0641 and (0.0388, 0.0971) in the SVt model.

In the SV-VOL and SVt-VOL models, we find that posterior mean and 95% interval of ϕ are 0.9281 and (0.9027, 0.9503), 0.9346 and (0.9112, 0.9551), respectively, which are slightly different than those in the SV and SVt models. As mentioned earlier, m_0 reflects the noisy component of trading volume generated by liquidity traders. The remaining part of trading volume that is induced by new information is represented by $m_1 e^{h_t}$. We find that the posterior mean of m_0 is 2.5924 and the distribution of m_1 has a posterior mean of 0.2655, in the SV-VOL model. In the SVt-VOL, the values are quite similar 2.5792 and 0.3018. In the heavier-tailed SVt and SVt-VOL models, the magnitude of the tail-fatness is measured by the degrees of freedom v. We found that the posterior mean of v and the posterior credibility interval of 95% are 20.3200 and (12.5986, 29.2941), 16.4514 and (10.6345, 24.3830), respectively.

DOI: 10.1002/asmb

Table III. BP data results. First row: posterior mean. Second row: posterior 95% credible interval. Third row: Monte Carlo error of the posterior mean. Fourth row: CD statistics.

Parameter	SV	SVt	SV-VOL	SVt-VOL
	0.0429	0.0404	0.0484	0.0375
α	(0.0233, 0.0674)	(0.0214, 0.0632)	(0.0295, 0.0696)	(0.0211, 0.0561)
	0.5×10^{-3}	0.5×10^{-3}	0.2×10^{-3}	0.2×10^{-3}
	0.011	−1.056	−0.601	−1.272
	0.9372	0.9322	0.9281	0.9346
ϕ	(0.9044, 0.9629)	(0.8987, 0.9596)	(0.9027, 0.9503)	(0.9112, 0.9551)
	0.7×10^{-3}	0.7×10^{-3}	0.3×10^{-3}	0.3×10^{-3}
	0.060	1.196	0.728	1.158
	0.0588	0.0641	0.0934	0.0834
σ^2	(0.0353, 0.0940)	(0.0388, 0.0971)	(0.0672, 0.1254)	(0.0599, 0.1129)
	0.9×10^{-3}	0.9×10^{-3}	0.5×10^{-3}	0.6×10^{-3}
	−0.336	−1.148	−1.142	−1.279
	—	—	2.5924	2.5792
m_0	—	—	(2.4442, 2.7320)	(2.4239, 2.7279)
	—	—	1.1×10^{-3}	1.5×10^{-3}
	—	—	−1.122	−1.782
	—	—	0.2655	0.3018
m_1	—	—	(0.2120, 0.3255)	(0.2381, 0.3735)
	—	—	0.5×10^{-3}	0.9×10^{-3}
	—	—	1.037	1.564
	—	20.3200	—	16.4514
v	—	(12.5986, 29.2941)	—	(10.6345, 24.3830)
	—	0.2083	—	0.4535
	—	1.154	—	−1.063

We now consider the estimation based on the MSSV-VOL and MSSVt-VOL models. In Figures 5–7, we show the sample paths, the autocorrelation function and the histograms for parameters, respectively, only for the SV-VOL model. Tables III and IV show that the persistence parameter, ϕ, drops from 0.9372 (SV) to 0.8052 (MSSV-VOL). This value, 0.8052, is slightly greater than the posterior mean of the MSSV model (0.7506). The 95% posterior interval of ϕ is different than that obtained from the MSSV model. We found similar results in the heavier-tailed version of the models. The posterior means of m_0 and m_1 are 2.5422 and 0.2952. They are similar to the values in the SV-VOL model and also the posterior 95% intervals (see Tables III and IV). Note that in both cases the heavier-tailed version of the model with volume has a degree of freedom smaller than the version without it, the same is true when the heavier-tailed MS version of the models are compared. When we compare models without and with heavier-tailed MS, the MS versions present a posterior mean greater than without MS. That is, part of tail effect has been captured by the MS volatility.

The posterior means of the transition probabilities p_{00} and p_{11} are 0.9898 and 0.9949, respectively, which are slightly different than these the values 0.9928 and 0.9874 from the MSSV model, indicating that the probability of switching between high- and low-volatility states is quite low.

Table IV. BP data results. First row: posterior mean. Second row: posterior 95% credible interval. Third row: Monte Carlo error of the posterior mean. Fourth row: CD statistics.

Parameter	MSSV	MSSVt	MSSV-VOL	MSSVt-VOL
	0.3026	0.2807	0.2576	0.2295
γ_0	(0.1848, 0.4972)	(0.1623, 0.4495)	(0.1713, 0.3660)	(0.1469, 0.3330)
	4.7×10^{-3}	3.6×10^{-3}	1.8×10^{-3}	1.7×10^{-3}
	0.160	−1.192	0.218	−0.829
	−0.2456	−0.2391	−0.1970	−0.1840
γ_1	(−0.4094, −0.1414)	(−0.3905, −0.1321)	(−0.2907, −0.1202)	(−0.2757, −0.1097)
	4.0×10^{-3}	3.1×10^{-3}	1.5×10^{-3}	1.4×10^{-3}
	−0.083	1.174	−0.058	0.718
	0.7506	0.7571	0.8052	0.8176
ϕ	(0.6133, 0.8349)	(0.6310, 0.8446)	(0.7386, 0.8609)	(0.7548, 0.8717)
	3.5×10^{-3}	3.5×10^{-3}	1.3×10^{-3}	1.2×10^{-3}
	0.060	1.058	−0.152	0.845
	0.1456	0.1337	0.1465	0.1332
σ^2	(0.1010, 0.2080)	(0.0925, 0.1906)	(0.1070, 0.1956)	(0.0968, 0.1772)
	1.4×10^{-3}	1.2×10^{-3}	1.1×10^{-3}	1.0×10^{-3}
	0.522	−0.699	0.710	−0.937
	—	—	2.5422	2.5259
m_0	—	—	(2.3992, 2.6791)	(2.3760, 2.6690)
	—	—	1.3×10^{-3}	1.7×10^{-3}
	—	—	0.372	−0.545
	—	—	0.2952	0.3272
m_1	—	—	(0.2387, 0.3592)	(0.2607, 0.4036)
	—	—	0.7×10^{-3}	1.1×10^{-3}
	—	—	−0.847	0.441
	—	28.7098	—	21.6961
v	—	(15.8200, 39.3200)	—	(14.0000, 31.8600)
	—	0.422	—	0.274
	—	−1.732	—	−1.496
	0.9928	0.9927	0.9898	0.9894
p_{00}	(0.9819, 0.9987)	(0.9811, 0.9987)	(0.9724, 0.9977)	(0.9699, 0.9977)
	0.1×10^{-3}	0.1×10^{-3}	0.1×10^{-3}	0.2×10^{-3}
	−0.074	0.367	0.231	−0.149
	0.9874	0.9946	0.9949	0.9950
p_{11}	(0.9875.9987)	(0.9875, 0.9988)	(0.9877, 0.9987)	(0.9873, 0.9988)
	0.5×10^{-3}	0.5×10^{-3}	0.4×10^{-3}	0.6×10^{-3}
	0.571	0.456	0.338	−1.142
	0.6013	0.6046	0.6241	0.6257
π_0	(0.1331, 0.9588)	(0.1363, 0.9584)	(0.1727, 0.9615)	(0.1744, 0.9613)
	1.1×10^{-3}	1.3×10^{-3}	1.1×10^{-3}	1.0×10^{-3}
	−1.152	0.338	0.514	1.004

DOI: 10.1002/asmb

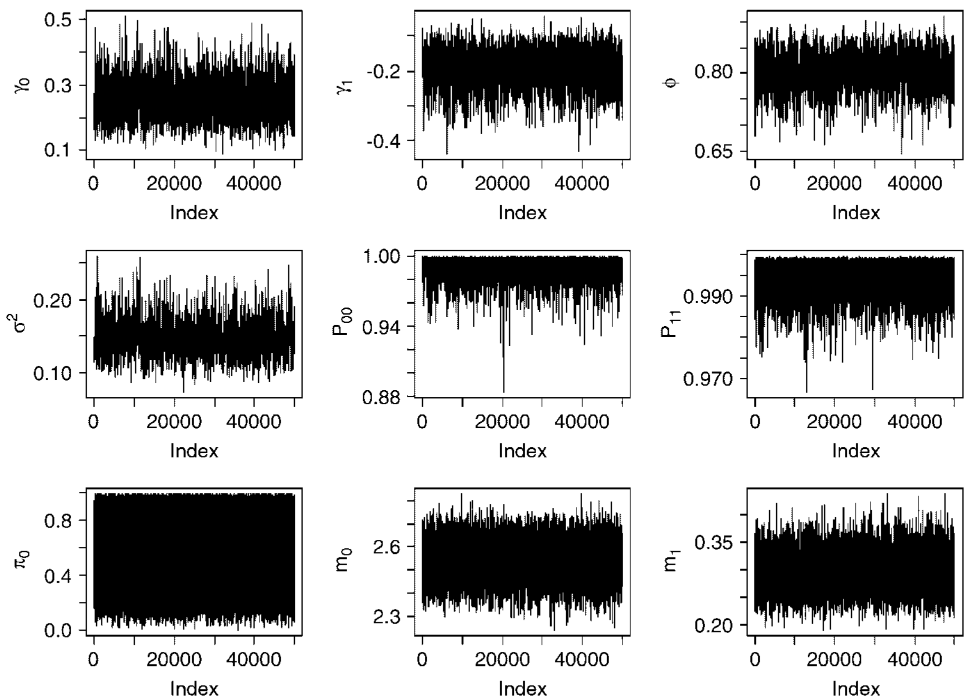

Figure 5. MSSV-VOL model, BP data: sample paths for parameters obtained from MCMC output.

In the MSSV-VOL model, a volatility shock lasts about 98 days in the high-volatility state compared with about 196 days in the low-volatility state. These values are 94 and 200 days in the MSSVt-VOL model. In the MSSV model, a volatility shock lasts about 138 and 79 days, respectively. In the MSSVt, the duration of the shock are 136 and 185 days, respectively. The expected duration of the shock in state i is obtained as $(1-p_{ii})^{-1}$.

Figure 8 depicts the posterior probabilities of the high-volatility state as inferred from the MSSV-VOL model (dotted line) and the MSSVt-VOL model (points). Following Hamilton [36], we consider an observation as belonging to a high-volatility state if the smoothed probability is higher than 0.5. Then, the high-volatilities periods are: 01/06/1999–05/17/1999,09/03/1999–01/25/2001, 10/04/2007–11/16/2007 and 12/27/2007–07/17/2008. We obtain the same results using the MSSVt-VOL model. The last two regimes of high volatility are explained by a series of events that caused the oil price to exceed \$92/barrel by October 2007, and \$99.29/barrel for December futures in New York on November 21, 2007. Throughout the first half of 2008, oil regularly reached record high prices. On February 29, 2008, oil prices peaked at \$103.05 per barrel, and reached \$110.20 on March 12, 2008, the sixth record in seven trading days. Prices on June 27, 2008, touched \$141.71/barrel, for August delivery in the New York Mercantile Exchange. The most recent price per barrel maximum of \$147.02 was reached on July 11, 2008. Figure 8 (solid line) shows the posterior probabilities of high-volatility state from the MSSV model. There is a significant difference between the period from 08/29/2001–12/03/2001, in which occurred the September 11 attacks. The trading volume gives us information to modify the regimes in the

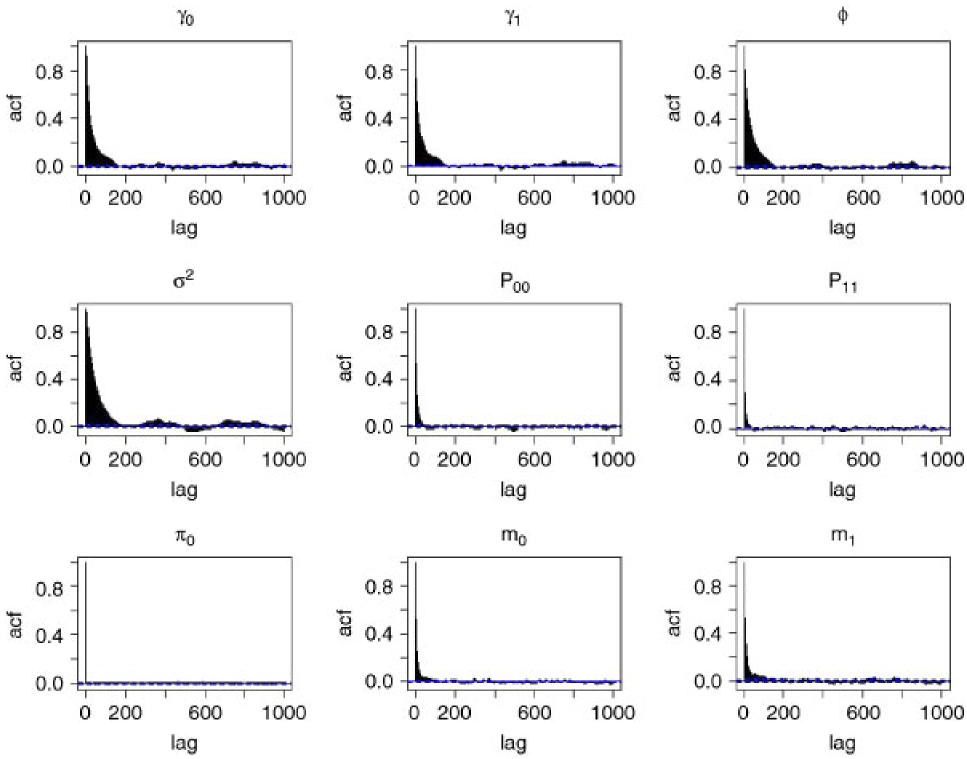

Figure 6. MSSV-VOL model, BP data: autocorrelation functions for parameters obtained from MCMC output.

MSSV model. We found similar results for other stocks such as IBM, Coca Cola and Kodak, although the results are not reported here.

To assess the goodness of the estimated models, we calculate the deviance information criterion, DIC (for more details about the DIC criterion see for example [37–39]). In this context, p_D is a measure of model complexity. As the quantity of information available is different between models with and without trading volume, we compare the SV, SVt, MSSV and MSSVt models and the SV-VOL, SVt-VOL, MSSV-VOL and MSSVt-VOL models. From Table V, according to the DIC the MSSV model best fits the return series of the BP stock. Model comparison for joint modeling of return and trading volume series according to the DIC criterion give us the MSSV as the best model to fit the BP data.

In Figure 9, we show the smoothed mean of e^{h_t} obtained from the MCMC output for MSSV-VOL and SV-VOL models (the best and worst to fit the BP data set). From a practical point of view, we are mainly interested in whether we find a significant difference between the two series. Therefore, we plot the difference between the smoothed means of the two series in the lower panel of Figure 9. This graph shows us that these series do not differ very much in most periods, but in some periods of high volatility, we observe difference in squared percentages of more than 4%, this is confirmed by the 95% posterior credibility intervals. This can have a substantial impact, for

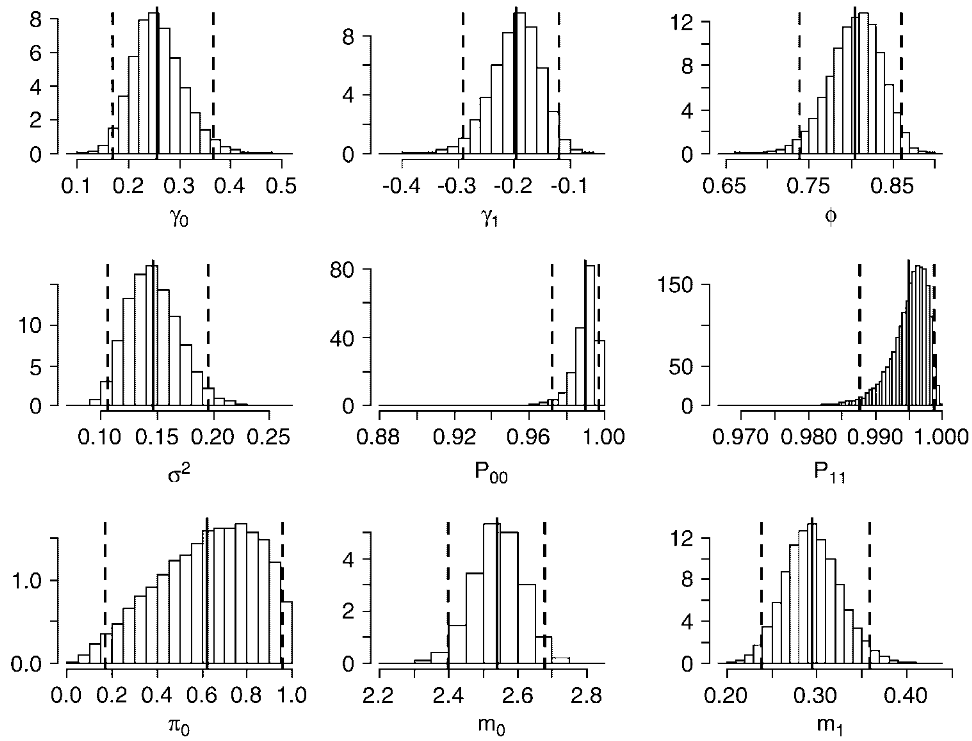

Figure 7. MSSV-VOL model, BP data: histograms for parameters obtained from MCMC output. The dotted lines indicate the 2.5 and 97.5*th* percentiles, respectively, and the solid line the posterior mean.

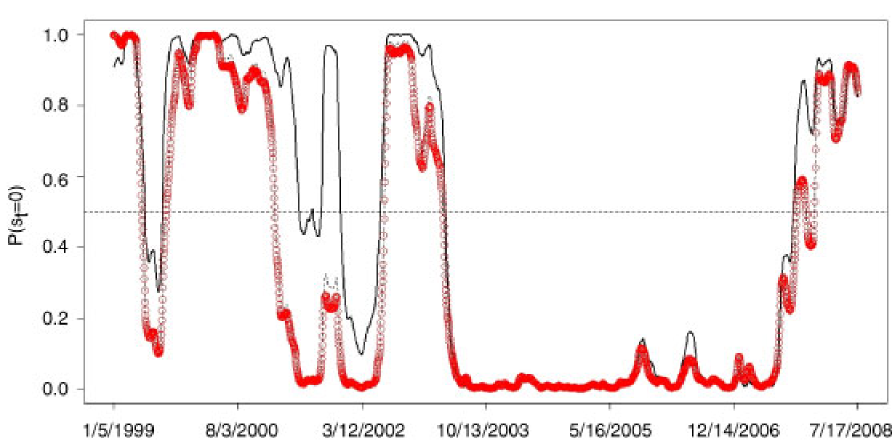

Figure 8. BP data: posterior probability of high-volatility state $P(S_t=0)$. MSSV model (solid line), MSSV-VOL model (dotted line), MSSVt-VOL model (points).

DOI: 10.1002/asmb

Table V. BP data. DIC: deviance information criterion, p_D: effective number of parameters.

Model	DIC	p_D	Rank
SV	8621.1	205.8	2
SVt	8621.8	280.5	3
MSSV	8615.9	249.9	1
MSSVt	8622.4	299.7	4
SV-VOL	16903.6	333.5	4
SVt-VOL	16897.3	430.4	3
MSSV-VOL	16847.1	364.9	1
MSSVt-VOL	16857.9	438.6	2

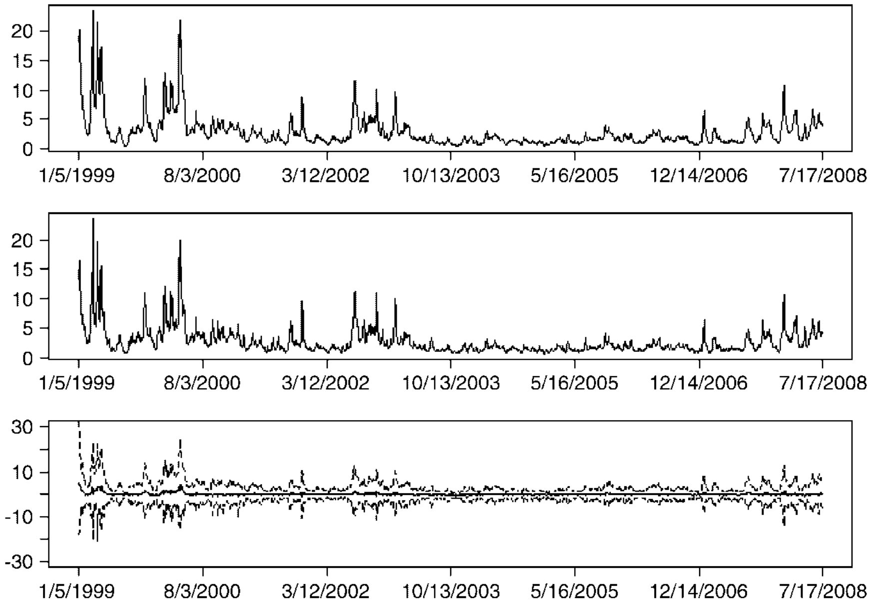

Figure 9. BP data. Top: smoothed mean of e^{h_t} from SV-VOL. Middle: smoothed mean of e^{h_t} from MSSV-VOL. Bottom: difference from both models. The dotted line indicates the 95% posterior credibility interval.

instance, in the valuation of derivative instruments and several strategic or tactical asset allocation topics.

6. CONCLUSIONS

This article studies the joint distribution of daily returns and trading volume based on the modified mixture model with Markov switching (MS) volatility specification. We extend this specification to heavier-tailed distribution for the returns using the Student t distribution. We have constructed an algorithm based on Markov Chain Monte Carlo (MCMC) simulation methods to estimate all the parameters and latent quantities in the model using the Bayesian approach. As a by product of the MCMC algorithm, we were able to produce an estimate of the latent information process, which can be used in financial modeling. Our estimation result shows that the estimate of the persistence parameter drops and the estimate of the variance error rises in the volatility specification. Markov switching models are supported by the BP data.

This article makes certain contributions, but several extensions are still possible. First, we focus on symmetrical distributions for ε_t, but skewed distributions, such as the skewed normal or the skewed t distribution can be used. Second, we specify the log of volatility as a simple AR(1) process with MS, but more elaborate models, such as long memory models, may be required to specify volatility. On the other hand, we can include a dynamic pattern in the parameters m_0 and m_1 considering them as time varying parameters, and finally we can extend the model to include many assets.

APPENDIX A: THE FULL CONDITIONALS

In this Appendix, we described the full conditional distributions for the parameters and the latent $\mathbf{S}_{0:T}$ indicators of the MSSVt-VOL model.

A.1. Full conditional distribution of ϕ, $\boldsymbol{\gamma}$ and σ^2

The prior distributions are set as: $\phi \sim \mathcal{N}_{(-1,1)}(\bar{\phi}, \sigma^2_{\phi})$, $\sigma^2_{\eta} \sim \mathcal{IG}(T_0/2, M_0/2)$, $\boldsymbol{\gamma} \sim N_{2(\gamma_1<0)}(\bar{\gamma}, \bar{B})$. According with (11), we have the following full conditional for ϕ:

$$p(\phi \,|\, \mathbf{y}_{1:T}, \mathbf{v}_{1:T}, \mathbf{h}_{0:T}, \mathbf{S}_{0:T}, \boldsymbol{\gamma}) \propto Q(\phi) \exp\left\{ -\frac{a}{2\sigma^2_{\eta}} \left(\phi - \frac{b}{a} \right)^2 \right\} I_{-1<\phi<1}$$

where

$$Q_{\phi} = \sqrt{1-\phi^2} \exp\left\{ -\frac{1}{2\sigma^2_{\eta}} (1-\phi^2) \left(h_0 - \frac{\alpha_{S_0}}{1-\phi} \right)^2 \right\}$$

$$a = \sum_{t=1}^{T} h^2_{t-1} + \frac{\sigma^2_{\eta}}{\sigma^2_{\phi}}, \quad b = \sum_{t=1}^{T} h_{t-1}(h_t - \alpha_{s_t}) + \bar{\phi} \frac{\sigma^2_{\eta}}{\sigma^2_{\phi}}$$

and $I_{-1<\phi<1}$ is an indicator variable. As $p(\phi \,|\, \mathbf{y}_{1:T}, \mathbf{v}_{1:T}, \mathbf{h}_{0:T}, \mathbf{S}_{0:T}, \boldsymbol{\gamma})$ does not have closed form, we sample from using the Metropolis–Hastings algorithm with proposal density the truncated $\mathcal{N}_{(-1,1)}(b/a, \sigma^2_{\eta}/a)$.

To calculate the conditional posterior of $\boldsymbol{\gamma}$, we define

$$\mathbf{Z} = \mathbf{X}\boldsymbol{\gamma} + \boldsymbol{\eta}$$

DOI: 10.1002/asmb

with $\mathbf{Z}' = ([(1-\phi^2)/\sigma^2]^{1/2} h_0,\ 1/\sigma[h_1 - \phi h_0], \ldots, 1/\sigma[h_T - \phi h_{T-1}])$

$$\mathbf{X}' = \begin{pmatrix} \left[\dfrac{1}{\sigma^2}\dfrac{1+\phi}{1-\phi}\right]^{1/2} & \dfrac{1}{\sigma} & \cdots & \dfrac{1}{\sigma} \\ \left[\dfrac{1}{\sigma^2}\dfrac{1+\phi}{1-\phi}\right]^{1/2} S_0 & \dfrac{1}{\sigma} S_1 & \ldots & \dfrac{1}{\sigma} S_T \end{pmatrix}$$

and $\boldsymbol{\eta} \sim \mathcal{N}(0, I_{K+1})$. Then, we have that the full conditional for $\boldsymbol{\gamma}$ is given by $\mathcal{N}_{2(\gamma_1<0)}(\mu_\gamma, B_1)$, where $\boldsymbol{\mu}_\gamma = (\bar{\mathbf{B}}^{-1} + \mathbf{X}'\mathbf{X})^{-1}(\bar{\mathbf{B}}^{-1}\bar{\boldsymbol{\gamma}} + \mathbf{X}'\mathbf{Z})$ and $\mathbf{B}_1 = (\bar{\mathbf{B}}^{-1} + \mathbf{X}'\mathbf{X})^{-1}$.

From (11), the conditional posterior of σ_η^2 is $\mathcal{IG}(T_1/2, M_1/2)$, where $T_1 = T_0 + T + 1$ and

$$M_1 = M_0 + \left[(1-\phi^2)\left(h_0 - \frac{\alpha_{S_0}}{1-\phi}\right)^2\right] + \sum_{t=1}^{T}(h_t - \alpha_{S_t} - \phi h_{t-1})^2$$

A.2. Full conditionals of m_0 and m_1

We assume that prior distributions are, respectively, $m_0 \sim \mathcal{G}(a_0, b_0)$ and $m_1 \sim \mathcal{G}(a_1, b_1)$. Then the full the conditionals follows:

$$p(m_0 \mid \mathbf{y}_{1:T}, \mathbf{v}_{1:T}, \mathbf{h}_{0:T}, m_1) \propto \exp\{-(b_0+T)m_0\} \times \exp\left\{\sum_{t=1}^{T} \log[m_0^{a_0-1/T}(m_0 + m_1\exp(h_t))^{v_t}]\right\}$$

$$p(m_1 | \mathbf{y}_{1:T}, \mathbf{v}_{1:T}, \mathbf{h}_{0:T}, m_1) \propto \exp\left\{-m_1\left(b_1 + \sum_{t=1}^{T}\exp h_t\right)\right\}$$
$$\times \exp\left\{\sum_{t=1}^{T}\log[m_1^{a_1-1/T}(m_0 + m_1\exp(h_t))^{v_t}]\right\}$$

Since the above full conditional distributions are not in any known closed form, we must simulate m_0 and m_1 using the Metropolis–Hastings algorithm. The proposal density used are $\mathcal{N}_{(m_i>0)}(\mu_{m_i}, \tau^2_{m_i})$, with $\mu_{m_i} = x - q'(x)/q''(x)$ and $\tau^2_{m_i} = (-q''(x))^{-1}$ for $i = 0, 1$, where x is the value of the previous iteration, $q(\cdot)$ is the logarithm of the conditional posterior density, and $q'(\cdot)$ and $q''(\cdot)$ are the first and second derivatives, respectively.

A.3. Full conditionals of p_{00}, p_{11} and π_0

The prior distributions for p_{00}, p_{11} and π_0 are, respectively, given by $p_{ii} \sim \mathcal{B}e(\lambda_{i_0}, \lambda_{i_1}), i = 0, 1$ and $\pi_0 \sim \mathcal{B}e(l_0, l_1)$. Then, the full conditional posteriors are: $p_{ii} \sim \mathcal{B}e(\lambda^*_{i_0}, \lambda^*_{i_1})$ and $\pi_0 \sim \mathcal{B}e(l^*_0, l^*_1)$, where $\lambda^*_{i_j} = \sum_{t=1}^{T} I(S_{t-1} = i, S_t = j) + \lambda_{i_j}$ for $i, j = 0, 1$ and $l^*_i = l_i + I(S_0 = i)$, for $i = 0, 1$.

A.4. Full conditional of λ_t and v

As $\lambda_t \sim \mathcal{G}(v/2, v/2)$, the full conditional of λ_t is given by

$$p(\lambda_t \mid y_t, h_t, v) \propto \lambda_t^{(v+1)/2-1} e^{-\lambda_t/2(y_t^2 e^{-h_t}+v)} \tag{A1}$$

which is the gamma distribution, $\mathcal{G}((v+1)/2, (y_t^2 e^{-h_t} + v)/2)$.

We assume the prior distribution of v as $\mathcal{G}(a_v, b_v)\mathbb{I}_{2<v\leqslant 40}$. Then, the full conditional of v is

$$p(v\,|\,\boldsymbol{\lambda}_{1:T}) \propto \frac{\left[\frac{v}{2}\right]^{Tv/2} v^{a_v-1} e^{-v/2[\sum_{t=1}^{T}(\lambda_t - \log\lambda_t)+2b_v]}}{\left[\Gamma\left(\frac{v}{2}\right)\right]^T}\mathbb{I}_{2<v\leqslant 40} \tag{A2}$$

We sample v by the Metropolis–Hastings acceptance–rejection algorithm [28, 29]. Let v^* denote the mode (or approximate mode) of $p(v\,|\,\boldsymbol{\lambda}_{1:T})$, and let $\ell(v) = \log p(v\,|\,\boldsymbol{\lambda}_{1:T})$. As $\ell(v)$ is concave, we use the proposal density $\mathcal{N}_{(2,40)}(\mu_v, \sigma_v^2)$, where $\mu_v = v^* - \ell'(v^*)/\ell''(v^*)$ and $\sigma_v^2 = -1/\ell''(v^*)$. $\ell'(v^*)$ and $\ell''(v^*)$ are the first and second derivatives of $\ell(v)$ evaluated at $v = v^*$. To proof the concavity of $\ell(v)$, we use the result of Abramowitz and Stegun [40], in which the $\log\Gamma(v)$ could be approximated as

$$\log\Gamma(v) = \frac{\log(2\pi)}{2} + \frac{2v-1}{2}\log(v) - v + \frac{\theta}{12v}, \quad 0<\theta<1 \tag{A3}$$

Taking the second derivative of $\ell(v)$ from (A2) and using (A3), we have that

$$\ell''(v) = -\frac{T\theta}{3v^3} - \frac{(T+2a_v-2)}{2v^2} < 0$$

A.5. *Full conditional of* $\mathbf{S}_{0:T}$

The states $\mathbf{S}_{0:T}$ are simulated as a block as proposed by Carter and Kohn [41] and employed by So *et al.* [13]. The joint conditional distribution, $p(\mathbf{S}_{0:T}\,|\,\boldsymbol{\theta}, \mathbf{y}_{1:T}, \mathbf{v}_{1:T}, \mathbf{h}_{0:T})$, can be decomposed as

$$p(\mathbf{S}_{0:T}\,|\,\boldsymbol{\theta}, \mathbf{y}_{1:T}, \mathbf{v}_{1:T}, \mathbf{h}_{0:T}) \propto p(S_T\,|\,\boldsymbol{\theta}, \mathbf{y}_{1:T}, \mathbf{v}_{1:T}, \mathbf{h}_{0:T}) \times \prod_{t=0}^{T-1} p(S_t\,|\,\boldsymbol{\theta}, \mathbf{y}_{1:T}, \mathbf{v}_{1:T}, \mathbf{h}_{0:T}, \mathbf{S}_{t+1:T})$$

where $\mathbf{S}_{t+1:T} = (S_{t+1}, \ldots, S_T)'$. This equation provides a scheme to draw S_T from $p(S_T\,|\,\boldsymbol{\theta}, \mathbf{y}_{1:T}, \mathbf{v}_{1:T}, \mathbf{h}_{0:T})$ and then, recursively, S_t is generated from $p(S_t\,|\,\boldsymbol{\theta}, \mathbf{y}_{1:T}, \mathbf{v}_{1:T}, \mathbf{h}_{0:T}, \mathbf{S}_{t+1:T})$, for $t = T-1, \ldots, 0$. Suppressing for notational convenience the dependence on $\boldsymbol{\theta}, \mathbf{y}_{1:T}, \mathbf{v}_{1:T}$, we have that

$$p(S_t\,|\,\mathbf{h}_{0:T}, \mathbf{S}_{t+1:T}) = p(S_t\,|\,\mathbf{h}_{0:t}, S_{t+1}) = \frac{p(S_t\,|\,\mathbf{h}_{0:t})\,p(S_{t+1}\,|\,S_t)}{p(S_{t+1}\,|\,\mathbf{h}_{0:t})}$$

Since $p(S_T\,|\,\mathbf{h}_{0:T})$ and $p(S_t\,|\,\mathbf{h}_{0:t}, S_{t+1})$ are discrete, $\mathbf{S}_{0:T}$ can be obtained by simulating $T+1$ random numbers from the uniform distributions over $[0, 1]$ [42].

ACKNOWLEDGEMENTS

The authors would like to thank the Editor and two anonymous referees for their useful comments, which improved the quality of this article. The first author gratefully acknowledges financial support from the FAPERJ under number E-26/171.092/2006. Remaining errors are ours alone.

REFERENCES

1. Karpoff JM. The relation between price changes and trading volume: a survey. *The Journal of Financial and Quantitative Analysis* 1987; **22**:109–126.
2. Clark PK. A subordinated stochastic process model with finite variance for speculative prices. *Econometrica* 1973; **41**:135–155.
3. Tauchen GE, Pitts M. The price variability-volume relationships in speculative markets. *Econometrica* 1983; **51**:485–506.
4. Richardson M, Smith T. A direct test of the mixture distribution hypothesis: measuring the daily flow of information. *The Journal of Financial and Quantitative Analysis* 1994; **15**:183–202.
5. Foster F, Wiswanathan S. Can speculative trading explain the trading volume relation. *Journal of Business and Economic Statistics* 1995; **13**:379–396.
6. Harris L. Transaction data test of the mixture distribution hypothesis. *The Journal of Financial and Quantitative Analysis* 1987; **22**:127–141.
7. Andersen T. Return volatility and trading volume: an information flow interpretation of stochastic volatility. *Journal of Finance* 1996; **51**:169–204.
8. Andersen T. Stochastic autoregressive volatility: a framework for volatility modeling. *Mathematical Finance* 1994; **4**:75–102.
9. Mahieu R, Bauer R. A Bayesian analysis of stock return volatility and trading volume. *Applied Financial Economics* 1998; **8**:671–687.
10. Watanabe T. Bayesian analysis of dynamic bivariate mixture models: can they explain the behavior of returns and trading volume? *Journal of Business and Economic Statistics* 2000; **18**:199–210.
11. Lamoureux C, Lastrapes G. Persistence in variance, structural change, and the GARCH. *Journal of Business and Economic Statistics* 1998; **8**:225–234.
12. Hamilton JD, Susmel R. Autoregressive conditional heteroskedasticity and changes in regime. *Journal of Econometrics* 1994; **45**:39–70.
13. So M, Lam K, Li W. A stochastic volatility model with Markov switching. *Journal of Business and Economic Statistics* 1998; **15**:183–202.
14. Kalimipalli M, Susumel R. Regime-switching stochastic volatility and short-term interest rates. *Journal of Empirical Finance* 2004; **11**:309–329.
15. Hamilton JD. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* 1989; **57**:357–384.
16. Glosten LR, Milgrom PR. Bid, ask, and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics* 1985; **14**:71–100.
17. Jacquier E, Polson N, Rossi P. Bayesian analysis of stochastic volatility models. *Journal of Business and Economic Statistics* 1994; **12**:371–418.
18. Kim S, Shepard N, Chib S. Stochastic volatility: likelihood inference and comparison with ARCH models. *Review of Economic Studies* 1998; **65**:361–393.
19. Mahieu R, Schotman PC. An empirical application of stochastic volatility models. *Journal of Applied Econometrics* 1998; **13**:333–360.
20. Abanto-Valle CA, Lopes HF, Migon HS. Stochastic volatility estimation: a modern dynamic model viewpoint. *Technical Report 210*, Departament of Statistics, Federal University of Rio de Janeiro, 2007.
21. Liesenfeld R, Jung RC. Stochastic volatility models: conditional normality versus heavy-tailed distrutions. *Journal of Applied Econometrics* 2000; **15**:137–160.
22. Chib S, Nardari F, Shepard N. Markov chain Monte Carlo methods for stochastic volatility models. *Journal of Econometrics* 2002; **108**:281–316.
23. Jacquier E, Polson N, Rossi P. Bayesian analysis of stochastic volatility models with Fat-tails and correlated errors. *Journal of Econometrics* 2004; **122**:185–212.
24. Shibata M, Watanabe T. Bayesian analysis of a Markov switching stochastic volatility model. *Journal of the Japan Statistical Society* 2005; **35**:205–219.
25. Shephard N, Pitt M. Likelihood analysis of non-Gaussian measurements time series. *Biometrika* 1997; **84**:653–667.
26. Watanabe T, Omori Y. A multi-move sampler for estimate non-Gaussian time series model: comments on Shepard and Pitt (1997). *Bimetrika* 2004; **91**:246–248.
27. Omori Y, Watanabe T. Block sampler and posterior mode estimation for asymmetric stochastic volatility models. *Computational Statistics and Data Analysis* 2008; **52**:2892–2910.

DOI: 10.1002/asmb

28. Tierney L. Markov chains for exploring posterior distributions (with discussion). *Annal of Statistics* 1994; **21**:1701–1762.
29. Chib S, Greenberg E. Understanding the Metropolis–Hastings algorithm. *The American Statistician* 1995; **49**: 327–335.
30. de Jong P, Shepard N. The simulation smoother for time series models. *Biometrika* 1995; **82**:339–350.
31. Koopman S. Disturbance smoothers for state space models. *Biometrika* 1993; **80**:117–126.
32. Pemstein D, Quinn KV, Martin AD. The Scythe statistical library: an open source C++ library for statistical computation. *Journal of Statistical Software* 2007; **V**:1–29.
33. Geweke J. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In *Bayesian Statistics*, Bernardo JM, Berger JO, David AP, Smith AFM (eds), vol. 4. Oxford University Press: Oxford, 1992; 169–193.
34. Heidelbelger P, Welch P. Simulation run length control in the presence of a initial transient. *Operations Research* 1983; **31**:1109–1144.
35. Carvalho C, Lopes HF. Simulation-based sequential analysis of Markov switching stochastic volatility models. *Computational Statistics and Data Analysis* 2007; **51**:4526–4542.
36. Hamilton JD. Rational expectations econometrics analysis of changes in regime. *Journal of Economic Dynamics and Control* 1988; **12**:385–423.
37. Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society*, *Series B* 2002; **64**:583–640.
38. Berg A, Meyer R, Yu J. Deviance information criterion for comparing stochastic volatility models. *Journal of Business and Economic Statistics* 2004; **22**:107–120.
39. Celeux G, Forbes F, Robert CP, Titterington DM. Deviance information criteria for missing data models. *Bayesian Analysis* 2006; **1**:651–674.
40. Abramowitz M, Stegun N. *Handbook of Mathematical Functions*. Dover Publications Inc.: New York, 1970.
41. Carter CK, Kohn R. On Gibbs sampling for state space models. *Biometrika* 1994; **81**:541–553.
42. Ripley BD. *Stochastic Simulation*. Wiley: New York, 1987.

DOI: 10.1002/asmb