Quantitative Finance and Economics

http://www.aimspress.com/journal/QFE

QFE, 2(2): 325–347.
DOI:10.3934/QFE.2018.2.325
Received: 31 August 2017
Accepted: 23 April 2018
Published: 09 May 2018

Research article

A new variant of estimation approach to asymmetric stochastic volatility model

Zhongxian Men[1,*] **and Tony S. Wirjanto**[2,3]

[1] Quantitative Research, JPMorgan Chase & Co., 277 Park Avenue, New York, USA
[2] Department of Statistics and Actuarial Science, University of Waterloo, 200 University Avenue West, Waterloo, Ontario, Canada
[3] School of Accounting and Finance, University of Waterloo, 200 University Avenue West, Waterloo, Ontario, Canada

* **Correspondence:** Email: Zhongxian.men@jpmchase.com; Tel: +13138981855.

Abstract: This paper proposes a novel simulation-based inference for an asymmetric stochastic volatility model. An acceptance-rejection Metropolis-Hastings algorithm is developed for the simulation of latent states of the model. A simple and efficient algorithm is also developed for estimation of a heavy-tailed stochastic volatility model. Simulation studies show that our proposed methods give rise to reasonable parameter estimates. Our proposed estimation methods are then used to analyze a benchmark data set of asset returns.

Keywords: stochastic volatility; leverage effect; Bayesian inference; acceptance-rejection; Metropolis-Hastings; slice sampler
JEL classification numbers: C10, C11, C13, C15, C22, C58, C53

1. Introduction

In the study of financial econometrics, stochastic volatility (SV) models have been developed to model the time varying and clustering volatility of asset returns. There are mainly two types of models studied in the literature. One type of these models includes autoregressive conditional heteroskedasticity (ARCH) models and generalized autoregressive conditional heteroskedasticity (GARCH) models and the other includes stochastic volatility (SV) models. The GARCH models, proposed by Bollerslev (1986), are extensions of the ARCH model studied by Engle (1982), and have been extended in various directions. In the GARCH-type models, volatility is often defined as a deterministic function of the previous observed asset returns and volatilities. In the SV models, initially studied in Taylor (1986), the log volatility of asset returns is often modeled as a latent first order stationary autoregressive (AR (1)) process. SV models are more attractive because they are close to the models often used in financial theory to represent the behaviour of financial prices. Comparing with the GARCH models, the SV models capture more realistically the main empirical properties often observed in daily behaviors of financial time series (see, for example, Broto and Ruiz, 2004; Carnero et al., 2003).

In the estimation of the univariate SV models, many methods have been proposed in the literature as well, such as quasi-maximum likelihood (QML) method by Harvey et al. (1997), numerical integration method in Kawakatsu (2007), Simulated maximum likelihood (SML) methods by Dobigeon and Tourneret (2010). Bayesian inference approaches based on Markov Chain Monte Carlo (MCMC) methods have been proposed in Jacquier et al. (2004); Kim et al. (1998); Omori et al. (2007); Zhang and King (2008); Men (2012); Men et al. (2017) and Wirjanto et al. (2016) for the SV models with or without leverage effects. The greatest advantage of the MCMC methodology is that a large dimensional problem can be divided into several lower dimensional simulation tasks in which the log volatilities are estimated simultaneously. Broto and Ruiz (2004) claim that MCMC approaches are more efficient among other estimation methods such as the QML and the generalized method of moments proposed in Melino and Turnbull (1990). As usual, in the MCMC methods for the SV models, posterior distributions of parameters and augmented parameters are assumed to be either known or proportional to some positive functions. These distributions generally can not be sampled directly and the simulation is usually carried out through the Metropolis-Hastings (MH) algorithm. It is well known that the performance of the MH algorithm depends critically on the selection of a proposal distribution. However, choosing an appropriate proposal distribution in general is difficult and different proposal distributions tend to give different acceptance rates. Due to the high correlation among the latent states, a careful simulation of the log volatilities is required, which is discussed in Jacquier et al. (2004). There are two main methods for simulating the latent states. One is called the single-move simulation method developed in Jacquier et al. (2004); Yu and Meyer (2006); Zhang and King (2008); Kim et al. (1998); Men et al. (2017) in which the states are simulated one at a time. The other method, named the block sampling, is introduced by Shephard and Pitt (1997) and has been employed in Pitt and Shephard (1999a) and Chib et al. (2006). In a block sampling algorithm, the latent states are divided into random blocks and the blocks are sampled via the MH algorithm. The proposal distribution is either a multivariate Gaussian as in Shephard and Pitt (1997) or a multivariate Student-*t* distribution as in Chib et al. (2006), where the modes of these proposal distributions can be found by use of the Newton-Raphson method. As discussed in Shephard and Pitt (1997), the Newton-Raphson algorithm may converge slowly, rendering the block sampling technique to be computationally highly intensive.

In this article, we focus on developing new MCMC estimation methods for univariate SV models where the residuals of the measurement equation follow either a univariate standard normal or a Student-*t* distribution. The non-zero correlation between the innovations of asset returns and the latent AR(1) process is permitted. The aim of our methods is to consider a simulation-based inference for the parameters and log volatilities. Our contributions to the literature are two-folds. First, we develop single-move algorithms for simulating latent states based on the acceptance-rejection Metropolis-Hastings (ARMH) algorithm introduced in Chib and Greenberg (1995). The proposal distribution is simulated by the slice sampler introduced in Neal (2003). The second contribution is to propose a simple but efficient algorithm for computing the heavy-tailed SV model. Comparing with the method where a mixture decomposition of the Student-t distribution (for examples, see Jacquier et al., 2004; Zhang and King, 2008) is used at each observation time, our method uses the Student-t distribution directly. As a result, in our approach, we do not need to estimate the extra parameters from the mixture decomposition. Therefore our proposed ARMH methods are simple and easier to implement.

The rest of the paper is organized as follows. In Section 2, we propose ARMH approaches for the SV and heavy-tailed SV models. To asses the goodness of fit, in addition to check the realized innovations of the measurement equation, we test the so-called probability integral transforms (PITs) calculated from the density forecast introduced by Diebold et al. (1998). Since it is extremely difficult to obtain the analytical conditional densities of observed data, we employ the auxiliary particle filter in Pitt and Shephard (1999b) to evaluate the likelihoods. Section 3 presents simulation studies for the SV model. In Section 4, we apply our estimation methods for the SV and heavy-tailed SV models to a benchmark data set of asset returns which has been studied for comparative purposes purposes. Concluding remarks are drawn in the last section.

2. Estimation of asymmetric stochastic volatility models

2.1. A brief introduction of the acceptance-rejection Metropolis-Hastings method

Suppose that we wish to generate samples from a distribution with the density function $\pi(x) \propto f(x)/k$, where k is the unknown normalizing constant. Let $g(x)$ be a density function that can be sampled by a known method, and suppose that there exists a known constant c satisfying $f(x) \leq cg(x)$ for any x. Then to generate a random sample from $f(x)$ we use the following acceptance-rejection (AR) procedure:

1. Generate a candidate y from $g(.)$ and a value u from a uniform distribution $\mathcal{U}(0, 1)$.
2. If $u \leq f(y)/(cg(y))$, then return $x = y$; else go to step 1.

It is easy to see that the expected number of iterations of the AR algorithm to generate a sample point from $f(x)$ is $1/c$, which means that this sampling method can be optimized simply by setting

$$c = \sup_x \left\{ \frac{f(x)}{g(x)} \right\}.$$

It seems that the AR method is simple and in most of cases the constant c is easier to find. But if c is too small, then the AR method would not be efficient as we would have to wait a long time to generate a sample from $f(x)$. Primarily because of this concern, Metropolis-Hastings (MH) methods have been widely used in this literature. The typical MH algorithm can be described as the follows. Suppose that $q(x, y)$ is a candidate density given x. Let $\alpha(x, y)$ be a move probability from $x \to y$. Then the acceptance probability of a newly generated value from $q(x, y)$ is

$$\alpha(x, y) = \begin{cases} \min\left\{\frac{\pi(y)q(y,x)}{\pi(x)q(x,y)}, 1\right\}, & \text{if } \pi(x)q(x, y) > 0, \\ 1, & \text{otherwise.} \end{cases}$$

The procedure of sampling $\pi(x)$ is the following,

1. Initialize the chain with $x^{(0)}$.

2. Repeat the following for $n = 1, ..., N$, where N is a positive integer.
   - Generate y from $q(x^{(n)}, y)$ and u from $\mathcal{U}(0, 1)$.
   - if $u \leq \alpha(x^{(n)}, y)$, then set $x^{(n+1)} = y$; else set $x^{(n+1)} = x^{(n)}$ keeping the sampled value from the last iteration.
3. Return the values $\mathbf{x} = \{x^{(1)}, x^{(2)}, ..., x^{(N)}\}$.

After discarding the first m generated values from $\mathbf{x}$ as a burn in, values $\{x^{(m+1)}, x^{(m+2)}, ..., x^{(N)}\}$ are treated as a sample generated from the target density $\pi(x)$.

It is noticed that the candidate distribution of $q(x, y)$ is usually difficult to find, which may require a Taylor expansion or the Newton-Raphson method, which is, in some cases, computationally too costly especially for SV models. In response to this issue, Chib and Greenberg (1995) introduced a combination of AR and MH methods, which is called, in this paper, an acceptance-rejection Metropolis-Hastings (ARMH) method described below.

The ARMH algorithm for sampling $\pi(x)$

1. Generate y using the AR algorithm described previously.
2. Let $C_1 = \{f(x) < cg(x)\}$ and $C_2 = \{f(y) < cg(y)\}$.
3.
   - If $C_1 = 1$, then $\alpha = 1$.
   - If $C_1 = 0$ and $C_2 = 1$, then $\alpha = cg(x)/f(x)$.
   - If $C_1 = 0$ and $C_2 = 0$, then $\alpha = \min\left\{\frac{f(y)g(x)}{f(x)g(y)}, 1\right\}$.
4. Generate u from $\mathcal{U}(0, 1)$. If $u < \alpha$, then return $x = y$; else keep x.

2.2. *Estimation of asymmetric stochastic volatility model*

We will first propose a Bayesian estimation method for an asymmetric stochastic volatility (ASV) model. There are two types of ASV models, one was proposed in Jacquier et al. (2004) and the other was formulated in Harvey and Shephard (1996). Yu (2005) studied these two types of ASV models and showed that the ASV model in Harvey and Shephard (1996) is able to capture the leverage effect between asset returns and their future volatilities. Two different formulations of leverage effect between asset returns and the latent volatilities were also studied in Men et al. (2017). Based on this, in this paper, we focus on estimation of the ASV model proposed by Harvey and Shephard (1996).

Let y_t denote the observed asset return at time t, $t \leq T$, where T is the sample size. Without loss of generality, we assume that the expectation of y_t is zero such that $E(y_t) = 0$. Then the ASV model can be expressed as

$$y_t = \exp(h_t/2)\epsilon_t, t = 1, ..., T, \tag{1}$$

$$h_{t+1} = \mu + \phi(h_t - \mu) + \sigma\eta_{t+1}, t = 1, ..., T-1, \tag{2}$$

$$h_0 \sim \mathcal{N}(\mu, \sigma^2/(1-\phi^2)), \tag{3}$$

where

$$\begin{pmatrix} \epsilon_t \\ \eta_{t+1} \end{pmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right). \tag{4}$$

The two innovations ϵ_t and η_{t+1} are assumed to jointly follow an independently and identically distributed (*i.i.d.*) bivariate standard normal distribution with a correlation coefficient given by

$\rho = corr(\epsilon_t, \eta_{t+1})$. In order for the ASV model to be weakly stationary, it is assumed that $|\phi| < 1$. A priori, the coefficient ρ is expected to have a negative sign, which means that the returns y_t and the future log volatilities h_{t+1} are expected to be negatively correlated. This is often interpreted as evidence of a leverage effect in time series of asset returns.

To estimate the parameters of the model, we usually perform a Cholesky decomposition on the correlation matrix in (4), and the ASV model is then transformed to have the following representation

$$y_t = \exp(h_t/2)\epsilon_t, t = 1, ..., T, \quad (5)$$
$$h_{t+1} = \mu + \phi(h_t - \mu) + \psi y_t \exp(-h_t/2) + \tau u_{t+1}, \quad (6)$$
$$h_0 \sim \mathcal{N}(\mu, \sigma/(1 - \phi^2)), \quad (7)$$

where $\psi = \sigma\rho$ and $\tau = \sigma\sqrt{(1 - \rho^2)}$. In the derived MCMC algorithm to estimate the parameters, instead of simply using the posterior distributions of ρ and σ directly, we propose to simulate the posterior distributions of ψ and τ^2, respectively. After these two parameters have been sampled, the original parameters are then obtained from the two equations: $\sigma = \sqrt{\psi^2 + \tau^2}$ and $\rho = \psi/\sigma$. Define by $\theta = (\mu, \phi, \psi, \tau)$ as the parameter vector of the ASV model. For convenience, we define $\mathbf{y} = (y_1, ..., y_T)$, and $\mathbf{h} = (h_1, ..., h_T)$.

The ASV model is completed by specifying the proper prior distributions for the parameters. We assume that all of the prior distributions of the parameters in the model are mutually independent. The prior distributions of μ and ψ are $\mu \sim \mathcal{N}(0, 10)$, and $\psi \sim \mathcal{N}(0, 10)$, respectively. These prior distributions result in reasonably flat densities over their support regions. All of the above prior distributions are conjugate which is convenient for the calculation of the posterior distributions. To impose a weak stationary condition on the latent process, the prior distribution of ϕ is a normal distribution truncated in the interval $(-1, 1)$. The prior distribution of τ^2 is an inverse Gamma distribution $\tau^2 \sim \mathcal{IG}(5, 0.05)$ as in Pitt and Shephard (1999a).

We present an MCMC algorithm for parameter estimation of the ASV model in Table 1 followed by a detailed description of the procedure.

Table 1. MCMC algorithm for the ASV model.

Step 0. Initialize $\mathbf{h}$, μ, ϕ, ψ and τ.

Step 1. Sample $h_t, t = 1, ..., T$.

Step 2. Sample ϕ.

Step 3. sample μ, ψ and τ^2.

Step 4. Go to Step 1.

Step 0. Initialize $\mathbf{h}$, μ, ϕ, ψ and τ. For the start of the MCMC algorithm, the parameters of the latent Markov process are set as $\mu = -0.5$, $\phi = 0.5$, $\psi = -0.5$ and $\tau = 0.5$, respectively. The initial values of $\mathbf{h}$ are generated from the latent first order autoregressive process with the above initialized parameters.

Step 1. Sample $\mathbf{h}$. The simulation is conducted via a single-move ARMH algorithm. The full conditionals of the latent random variables are expressed as

$$
\begin{aligned}
f(h_1|\mathbf{y}, h_2, \theta) &\propto f(y_1|h_1) f(h_1|\theta) f(h_1|h_2, y_1, \theta),\\
f(h_t|\mathbf{y}, h_{t-1}, h_{t+1}, \theta) &\propto f(y_t|h_t) f(h_t|h_{t-1}, y_{t-1}, \theta) f(h_t|h_{t+1}, y_t, \theta),\\
f(h_T|\mathbf{y}, h_{T-1}, \theta) &\propto f(y_T|h_T) f(h_T|h_{T-1}, y_{T-1}, \theta),
\end{aligned}
$$

where $f(y_t|h_t), t = 1, ..., T$, are the conditional densities of y_t at discrete time points and $f(h_1|\theta)$ is the density of the latent log volatility h_1, $f(h_t|h_{t-1}, y_{t-1}, \theta)$ and $f(h_t|h_{t+1}, y_t, \theta)$ are the conditional densities of h_t given h_{t-1} and of h_t given h_{t+1} by the latent equation (6), respectively. Since y_T is the last observation, the posterior distribution of h_T depends only on $y_{T-1}$and h_{T-1}.

We only present the full conditionals of $h_t, t = 2, ..., T - 1$. The full conditionals of h_1 and h_T are easily derived and therefore not provided here.

The full conditional of h_t is

$$
\begin{aligned}
&f(h_t|\mathbf{y}, h_{t-1}, h_{t+1}, \theta)\\
&= c_{1t} f(y_t|h_t) f(h_t|h_{t-1}, y_{t-1}, \theta) f(h_t|h_{t+1}, y_t, \theta)\\
&= c_{2t} \exp\left\{ -\frac{h_t}{2}\right\} \exp\left\{ -\frac{y_t^2 \exp(-h_t)}{2}\right\}\\
&\quad \times \exp\left\{ -\frac{[(h_t-\mu) - \phi(h_{t-1}-\mu) - \psi y_{t-1}\exp(-h_{t-1}/2)]^2}{2\tau^2}\right\}\\
&\quad \times \exp\left\{ -\frac{[(h_{t+1}-\mu) - \phi(h_t-\mu) - \psi y_t\exp(-h_t/2)]^2}{2\tau^2}\right\} \qquad (8)\\
&< c_{2t} \exp\left\{ -\frac{h_t}{2}\right\} \exp\left\{ -\frac{y_t^2 \exp(-h_t)}{2}\right\}\\
&\quad \times \exp\left\{ -\frac{[(h_t-\mu) - \phi(h_{t-1}-\mu) - \psi y_{t-1}\exp(-h_{t-1}/2)]^2}{2\tau^2}\right\} \qquad (9)
\end{aligned}
$$

where c_{1t} and c_{2t} are the two normalizing constants. The inequality sign in (9) holds because the last part of the right-hand side of equation in (8) is less than 1. Suppose that the dominate distribution can be sampled by some sophisticated method. Then the full conditional of h_t can be simulated by the acceptance-rejection method. By looking at the inequality of (9), we notice that c_{2t} can be obtained as follows,

$$
c_{2t} = \sup_{h_t} \exp\left\{ -\frac{[(h_{t+1}-\mu) - \phi(h_t-\mu) - \psi y_t\exp(-h_t/2)]^2}{2\tau^2}\right\}.
$$

As is usually observed, c_{2t} is very small meaning that the number of iterations for the acceptance-rejection method to generate a point from the conditional distribution of h_t is really large. In other words, if we use the acceptance-rejection algorithm in this way, the algorithm would be very inefficient. For this reason, we use a modified ARMH procedure described in Chib and Greenberg (1995).

Focusing on the last term of the full conditional of h_t, we define

$$
g(h_t|.) \propto \exp\left\{ -\frac{h_t}{2}\right\} \exp\left\{ -\frac{y_t^2 \exp(-h_t)}{2}\right\}
$$

$$\times \exp\left\{-\frac{[(h_t-\mu)-\phi(h_{t-1}-\mu)-\psi y_{t-1}\exp(-h_{t-1}/2)]^2}{2\tau^2}\right\}$$
$$\times \exp\left\{-\frac{[(h_{t+1}-\mu)-\phi(h_t-\mu)]^2}{2\tau^2}\right\}, \tag{10}$$

and define

$$c_{3t} = \sup_{h_t}\left\{\frac{f(h_t|.)}{g(h_t|.)}\right\}.$$

Clearly the function $c_{3t}g(h_t|.)$ does not necessarily dominate $f(h_t|.)$. We will use the ARMH method described earlier to sample the full conditional of h_t, where the density $g(h_t|.)$ can be sampled by the slice sampler described in Neal (2003).

Algorithm of the slice sampler for $g(h_t|.)$

It is easy to verify that (10) can be expressed by

$$\begin{aligned}
g(h_t|.) \propto \exp\left\{-\frac{h_t}{2}\right\}\exp\left\{-\frac{y_t^2\exp(-h_t)}{2}\right\}& \\
\times \exp\left\{-\frac{[(h_t-\mu)-\phi(h_{t-1}-\mu)-\psi y_{t-1}\exp(-h_{t-1}/2)]^2}{2\tau^2}\right\}& \\
\times \exp\left\{-\frac{[h_t-\mu-(h_{t+1}-\mu)/\phi]^2}{2\tau^2/\phi^2}\right\}& \\
\propto \exp\left\{-\frac{y_t^2\exp(-h_t)}{2}\right\}\exp\left\{-\frac{(h_t-\mu_t)^2}{2c}\right\}&,
\end{aligned} \tag{11}$$

where

$$\begin{aligned}
c &= 1/a,\\
\mu_t &= -\frac{c}{2}+\frac{b_t}{a},\\
a &= \frac{1}{\tau^2}+\frac{\phi^2}{\tau^2},\\
b_t &= \frac{\mu+\phi(h_{t-1}-\mu)+\psi y_{t-1}\exp(-h_{t-1}/2)}{\tau^2}+\frac{\phi^2\mu-\phi(h_{t+1}-\mu)}{\tau^2}.
\end{aligned}$$

0. Given $h_t^{(k)}$, the sampled point after the k-iteration of the MCMC algorithm.

1. Draw u_1 uniformly from the interval $(0, 1)$.

Let $u_2 = u_1\exp\left\{-\frac{y_t^2\exp(-h_t^{(k)})}{2}\right\}$ and set

$$u_2 \le \exp\left\{-\frac{y_t^2\exp(-h_t)}{2}\right\}.$$

Then we have

$$h_t \ge \log(-y_t^2/(2\log(u_2))),\ \text{if } y_t \ne 0. \tag{12}$$

2. Draw u_3 uniformly from the interval $(0, 1)$.
Let $u_4 = u_3 \exp\left\{-\frac{(h_t^{(k)}-\mu_t)^2}{2c}\right\}$ and set

$$u_4 < \exp\left\{-\frac{(h_t - \mu_t)^2}{2c}\right\}.$$

Then we have

$$\mu_t - \sqrt{-2c\log(u_4)} \leq h_t \leq \mu_t + \sqrt{-2c\log(u_4)}. \tag{13}$$

3. Draw h_t uniformly from the interval determined by the inequalities (12) and (13) such as

$$h_t \sim \mathcal{U}\left(\max\left\{\log(-y_t^2/(2\log(u_2))),\ \mu_t - \sqrt{-2c\log(u_4)}\right\},\ \mu_t + \sqrt{-2c\log(u_4)}\right), \text{if } y_t \neq 0,$$
$$h_t \sim \mathcal{U}\left(\mu_t - \sqrt{-2c\log(u_4)},\ \mu_t + \sqrt{-2c\log(u_4)}\right), \text{if } y_t = 0.$$

The advantage of the slice sampler is that it can adapt to the underlying density function, and therefore is likely to be more efficient. Under some sufficient conditions, Roberts and Rosenthal (1999) show that the slice algorithm has extremely robust geometric ergodicity properties. Mira and Tierney (2002) prove that the slice sampler has a smaller second-largest eigenvalue, which ensures faster convergence to the underlying distribution. The single-move simulation method is popular in the literature. See for instance, in Jacquier et al. (2004); Zhang and King (2008); Men (2012); Men et al. (2015); Men et al. (2017), among others.

Compared with other MCMC methods proposed in the literature, our proposed estimation methods focus on a direct simulation of the laten states based on the density functions. The method proposed by Jacquier et al. (2004), which is also a single move MCMC algorithm, is based on the approximation of the posterior distribution. They approximate the log-normal kernel by an inverse gamma with the same mean and variance. The MCMC method proposed in Omori et al. (2007) is a block sampling method for latent states. Compared with their method, our method may be a bit slower, but as we implemented our algorithm in C, the estimation speed is not a serious concern.

We believe that our proposed MCMC estimation methods is flexible enough and can be generalized to deal with more flexible dynamics. The model proposed in Eraker et al. (2003) introduces Poisson processes to the measurement and volatility processes, respectively. To fit their model in an MCMC framework, we need to know proportionally the posterior distribution of each latent state, which can be derived conditional on other parameters in the model that have been previously simulated. For example, once we have simulated the two jumps in both the measurement and volatility equations, the two jumps are treated as constants and the posterior distribution of the corresponding latent state can be simulated by our proposed methods. Liesenfeld and Richard (2003) introduce a one-factor multivariate statistic volatility model which is estimated by incorporating an efficient importance sampling scheme in the Quasi Maximum Likelihood approach. Our proposed methods can be easily applied to such a model when simulating the latent states of that model. Liesenfeld and Richard (2003) also propose a two-component SV model obtained by extending the univariate SV model to allow the dynamics of volatility to be driven by two independent components. Our proposed method can be directly applied to this model as well. Conditioning on that one element at time t and the two coefficients of elements which have been previously simulated, the posterior distribution of the other element at time t can be calculated based on a univariate SV model and therefore can be simulated by our proposed method. Lastly, for the heavy-tailed SV model in Liesenfeld and Richard (2003), we notice that there is no leverage effect permitted between the two innovation processes of the two equations. Our proposed estimation method can be applied to that model after a simple modification.

Step 2. Sample ϕ. Given the truncated normal prior distribution $\phi \sim \mathcal{N}(\alpha_\phi, \beta_\phi^2)$, the full conditional of ϕ is

$$\begin{aligned} f(\phi|\mathbf{y}, \mu, \psi, \tau) &\propto f(h_1|\theta) \prod_{t=1}^{T-1} f(h_{t+1}|h_t, \theta, y_t) \exp\Big\{ -\frac{(\phi-\alpha_\phi)^2}{2\beta_\phi^2} \Big\} \\ &\propto \mathcal{N}\Big(\frac{d}{c}, \frac{1}{c}\Big)(1-\phi^2)^{\frac{1}{2}}, \end{aligned}$$

where

$$\begin{aligned} c &= \frac{-(h_1-\mu)^2 + \sum_{t=1}^{T-1}(h_t-\mu)^2}{\tau^2} + \frac{1}{\beta_\phi^2}, \\ d &= \frac{\sum_{t=1}^{T-1}(h_t-\mu)(h_{t+1}-\mu-\phi(h_t-\mu)+\psi\exp(-h_t/2)y_t)}{\tau^2} + \frac{\alpha_\phi}{\beta_\phi^2}, \end{aligned}$$

which is proportional to the product of a univariate normal distribution and a positive function. This full conditional can also be sampled by the slice sampler.

Step 3. Sample parameters μ, ψ and τ. Instead of sampling τ directly, we sample τ^2. Since the priors for these parameters are conjugate normal and an inverse Gamma distribution, respectively, the sampling can be done relatively straightforwardly.

2.3. Estimation of the heavy-tailed ASV model

Assuming that the innovations ϵ_t in the ASV model follow a Student-t distribution with unknown v degrees of freedom, the heavy-tailed ASV (ASVt) model is specified as

$$y_t = \exp(h_t/2)\epsilon_t, \qquad \epsilon_t \sim t(v), \tag{14}$$

$$h_{t+1} = \mu + \phi(h_t - \mu) + \tau\eta_{t+1}, \qquad u_{t+1} \sim \mathcal{N}(0,1). \tag{15}$$

In the literature, the Student-t distribution, $t(v)$, is usually decomposed as $\epsilon_t = \sqrt{\lambda_t}e_t$, where $\lambda_t \sim \mathcal{IG}(\alpha,\beta)$, an inverse Gamma distribution and $e_t \sim \mathcal{N}(0,1)$. See for instance, Jacquier et al. (2004), and Zhang and King (2008). Instead of using this mixture decomposition, we consider estimating this ASVt model with the Student-t distribution directly. In order to introduce a correlation structure between y_t and the log volatility, we rewrite equation (15) as

$$h_{t+1} = \mu + \phi(h_t - \mu) + \psi y_t \exp(-h_t/2) + \tau u_{t+1}. \tag{16}$$

The correlation between the asset returns and the latent volatilities is captured by a coefficient, denoted as ψ. Upon our specification, there is only one extra parameter v to be estimated compared with the ASV model, where in the mixture case we estimate this v and plus T augmented parameters λ_t. Again, we propose an ARMH algorithm to fit this heavy-tailed ASV model. Denote by $\theta = (\mu, \phi, \psi, \tau, v)$ the parameter vector of the ASVt model. We only provide the methods to simulate the latent states h_t and v, respectively.

• Sample the latent states $h_t, t = 1, ..., T-1$. The simulation of h_1 and h_T are similar. The full conditionals of $h_t, t = 2, ..., T-1$, is

$$\begin{aligned}
& f(h_t|\mathbf{y}, h_{t-1}, h_{t+1}, \theta) \\
& = c_{1t} f(y_t|h_t) f(h_t|h_{t-1}, y_{t-1}, \theta) f(h_t|h_{t+1}, y_t, \theta) \\
& < c_{2t} e^{-h_t/2}\Big(1 + \frac{y_t^2 e^{-h_t}}{v}\Big)^{-\frac{v+1}{2}} \\
& \qquad \times \exp\Big\{ - \frac{\left[(h_t-\mu) - \phi(h_{t-1}-\mu) - \psi \exp(-h_{t-1}/2) y_{t-1}\right]^2}{2\tau^2}\Big\},
\end{aligned}$$

where c_{1t} is a normalizing constant. Similarly to the previous steps we provided in the estimation of the ASV model, we define

$$c_{2t} = \sup_{h_t} \exp\Big\{ - \frac{\left[(h_{t+1}-\mu) - \phi(h_t-\mu) - \psi y_t \exp(-h_t/2)\right]^2}{2\tau^2}\Big\}.$$

Foucsing on the last term of the full conditional of h_t, we define

$$\begin{aligned}
g(h_t|.) & \propto \exp\Big\{ -\frac{h_t}{2}\Big\}\Big(1 + \frac{y_t^2 e^{-h_t}}{v}\Big)^{-\frac{v+1}{2}} \\
& \qquad \times \exp\Big\{ - \frac{\left[(h_t-\mu) - \phi(h_{t-1}-\mu) - \psi y_{t-1} \exp(-h_{t-1}/2)\right]^2}{2\tau^2}\Big\} \\
& \qquad \times \exp\Big\{ - \frac{\left[(h_{t+1}-\mu) - \phi(h_t-\mu)\right]^2}{2\tau^2}\Big\} \\
& \propto \Big(1 + \frac{y_t^2 e^{-h_t}}{v}\Big)^{-\frac{v+1}{2}} \exp\Big\{ - \frac{(h_t-\mu_t)^2}{2c}\Big\},
\end{aligned} \tag{17}$$

where

$$\begin{aligned}
c &= 1/a, \\
\mu_t &= -\frac{c}{2} + \frac{b_t}{a}, \\
a &= \frac{1}{\tau^2} + \frac{\phi^2}{\tau^2}, \\
b_t &= \frac{\mu + \phi(h_{t-1}-\mu) + \psi y_{t-1}\exp(-h_{t-1}/2)}{\tau^2} + \frac{\phi^2\mu - \phi(h_{t+1}-\mu)}{\tau^2}.
\end{aligned}$$

and define

$$c_{3t} = \sup_{h_t} \Big\{ \frac{f(h_t|.)}{g(h_t|.)} \Big\}.$$

It is easy to see that the function $c_{3t}g(h_t|.)$ does not necessarily dominate $f(h_t|.)$. We again use the ARMH method described earlier to sample the full conditional of h_t, where the density $g(h_t|.)$ can be sampled by the slice sampler described in Neal (2003).

Based on the full conditionals of h_t, $t = 1, ..., T$, we only give the slice sampler for sampling h_t.

0. Given $h_t^{(k)}$, the sampled point after the k-iteration of the MCMC algorithm, we go to the next step.

1. Draw $u_1 \sim \mathcal{U}(0, 1)$. Let $u_2 = u_1 * \left\{1 + \frac{y_t^2 \exp(-h_t^{(k)})}{v}\right\}^{-\frac{v+1}{2}}$ and let

$$u_2 \leq \left\{1 + \frac{y_t^2 \exp(-h_t)}{v}\right\}^{-\frac{v+1}{2}}.$$

If $y_t \neq 0$ then we have

$$\exp(-h_t) \leq v\Big(\Big(\frac{1}{u_2}\Big)^{\frac{2}{v+1}} - 1\Big)\Big/y_t^2,$$

and then,

$$h_t \geq -\log\left[v\Big(\Big(\frac{1}{u_2}\Big)^{\frac{2}{v+1}} - 1\Big)\Big/y_t^2\right]. \tag{18}$$

2. Draw $u_3 \sim \mathcal{U}(0, 1)$.
Let $u_4 = u_3 * \exp\left\{-\frac{(h_t^{(k)} - \mu_t)^2}{2\tau}\right\}$ and let

$$u_4 \leq \exp\left\{-\frac{(h_t - \mu_t)^2}{2\tau}\right\},$$

and then we have,

$$(h_t - \mu_t)^2 \leq -2\tau \log(u_4).$$

Then we have

$$\mu_t - \sqrt{-2\tau \log(u_4)} \leq h_t \leq \mu_t + \sqrt{-2\tau \log(u_4)}. \tag{19}$$

3. If $y_t \neq 0$ draw $h_t^{(k+1)}$ uniformly from the interval determined by the inequalities (18) and (19),

$$h_t \sim \mathcal{U}\Big(\max\Big\{-\log\left[v\Big(\Big(\frac{1}{u_2}\Big)^{\frac{2}{v+1}} - 1\Big)\Big/y_t^2\right],\ \mu_t - \sqrt{-2\tau \log(u_4)}\ \Big\},\ \mu_t + \sqrt{-2\tau \log(u_4)}\ \Big),$$

otherwise,

$$h_t \sim \mathcal{U}\Big(\mu_t - \sqrt{-2\tau \log(u_4)},\ \mu_t + \sqrt{-2\tau \log(u_4)}\ \Big).$$

• Sample v, the degrees of freedom of the Student-t distribution. The full conditional of v is given by

$$f(v|\mathbf{y}, \mathbf{h}, \mu, \phi, \sigma^2) \propto f(\mathbf{y}|\mathbf{h}, v)f(v)$$

$$= f(v)\prod_{t=1}^{T}\frac{v^{v/2}\Gamma((v+1)/2)}{\Gamma(v/2)\Gamma(1/2)}(v+y_t^2\exp(-h_t))^{-(v+1)/2}, \quad (20)$$

where $f(v)$ is a prior density of v. In the literature, there are several ways to specify this prior distribution. For instance, Jacquier et al. (2004) propose a discrete prior distribution $\mathcal{U}[3,40]$ from which the full conditional can be sampled directly from a multinomial distribution. Geweke (1993) suggests $\alpha\exp(-\alpha v)$ with $\alpha = 0.2$ as an alternative, while Zhang and King (2008) choose a normal distribution $v \sim \mathcal{N}(20,25)$. Bauwens and Lubrano (1998) use a Cauchy prior proportional to $1/(1+v^2)$. In our procedure we use a normal prior proposed in Zhang and King (2008). Since this full conditional is an unknown distribution, we use a random-walk Metropolis-Hastings algorithm, in which the proposal density is a standard Gaussian density and the acceptance probability is computed using equation (20).

2.4. Associated particle filter

In the literature of stochastic volatility models, model comparison is usually based on AIC and BIC criteria, which needs the calculation of model likelihood. For ASV and ASVt models, model likelihood is infeasible to compute because of the non-linear structure of the model, which requires that we integrate out the latent states. In this paper we employ an auxiliary particle filter (APF) proposed in Shephard and Pitt (1997) to perform this task, which is a recursive efficient algorithm to approximate the filter and one-step ahead predication distributions of the latent states of the model. The likelihood of the specific ASV model via the successive conditional decomposition is

$$f(\mathbf{y}|\theta) = f(y_1|\theta)\prod_{t=2}^{T} f(y_t|\mathcal{I}_{t-1},\theta), \quad (21)$$

where $\mathcal{I}_t = (y_1,...,y_t)$ is the information known at time t. The conditional density of y_{t+1} given θ and $\mathcal{I}_t$ has the following expression

$$\begin{aligned} f(y_{t+1}|\mathcal{I}_t,\theta) &= \int f(I_{t+1}|h_{t+1},\theta)dF(h_{t+1}|\mathcal{I}_t,\theta) \\ &= \int f(y_{t+1}|h_{t+1},\theta)f(h_{t+1}|h_t,\theta)dF(h_t|\mathcal{I}_t,\theta). \end{aligned} \quad (22)$$

In general it is impossible to have an analytical solution for (22), instead numerical methods such as the APF method have to be employed. Suppose that we have a particle sample $\{h_t^{(i)}, i = 1,...,N\}$ of h_t from the filtered distribution of $(h_t|\mathcal{I}_t,\theta)$ with weights $\{\pi_t^{(i)}, i = 1,...,N\}$ such that $\sum_{i=1}^{N}\pi_t^{(i)} = 1$. Upon this sample, the one-step ahead predictive density of h_{t+1} is

$$f(h_{t+1}|\mathcal{I}_t,\theta) \approx \sum_{i=1}^{N}\pi_t^{(i)} f(h_{t+1}|h_t^{(i)},\theta). \quad (23)$$

Then the one-step ahead prediction distribution of h_{t+1} can be sampled and the conditional density (22) can be evaluated numerically by

$$f(y_{t+1}|\mathcal{I}_t,\theta) \approx \sum_{i=1}^{N}\pi_t^{(i)} f(y_{t+1}|h_{t+1}^{(i)},\theta), \quad (24)$$

where $h_{t+1}^{(i)}$ are particles from the prediction distribution of $(h_{t+1}|\mathcal{F}_t, \theta)$. For the assumption for the above evaluations (23) and (24) to be valid, the prediction density of h_{t+1} must be known or at least approximately. This assumption is satisfied by our procedure since from the latent AR (1) process h_{t+1} has a conditional normal distribution $h_{t+1} \sim \mathcal{N}(\mu + \phi(h_t - \mu) + \psi y_t \exp(h_t/2), \tau^2)$, which can also be used for volatility forecast.

Now the question is how to sample $(h_{t+1}|\mathcal{I}_{t+1}, \theta)$ given that we have a particle sample from the filter distribution of $(h_t|\mathcal{I}_t, \theta)$. We present an algorithm for the ASV and ASVt models based on the procedure in Chib et al. (2006).

Step 1. Given a sample $\{h_t^{(i)}, i = 1, ..., N\}$ from $(h_t|\mathcal{F}_t, \theta)$, we calculate the expectation $\hat{h}_{t+1}^{*(i)} = \mathrm{E}(h_{t+1}|h_t^{(i)})$ and

$$\pi_t^{(i)} = f(y_{t+1}|\hat{h}_{t+1}^{*(i)}, \theta), i = 1, ..., N. \tag{25}$$

Sample N times with replacement the integers of $1, ..., N$ with probability $\hat{\pi}_t^{(i)} = \pi_t^{(i)} / \sum_{i=1}^N \pi_t^{(i)}$. Define the sampled indexes $n_1, ..., n_N$ and associate these with particles $\{h_t^{(n_1)}, ..., h_t^{(n_N)}\}$.

Step 2. For each value n_i from Step 1, sample the values $\{h_{t+1}^{*(1)}, ..., h_{t+1}^{*(N)}\}$ from

$$h_{t+1}^{*(i)} = \mu + \phi(h_t^{(n_i)} - \mu) + \psi \exp(-h_t^{(n_i)}/2) y_t + \tau v_{t+1}, i = 1, ..., N. \tag{26}$$

where $v_{t+1} \sim \mathcal{N}(0, 1)$.

Step 3. Calculate the weights of the values $\{h_{t+1}^{*(1)}, ..., h_{t+1}^{*(N)}\}$ as

$$\pi_t^{*(i)} = \frac{f(y_{t+1}|h_{t+1}^{*(i)}, \theta)}{f(y_{t+1}|\hat{h}_{t+1}^{*(i)}, \theta)}, i = 1, ..., N, \tag{27}$$

and resample the values $\{h_{t+1}^{*(1)}, ..., h_{t+1}^{*(N)}\}$ N times with replacement using these weights to obtain a fair sample $\{h_{t+1}^{(1)}, ..., h_{t+1}^{(N)}\}$ with weights $1/N$ from the filter distribution of $(h_{t+1}|\mathcal{I}_{t+1}, \theta)$.

In our experience $N = 3000$ is sufficient for our simulation studies and the real stock returns data that we use to illustrate our estimation methods.

2.5. Diagnostics

There are many statistical tools that can be used to check the overall fit of the ASV and ASVt models. One such tool is the Kolmogorov-Smirnov (KS) test which assesses whether the realized observation errors come from the correspondingly assumed distribution. An alternative way is to calculate the probability transform integrals (PITs) proposed by Diebold et al. (1998).

Suppose that $\{f(y_t|\mathcal{I}_{t-1})\}_{t=1}^T$ is the sequence of conditional densities that guides the time series of y_t, and let $\{p(y_t|\mathcal{I}_{t-1})\}_{t=1}^T$ be the corresponding sequence of one-step ahead density forecasts. The PIT of y_t is defined as

$$u(t) = \int_{-\infty}^{y_t} p(z|\mathcal{I}_{t-1}) dz. \tag{28}$$

Under the null hypothesis that the sequence $\{p(y_t|\mathcal{I}_{t-1})\}_{t=1}^T$ coincides with $\{f(y_t|\mathcal{I}_{t-1})\}_{t=1}^T$, the sequence $\{u(t)\}_{t=1}^T$ is an *i.i.d.* Uniform $(0, 1)$. In our ASV model, the PITs can be calculated by the following equations

$$u(t) \approx \frac{1}{N}\sum_{i=1}^{N}\int_{-\infty}^{y_t}\frac{1}{\sqrt{2\pi}}\exp\left\{\frac{-h_t^{(i)}}{2}\right\}\exp\left\{-\frac{y_t^2\exp(-h_t^{2(i)})}{2}\right\}dz. \tag{29}$$

In our ASVt model, the PITs can be calculated by the following equations

$$u(t) \approx \frac{1}{N}\sum_{i=1}^{N}\int_{-\infty}^{y_t}\frac{\nu^{\nu/2}\Gamma((\nu+1)/2)}{\Gamma(\nu/2)\Gamma(1/2)}(\nu+z^2\exp(-h_t))^{-(\nu+1)/2}dz. \tag{30}$$

In the computation of $u(t)$, $h_t^{(i)}$ are particles from the corresponding predictive distribution of h_t with weights set equal to $1/N$.

3. Simulation studies for the ASV model

In this section, we conduct simulation studies only for the ASV model. To generate asset returns, we make a Cholesky transformation to the two innovations of the ASV model and obtain

$$y_t = \exp(h_t/2)\rho(h_{t+1}-\mu-\phi(h_t-\mu))/\sqrt{\sigma}+\exp(h_t/2)\sqrt{1-\rho^2}v_t, \tag{31}$$

$$h_{t+1} = \mu+\phi(h_t-\mu)+\sqrt{\sigma}\eta_{t+1}, \tag{32}$$

where v_t and η_{t+1} are independent and *i.i.d.* with $v_t \sim \mathcal{N}(0,1)$ and $\eta_{t+1} \sim \mathcal{N}(0,1)$. For given θ, the following equations will be used to generate $\mathbf{h}$ and $\mathbf{y}$.

$$h_{t+1} \sim \mathcal{N}(\mu+\phi(h_t-\mu),\sigma^2), \tag{33}$$

$$y_t \sim \mathcal{N}(\exp(h_t/2)\rho(h_{t+1}-\mu-\phi(h_t-\mu))/\sqrt{\sigma},\exp(h_t)(1-\rho^2)), \tag{34}$$

where $h_1 \sim \mathcal{N}(\mu,\sigma/(1-\phi^2))$ and $y_T \sim \mathcal{N}(0,\exp(h_T))$.

The parameters used to generate the asset returns are reported in the second column of Table 2. We generated 2000 observations from the ASV model. Our proposed estimation algorithm was iterated 50,000 iterations and the first 10,000 sampled points were discarded as the burn in prior to conducting a Bayesian inference. Figure 1 includes the histograms and time series of simulated values from the conditional posterior distributions of the parameters of the ASV model. These time series converge very fast indicating that a very short burn in period is necessary prior to conducting Bayesian estimation. In Figure 2 we compare the absolute returns with the estimated volatilities of the returns, while in Figure 3 we check the assumption of the model by analyzing the PITs of the fitted ASV model. We plot the scatter plot of the PITs and histogram plot of $u(t)$ which depicts the empirical distributions of the PITs. The solid lines represent the 95% confidence intervals of the uniformity, which calculation was detailed in Diebold et al. (1998). In Figure 4 we compare the cumulative distribution function (CDF) of PITs with the theoretical CDF of a uniform distribution over the interval (0,1). The KS test statistic is 0.0204 with the corresponding p-value 0.3716; so we can not reject the null hypothesis that the fitted ASV model agrees with the artificially generated asset returns. Table 2 includes summaries in terms of standard errors and Bayesian highest probability density (HPD) confidence intervals for the estimated parameters. The estimated parameters are close to their true values with relatively small standard errors.

Table 2. True and estimated parameters of the ASV model based on the simulated returns data.

Parameter	True	Est.	Std.	HPD CI(95%)
μ	**-10.45**	-10.7423	0.2749	(-11.2547, -10.1877)
ϕ	**0.98**	0.9820	0.0053	(0.9713, 0.9921)
ρ	**-0.41**	-0.4138	0.0759	(-0.5706, -0.2660)
σ	**0.19**	0.2096	0.0198	(0.1667, 0.2440)

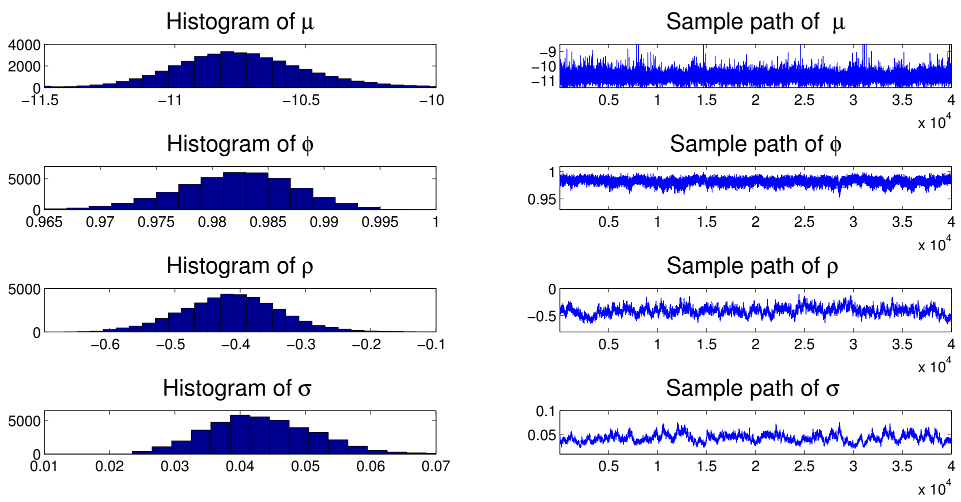

Figure 1. Histogram and time series of simulated values from the conditional posterior distributions of the parameters of the ASV model.

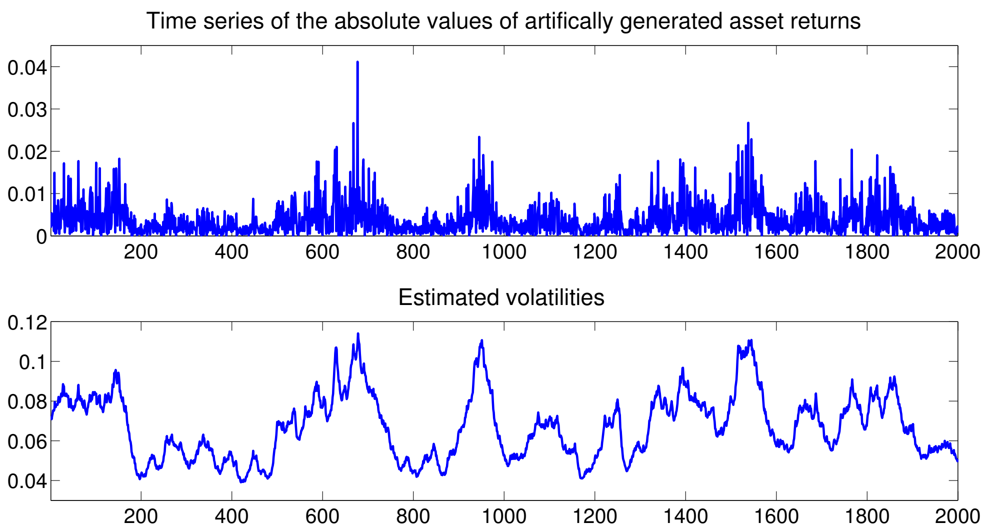

Figure 2. Comparison between the absolute returns and the estimated volatilities of the ASV model based on the generated asset returns data.

Overall the simulation studies performed in this section show that our proposed estimation methods work quite well in terms of parameter and volatility estimations of asymmetric stochastic volatility models.

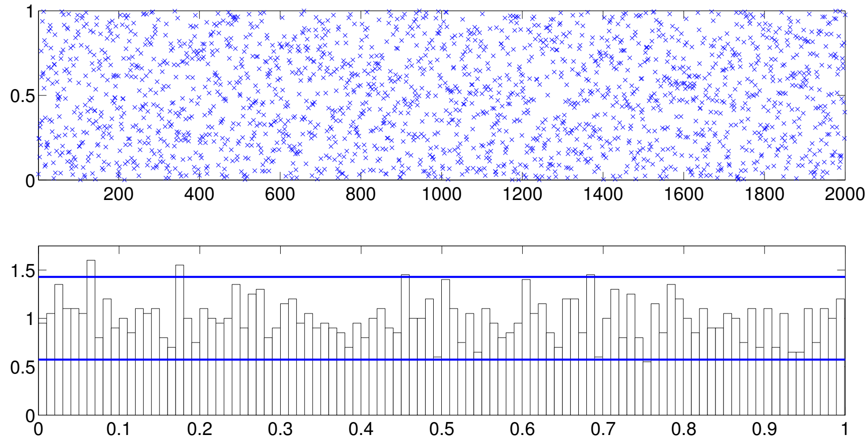

Figure 3. Analysis of the PITs from the fitted ASV model based on the simulated data. Top Panel: scatter plot of $u(t)$. Bottom Panel: histogram plot of $u(t)$ which depicts the empirical distributions of the PITs. The solid lines represent the 95% confidence intervals of the uniformity, which calculation was detailed in Diebold et al. (1998).

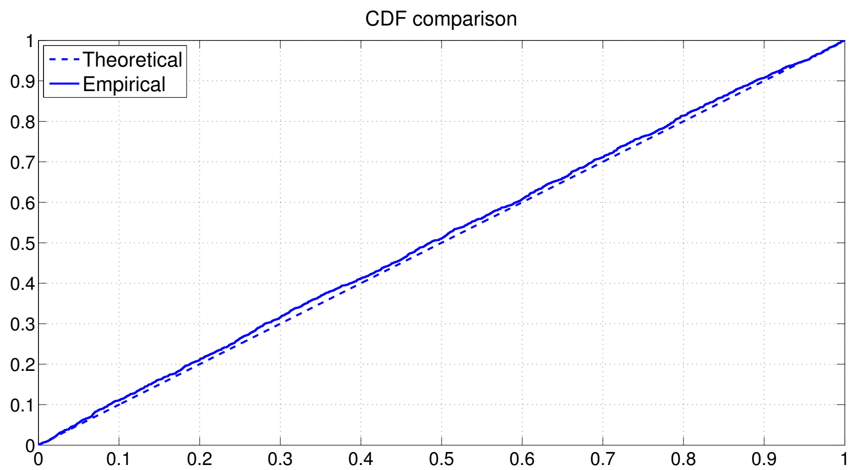

Figure 4. Comparison between the theoretical uniform CDF and the empirical CDF of the PITs from generated returns data.

4. Empirical analysis

In this section, we present empirical results of the proposed estimation methods of ASV and ASVt models to a benchmark data set of asset returns. The data set includes the daily returns of the Australian

All Ordinaries stock index.* There are totally 1508 observations covering the period from January 2, 2000 to December 30, 2005, excluding weekends and holidays. This data set is called AUX data hereafter.

4.1. Analysis of the AUX data based on the ASV model

Table 3 reports the estimated parameters from the ASV models for the the AUX data. The correlation between the two innovations of the ASV model is found to be statistically significant with $\rho = -0.6417$ indicating that there is a strong leverage between asset returns and the latent volatilities of the returns.

Table 3. Estimated parameters of the ASV models based on the AUX data.

Parameter	Est.	Std.	HPD CI(95%)
μ	-0.6297	0.1115	(-0.8421, -0.4014)
ϕ	0.9621	0.0101	(0.9411, 0.9796)
ρ	-0.6417	0.0725	(-0.7747, -0.4915)
σ	0.1907	0.0254	(0.1420, 0.2406)

To assess the goodness-of-fit of the ASV model to the data set, we calculate the PITs originated from the fitted model. Figure 5 provides scatter and histogram plots of the PITs, while in Figure 6 we compare the CDF of PITs with the theoretical CDF of a uniform distribution over the interval (0,1). The KS test statistic is 0.0235 with the corresponding p-value 0.3711. We can not reject the null hypothesis that the PITs obtained from the fitted ASV model follow a uniform distribution over the interval (0,1). Figure 7 compares the absolute values of asset returns with the estimated volatilities of the returns based on the ASV model.

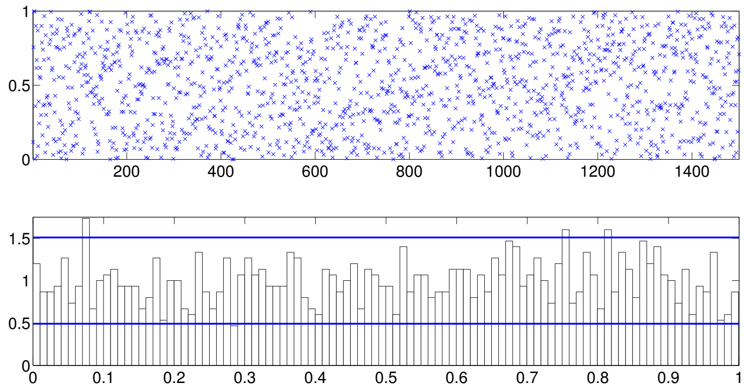

Figure 5. Goodness-of-fit test via the scatter plot (top) and the histogram (bottom) of the PITs produced by the fitted ASV model based on the AUX data. The two horizontal lines in the histogram plot are the 95% confidence intervals of the uniformity, constructed under the normal approximation of a binomial distribution, the calculation of which is detailed in Diebold et al. (1998).

*We thank Professor Xibin Zhang for kindly providing us this data set, which was analyzed in Zhang and King (2008).

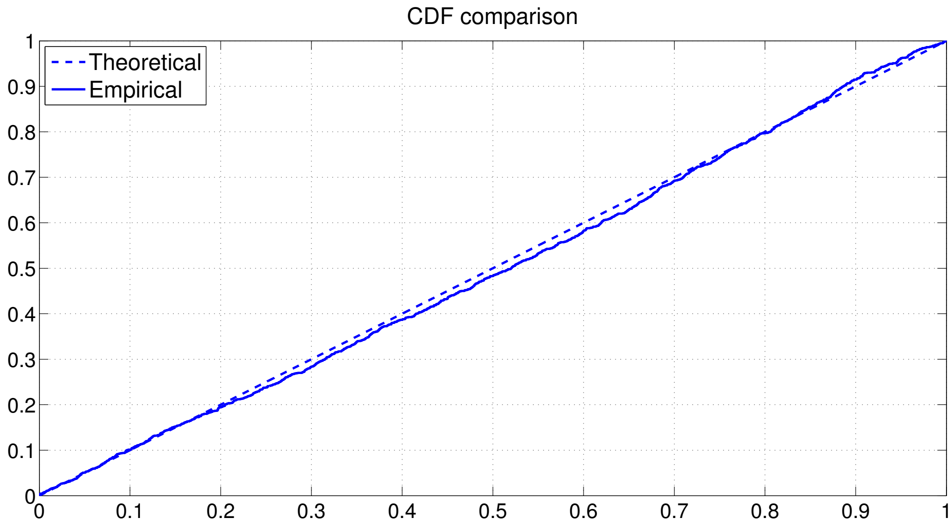

Figure 6. Comparison between the empirical CDF of the PITS produced by the fitted ASV model and the theoretical CDF of a uniform distribution over the interval (0,1) based on the AUX data.

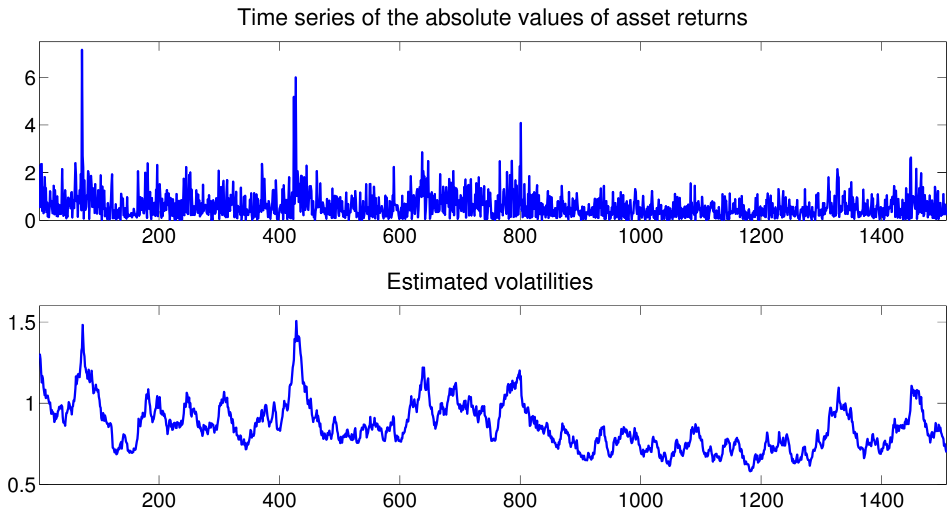

Figure 7. Comparison between the absolute returns and the estimated volatilities under the ASV model based on the AUX data.

4.2. Analysis of the AUX data based on the ASVt model

Table 4 reports the estimated parameters from the ASVt model for the AUX data. The correlation between the returns and the log volatilities is captured by $\psi = -0.7214$, which again indicates that there is a significant leverage effect between the asset return and the future volatility of the returns. Similar to the empirical analysis of the ASV model, to assess the goodness-of-fit of the ASVt models for this data set, we calculate the PITs originated from the fitted model. Figure 8 provides scatter and histogram plots of the PITs, while in Figure 9 we compare the CDF of PITs with the theoretical CDF of a uniform distribution over the interval (0,1). The KS test statistic is 0.0304 with the corresponding p-value 0.12. We can not reject the non hypothesis that the PITs obtained from the fitted ASVt model follow a uniform distribution over the interval (0,1). Figure 10 compares the absolute values of asset returns with the estimated volatilities of the returns based on the ASVt model.

Table 4. Estimated parameters of the ASVt models based on the AUX data.

Parameter	Est.	Std.	HPD CI(95%)
μ	-0.7532	0.1147	(-0.8531, -0.4104)
ϕ	0.9670	0.0102	(0.9429, 0.9818)
ψ	-0.7214	0.0774	(-0.7868, -0.4893)
τ	0.1547	0.0265	(0.1456, 0.2433)
ν	19.1603	4.4612	(10.9462, 27.9079)

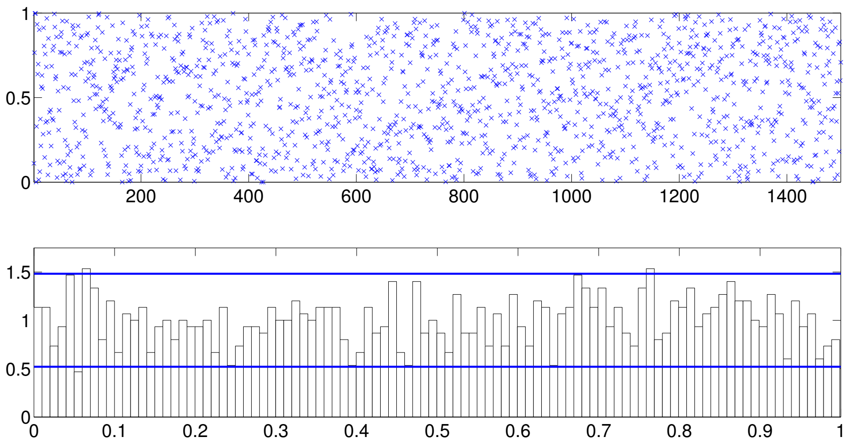

Figure 8. Goodness-of-fit test via the scatter plot (top) and the histogram (bottom) of the PITs produced by the fitted ASVt model based on the AUX data. The two horizontal lines in the histogram plot are the 95% confidence intervals of the uniformity, constructed under the normal approximation of a binomial distribution, the calculation of which is detailed in Diebold et al. (1998).

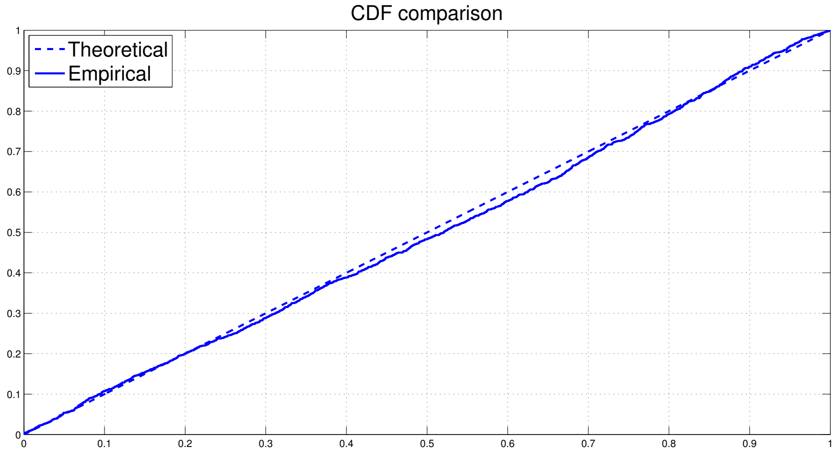

Figure 9. Comparison between the empirical CDF of the PITS produced by the fitted ASVt model and the theoretical CDF of a uniform distribution over the interval (0,1) based on the AUX data.

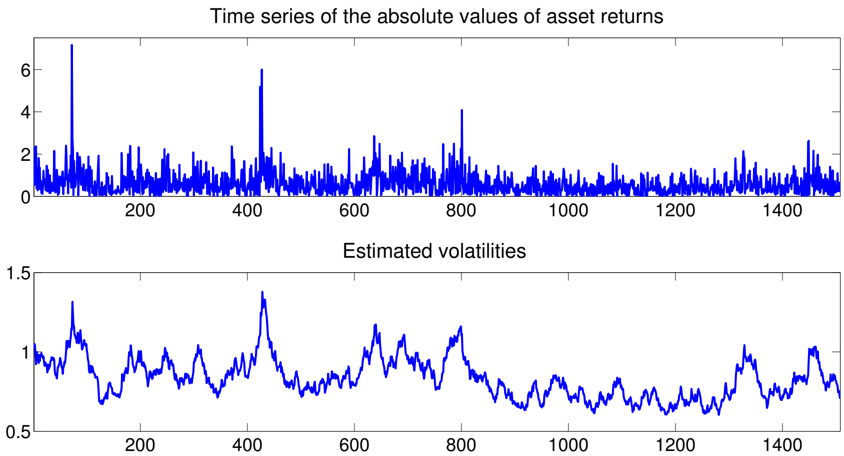

Figure 10. Comparison between the absolute returns and the estimated volatilities under the ASVt model based on the AUX data.

5. Conclusions

In this paper we have proposed using an acceptance-rejection Metrologies-Hastings method to fit asymmetric stochastic volatility models either with or without allowing for heavy tails in the return distribution. The proposal distributions are simulated by using the slice sampler. Simulation studies and empirical analysis show that our proposed estimation methods work reasonably quite well for parameter and volatility estimations of the stochastic volatility models studied in this paper.

Acknowledgments

The author, Zhongxian Men, would like to declare: The statements, views and opinions expressed in this article are my own and do not necessarily reflect those of my employer, JPMorgan Chase & Chase Co., its affiliates, other employees or clients.

Conflict of interest

The authors declare no conflict of interest.

References

Bollerslev T (1986) Generalized autoregressive conditional heteroskedasticity. *J Econometrics* 31: 307–327.

Bauwens L, Lubrano M (1998) Bayesian inference on GARCH models using the Gibbs sampler. *Economet J* 1: C23–C26.

Broto C, Ruiz E (2004) Estimation methods for stochastic volatility models: a survey. *J Econ Surv* 18: 613–649.

Carnero A, Pena D, Ruiz E (2003) Persistence and kurtosis in GARCH and stochastic volatility models *J Financ Economet* 2: 319–342.

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

Chib S, Nardarib F, Shephard N (2006) Analysis of high dimensional multivariate stochastic volatility models. *J Econometrics* 134: 341–371.

Dobigeon N, Tourneret J (2010) Bayesian orthogonal component analysis for sparse representation. *IEEE T Signal Proces* 58: 2675–2685.

Diebold FX, Guther TA, Tay AS (1998) Evaluating density forecasts with applications to financial risk management. *Int Econ Rev* 39: 863–883.

Engle RF (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* 50: 987–1007.

Eraker B, Johannes M, Polson N (2003) The Impact ofJumps inVolatility and Returns. *J Financ* 58: 1269–1300.

Geweke J (1993) Bayesian treatment of the independent Student-t linear model. *J Appl Econom* 8: S19–S40.

Harvey AC, Shephard N (1996) Estimation of an asymmetric stochastic volatility model for asset returns. *J Bus Econ Stat* 14: 42-434.

Jacquier E, Polson NG, Rossi PE (2004) Bayesian analysis of stochastic volatility models with fat-tails and correlated errors. *J Econometrics* 122: 185–212.

Kawakatsu H (2007) Numerical integration-based Gaussian mixture filters for maximum likelihood estimation of asymmetric stochastic volatility models. *Economet J* 10: 342–358.

Kim S, Shephard N, Chib S (1998) Stochastic volatility: Likelihood inference and comparison with ARCH models. *Rev Econ Stud* 65: 361–393.

Liesenfeld R, Richard J (2003) Univariate and multivariate stochastic volatility models: estimation and diagnostics. *J Empiri Financ* 45: 505–531.

Melino A, Turnbull SM (1990) Pricing foreign currencyoptions with stochastic volatility. *J Econometrics* 45: 239–265.

Men Z (2012) Bayesian Inference for Stochastic Volatility Models. Ph.D. thesis, Department of Statistics and Actuarial Science at the University of Waterloo.

Men Z, McLeish D, Kolkiewicz A, et al. (2017) Comparison of Asymmetric Stochastic Volatility Models under Different Correlation Structures. *J Appl Stat* 44: 1350–1368.

Men Z, Kolkiewicz A, Wirjanto TS (2015) Bayesian Analysis of Asymmetric Stochastic Conditional Duration Model. *J Forecasting* 34: 36–56.

Mira A, Tierney L (2002) Efficiency and Convergence Properties of Slice Samplers. *Scand J Stat* 29: 1–12.

Neal RN (2003) Slice sampling. *Annals Stat* 31: 705–767.

Omori Y, Chib S, Shephard N, et al. (2007) Stochastic volatility with leverage: Fast and efficient likelihood inference. *J Econometrics* 140: 425–449.

Pitt MK, Shephard N (1999a) Time varying covariances: A factor stochastic volatility approach. *Bayesian Stat* 6: 547–570.

Pitt M, Shephard N (1999b) Filtering via simulation: Auxiliary particle filters. *J Am Stat Assoc* 94: 590–599.

Roberts GO, Rosenthal JS (1999) Convergence of Slice Sampler Markov Chains. *J R Stat Soc B* 61: 643–660.

Shephard N, Pitt MK (1997) Likelihood Analysis of non-Gaussian Measurement Time Series. *Biometrika* 84: 653–667.

Taylor SJ (1986) Modelling Financial Time Series, Chichester: Wiley.

Wirjanto TS, Kolkiewicz A, Men Z (2016) Bayesian Analysis of a Threshold Stochastic Volatility Model. *J Forecasting* 35: 462–476.

Yu J (2005) On leverage in a stochastic volatility model. *J Econometrics* 127: 165–178.

Yu J, Meyer R (2006) Multivariate stochastic volatility models: Bayesian estimation and model comparison. *Economet Rev* 51: 2218–2231.

Zhang X, King L (2008) Box-Cox stochastic volatility models with heavy-tails and correlated errors. *J Empiri Financ* 15: 549–566.