Multiscale Stochastic Volatility Model with Heavy Tails and Leverage Effects

Zhongxian Men\textsuperscript{a}, Tony S. Wirjanto\textsuperscript{a,b,}\textsuperscript{*}, Adam W. Kolkiewicz\textsuperscript{a}

\textsuperscript{a}Department of Statistics and Actuarial Science, University of Waterloo, 200 University Avenue West, Waterloo, Ontario, Canada N2L 3G1
\textsuperscript{b}School of Accounting and Finance, University of Waterloo, 200 University Avenue West, Waterloo, Ontario, Canada N2L 3G1

July 2014

Abstract

This paper extends the multiscale stochastic volatility (MSSV) models to allow for heavy tails of the marginal distribution of the asset returns and correlation between the innovation of the mean equation and the innovations of the latent factor processes. Novel algorithms of Markov Chain Monte Carlo (MCMC) are developed to estimate parameters of these models. Results of simulation studies suggest that our proposed models and corresponding estimation methodology perform quite well. We also apply an auxiliary particle filter technique to construct one-step-ahead in-sample and out-of-sample volatility forecasts of the fitted models. In addition the models and MCMC methods are applied to data sets of asset returns from both foreign currency and equity markets.

\textit{Keywords:} Stochastic Volatility; Bayesian Inference; Markov Chain Monte Carlo; Leverage Effect; Acceptance-rejection; Slice Sampler.

\textsuperscript{*}Corresponding author

\textit{Email address:} twirjant@uwaterloo.ca (Tony S. Wirjanto)

Research support by the Society of Actuaries CAE Grant is gratefully acknowledged.
1. Introduction

There is a large volume of studies on volatility of asset returns. Since the initial work of Taylor (1986), stochastic volatility (SV) models have been heavily researched. The main feature of the canonical SV model is that the logarithm of the conditional volatility of asset returns is assumed to have a latent autoregressive process driven by a noise process from a univariate normal distribution, while the innovation of the asset return process follows a standard normal distribution. In order to accommodate the heavy tail property of the marginal distribution of asset returns, a Student-$t$ distribution is often assumed for the innovation of the asset return equation. As the SV models have a hierarchical structure, parameter estimation of these models are challenging. General methods of moments (GMM), simulated methods of moments (SMM), and important sampling methods, among others, have been proposed to estimate the parameters of the SV models. Bayesian Monte Carlo methods, in particular Markov Chain Monte Carlo (MCMC) algorithms, have become popular methods to fit the SV models. For a review of these approaches, see, for instance, Chib, Omori and Asai (2009) and Lopes and Polson (2010).

Molina, Han and Fouque (2010) propose a multiscale SV (MSSV) model to model different scales of the logarithm of the conditional volatilities of asset returns. The conditional volatility is driven by two equally weighted factors where each factor follows a first order autoregressive (AR(1)) process. The innovations of the return process and the two latent AR(1) processes are assumed to be independent and follow univariate normal distributions. Given that the marginal distribution of asset returns often appears to have heavy tails, we extend the MSSV model by assuming a Student-$t$ distribution for the innovation of the mean equation from which the heavy tail of the asset returns can be captured. This is our first contribution to the literature. Our second contribution is to assume that the innovation of the mean equation of
the model is correlated with the innovations of the latent factor processes. This is intended to capture the so-called leverage effect between the asset returns and the future volatilities. Our third contribution to the literature consists of suitable MCMC methods for parameter estimation of the model. It is worth noting that the proposed MCMC method in this paper is different from that used in Molina, Han and Fouque (2010), where the authors utilize the method advocated in Harvey and Shephard (1996) by taking the logarithm of the squared measurement equation. In this paper we use a posterior distribution of the latent states directly and the states are then simulated via a Metropolis-Hastings (MH) method where the proposed distribution is simulated by a slice sampler. The fourth contribution of this paper is our use of an auxiliary particle filter (APF) to perform one-step-ahead in-sample and out-of-sample volatility forecasts.

The remaining parts of the paper are organized as follows. Section 2 reviews the SV model and presents the multiscale asymmetric SV (MSASV) model. Then the MSASV model is extended to accommodate for heavy tails of the marginal distribution of asset returns. In particular, we assume that the innovations of the return time series has a Student-\(t\) distribution. In this paper we also introduce correlation between the innovation of the mean equation and the innovations of the latent factor processes in the model. Section 3 presents MCMC algorithms for parameter estimation of the model. Simulation studies are conducted in Section 4 to illustrate the ability that our MCMC algorithms to recover the true parameters. Empirical applications are provided in Section 5 to illustrate the performance of our model and estimation method with asset return data sets from both foreign currency and equity markets, and Section 6 concludes the paper.
2. The MSASV Model

2.1. The SV model

A canonical SV model studied in the literature is a one-factor SV model where the conditional volatility of asset returns is assumed to follow a first order autoregressive AR(1) process. The Multiscale SV model proposed by Molina, Han and Fouque (2010) is an extension of this SV model. Therefore before presenting the MSASV model, we first briefly review the one-factor SV model as a reference point.

The SV model was originally proposed by Taylor (1986) to model the time-varying volatility of asset returns. Define by $y_t$ the asset return at time $t$, the dynamics of $y_t$ is assumed to be of the following form

$$y_t = e^{h_t/2} \epsilon_t, \quad t = 1, \ldots, T,$$

(1)

$$h_{t+1} = \phi h_t + \sigma \eta_{t+1}, \quad t = 1, \ldots, T - 1,$$

(2)

$$h_0 \sim N(0, \sigma^2/(1 - \phi^2)),$$

(3)

where $\eta_t$'s are assumed to be independent variables such that $\eta_t \sim N(0, 1)$, with $N(a, b)$ denoting a normal distribution with mean $a$ and variance $b$. It is assumed that $\epsilon_t$'s are mutually independent with a common univariate normal distribution $N(0, \delta^2)$, and the innovation processes $\epsilon_t$ and $\eta_t$ are assumed to be independent. To ensure that the latent AR(1) process is weakly stationarity, we impose the condition that $|\phi| < 1$.

As the SV model is hierarchical and the mean equation defined in (1) is highly non-linear, the likelihood function of the model does not have an analytically closed-form representation, and it is highly intractable to integrate out the $T$ latent random variables from this likelihood function. In the face of this difficulty, MCMC and simulated MC maximum likelihood methods have been proposed for parameter estimation.
of SV models.

2.2. The MSASV model

The MSSV model, proposed by Molina, Han and Fouque (2010), is a natural extension of the one-factor SV model by assuming that the process of \( y_t \) is determined by multiple additive factors, which is defined as follows

\[
y_t = e^{(1'h_t/2)}\epsilon_t, \quad t = 1, 2, ..., T, \quad (4)
\]

\[
h_{t+1} = \Phi h_t + \Sigma^{1/2}\eta_{t+1}, \quad t = 1, ..., T - 1, \quad (5)
\]

\[
h_0 \sim N(0, \Omega), \quad (6)
\]

where

- The error processes, \( \epsilon_t \) and \( \eta_{t+1}, \quad t = 1, 2, \ldots \), are assumed to be mutually independent, \( \eta_t = (\eta_{1,t}, \ldots, \eta_{K,t})' \) is a vector of multivariate normal variates such that \( \eta_{t+1} \sim N(0, I_K) \), where \( I_K \) is a \( K \) by \( K \) unit matrix, and \( \epsilon_t \)'s are assumed to be mutually independent with a common univariate normal distribution, \( N(0, \delta^2) \).

- \( h_t = (h_{1,t}, ..., h_{K,t})' \) is a vector of \( K \) latent states at time \( t \), and \( 1 \) denotes a \( K \)-dimensional vector of ones.

- The innovations of the latent process \( h_t, \quad t = 1, 2, \ldots, \) are assumed to be mutually independent; that is, \( \Sigma \) is a \( K \times K \) diagonal matrix with the \( k \)-th diagonal element being equal to \( \sigma_k^2, \quad \sigma_k^2 > 0 \), and \( \Phi \) is a \( K \times K \) diagonal matrix with mean reversion parameters, such that \( |\phi_k| < 1, \) for \( k = 1, \ldots, K \).

- The covariance matrix of the initial latent variable vector \( h_0 \) is equal to the implied stationary marginal covariance matrix \( \Omega \) of the latent process, which satisfies the well-known condition that \( \Omega = \Phi \Omega \Phi + \Sigma \).
Note the similarity between this model and models that allow for jumps in volatility. If we set $\phi_i = 0$ for some $i$, the implied model reduces to a model that allows for a permanent (log-normal) source of independent jumps in the volatility series, as the $h_{i,t}$ would be (temporally independent) normals augmented to the (log)variance process driving the return series. augmented to the (log)variance process driving the return series.

Molina, Han and Fouque (2010) motivate the model in (4) - (6) as a discrete-time approximation to the underlying continuous-time SV models, where the volatility is assumed to take the form of an exponential function of a sum of multiple reverting diffusion processes (more precisely Ornstein-Uhlenbeck processes) with the mean reverting processes varying on well-separated time scales. See Appendix A. Following Molina, Han and Fouque (2010), we impose the condition that $\phi_1 > \ldots > \phi_K$ in order to ensure that the MSSV model is identifiable. Under this restriction, all of the components of the latent process in (5) are ensured to evolve on different time scales. Note that we do not include a location parameter in this process, since the innovations $\epsilon_t$ in our formulation are assumed to have a non-unit variance.

The MSSV model proposed by Molina, Han and Fouque (2010), does not allow for correlation between the innovations $\epsilon_t$ and $\eta_{t+1}$. In the equity markets, asset returns are known to have a negative correlation with their logarithms of conditional volatilities. As a result, in this paper, we allow for correlation between the innovation of the mean equation and the innovations of the latent factor processes. In principle, we can also allow for correlation between the innovations of the latent AR(1) processes. To simplify the development of the MCMC estimation algorithm, in this paper, we do not consider this possibility. Another important observation in financial time series is the heavy-tail property of the marginal distribution of the asset returns, which is often captured by assuming that the innovation of the mean equation follows a Student-
distribution. Accordingly we assume that $\epsilon_t \sim t(v)$ with $v$ degrees of freedom. The MSASV model with the Student-$t$ distributional assumption to the innovation of the mean equation is called the MSASV-$t$ model hereafter.

For the convenience of deriving the MCMC algorithm, we reparameterize the latent AR(1) process of the MSASV model as follows

$$h_{k,t+1} = \phi_k h_{k,t} + \psi_k y_t e^{-\frac{1}{2} \sum_{k=1}^{K} h_{k,t}} + \tau_k e_{k,t+1}, \quad k = 1, \ldots, K,$$

where $e_{k,t}, k = 1, \ldots, K$, are independent univariate standard normal noises, $\psi_k = \sigma_k \rho_k$ and $\tau_k = \sigma_k \sqrt{1 - \rho_k^2}, k = 1, \ldots, K$. In this reparameterized form, it is clear that we allow for the leverage effects between the innovation of the mean equation and those of the latent factor processes, as conventionally defined in the one-factor SV literature. But we do not assume correlation between the latent innovations. The latter is done for computational tractability. Given (7), instead of sampling $\rho_k$ and $\sigma_k, k = 1, \ldots, K$, we sample $\psi_k$ and $\tau_k, k = 1, \ldots, K$, and then proceed backwards to obtain samples of $\rho_k$ and $\sigma_k$.

2.3. MCMC Estimation

In the remainder of the paper, we focus on the MSASV and the MSASV-$t$ models with two factors, that is, we assume that $K = 2$, for computational tractability. Define $\theta = (\phi_1, \phi_2, \sigma_1, \sigma_2, \rho_1, \rho_2, \delta)'$ as the vector of parameters for the MSASV model, $\theta_t = (\phi_1, \phi_2, \sigma_1, \sigma_2, \rho_1, \rho_2, v)'$ as the vector of parameters of the MSASV-$t$ model, and $h = \{h_1, \ldots, h_T\}$ as the set of the corresponding latent states.

The MSASV and the MSASV-$t$ models are completed by specifying proper prior distributions for the parameters of the model. For simplicity, we assume that all prior distributions of the parameters in each of the two multiscale SV models are mutually independent. To impose stationary condition on the latent processes, the
prior distributions for $\phi_1$ and $\phi_2$ are assumed to be normal $N(0,10)$ truncated in the interval $(-1,1)$. These prior distributions result in reasonably flat densities over their support regions. In the MCMC algorithm, we sample $\sigma_i^2$ instead of $\sigma_i$, $i = 1,2$, and for this we use an inverse Gamma distribution $IG(5,0.05)$. For the prior distributions of $v$, we use the half-Cauchy prior with the probability density function

$$p(v) \propto \frac{1}{1 + v^2}, \quad v > 0. \quad (8)$$

In our implementation of the MCMC algorithm, the latent states $h$ are augmented as a vector of parameters and estimated as a by-product of the estimation process.

2.4. Estimation of the MSASV model

An outline of the MCMC algorithm is listed in Table 1. Below we provide additional explanations.

<table>
<thead>
<tr>
<th>Table 1: MCMC algorithm for the MSASV model.</th>
</tr>
</thead>
</table>

**Step 0.** Initialize $h, \phi_k, \sigma_k, k = 1,2,$ and $\delta$.

**Step 1.** Sample $h_{k,t}, k = 1,2, t = 1,...,T$.

**Step 2.** Sample $\phi_1$ and $\phi_2$.

**Step 3.** Sample $\psi_1$ and $\psi_2$.

**Step 4.** Sample $\tau_1^2$ and $\tau_2^2$.

**Step 5.** Sample $\delta^2$.

**Step 6.** Go to **Step 1**.

**Step 0.** Initialize $\phi_k, \psi_k, \sigma_k, k = 1,2,$ and $\delta$ using the corresponding prior distributions. To set an initial value of the vector $h$, we set the parameters of the latent process as $\phi_k = 0.5, \sigma_k = 0.12, k = 1,2, v = 0.5$ and $\gamma = 0.5$, and then generate the
initial value of $h$ using the definition (5)–(6) of the process.

**Step 1.** Sample $h$. The simulation is conducted via a single-move acceptance-rejection algorithm.

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 $h_{1,t}$ is

$$f(h_{1,t} | y_t, h_{t-1}, h_{t+1}, \theta)$$

$$= c_{1t} f(y_t | h_t) f(h_t | h_{t-1}, y_{t-1}, \theta) f(h_{t+1} | y_t, \theta)$$

$$= c_{2t} \exp(-h_{1,t}/2) \times \exp \left\{ - \frac{(y_t^2 \exp(-h_{1,t} - h_{2,t}))}{2\delta^2} \right\}$$

$$\times \exp \left\{ - \frac{(h_{1,t} - \phi h_{1,t-1} - \psi_1 y_{t-1} \exp(-h_{1,t-1}/2 - h_{2,t-1}/2))^2}{2\tau_1^2} \right\}$$

$$\times \exp \left\{ - \frac{(h_{2,t+1} - \phi h_{2,t} - \psi_2 y_t \exp(-h_{1,t}/2 - h_{2,t}/2))^2}{2\tau_2^2} \right\}$$

$$< c_{2t} \exp(-h_{1,t}/2) \times \exp \left\{ - \frac{(y_t^2 \exp(-h_{1,t} - h_{2,t}))}{2\delta^2} \right\}$$

$$\times \exp \left\{ - \frac{(h_{1,t} - \phi h_{1,t-1} - \psi_1 y_{t-1} \exp(-h_{1,t-1}/2 - h_{2,t-1}/2))^2}{2\tau_1^2} \right\}$$

$$\times \exp \left\{ - \frac{(h_{2,t+1} - \phi h_{2,t} - \psi_2 y_t \exp(-h_{1,t}/2 - h_{2,t}/2))^2}{2\tau_2^2} \right\}$$

(9) (10)

where $c_{1t}$ and $c_{2t}$ are the two normalizing constants. The inequality sign holds because the last two parts of the right-hand side of equation in (9) is less than 1. We also note that the dominant distribution in (10) is an unknown distribution and therefore can not be simulated directly. Instead we employ the slice sampler introduced in Neal (2003) to sample the dominant distribution. Then the full conditional of $h_{1,t}$ can be simulated by the acceptance-rejection method. This is a novel feature of our algorithm.
Algorithm of the slice sampler for $h_{1,t}$

It is easy to verify that the right-hand side of (10) can be expressed as

$$g(h_{1,t}) \propto \exp \left\{ -\frac{(y_{t}^2 \exp(-h_{1,t} - h_{2,t}))}{2\delta^2} \right\} \exp \left\{ -\frac{(h_{1,t} - \mu_{1,t})^2}{2\tau_1^2} \right\},$$

where

$$\mu_{1,t} = \frac{\tau_1^2}{2} + \phi_1 h_{1,t-1} + \psi_1 y_{t-1} \exp(-h_{1,t-1}/2 - h_{2,t-1}/2).$$

1. Draw $u_1$ uniformly from the interval $(0, 1)$. Let $u_2 = u_1 \exp \left\{ -\frac{y_{t}^2}{2\delta^2} \exp(-h_{1,t} - h_{2,t}) \right\}$. If $y_{t} \neq 0$, then we have

$$h_{1,t} \geq -\log \left( \frac{-2\delta^2 \log(u_2)}{y_{t}^2 \exp(-h_{2,t})} \right). \tag{11}$$

2. Draw $u_3$ uniformly from the interval $(0, 1)$. Let $u_4 = u_3 \exp \left\{ -\frac{(h_{1,t} - \mu_{1,t})^2}{2\tau_1^2} \right\}$ and

$$u_4 < \exp \left\{ -\frac{(h_{1,t} - \mu_{1,t})^2}{2\tau_1^2} \right\}$$

Then we have

$$\mu_{1,t} - \sqrt{-2\tau_1^2 \log(u_4)} \leq h_{1,t} \leq \mu_{1,t} + \sqrt{-2\tau_1^2 \log(u_4)}. \tag{12}$$

3. If $y_{t} \neq 0$, draw $h_{1,t}$ uniformly from the interval determined by the inequalities (11) and (12) such as

$$h_{1,t} \sim U \left( \max \left\{ -\log \left( \frac{-2\delta^2 \log(u_2)}{y_{t}^2 \exp(-h_{2,t})} \right), \mu_{1,t} - \sqrt{-2\tau_1^2 \log(u_4)} \right\}, \mu_{1,t} + \sqrt{-2\tau_1^2 \log(u_4)} \right),$$
otherwise,

\[ h_{1,t} \sim U \left( \mu_{1,t} - \sqrt{-2\tau_1^2 \log(u_4)}, \mu_{1,t} + \sqrt{-2\tau_1^2 \log(u_4)} \right). \]

The single-move simulation method is popular in the literature and used in Jacquier, Polson and Rossi (1994, 2004), Yu, Yang and Zhang (2006), Zhang and King (2008) and among others. The advantage of the slice sampler is that each iteration can give us a point from the underlying distribution unlike the MH algorithm where many generated points have to be discarded.

**Step 2.** Sampling \( \phi_k, k = 1, 2 \). Given the conjugate prior distribution \( \phi_k \sim \mathcal{N}(\alpha_{\phi_k}, \beta_{\phi_k}^2) \), the full conditional of \( \phi_k \) is

\[
 f(\phi_k|y, \psi_k, \tau_k) \propto p(h_{k,1}|\theta_{-\phi_k}) \prod_{t=1}^{T-1} p(h_{k,t+1}|h_{k,t}, \theta_{\phi_k}, y_t) \exp \left\{ -\frac{(\phi_k - \alpha_{\phi_k})^2}{2\beta_{\phi_k}^2} \right\}
\]

\[
 \propto \mathcal{N}\left(\frac{d}{c}, \frac{1}{c}\right)(1 - \phi_k^2)^{\frac{1}{2}},
\]

where

\[
 c = \frac{-h_{k,1}^2}{\sigma_k} + \sum_{t=1}^{T-1} \frac{h_{k,t}^2}{\tau_k^2} + \frac{1}{\beta_{\phi_k}^2},
\]

\[
 d = \sum_{t=1}^{T-1} h_{k,t} \left( h_{k,t+1} - \psi_k y_t \exp \left( -h_{1,t}^2/2 - h_{2,t}^2/2 \right) \right) + \frac{\alpha_{\phi_k}}{\beta_{\phi_k}^2}.
\]

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

**Step 3, 4, 5.** Sampling parameters \( \psi_k \) and \( \tau_k, k = 1, 2 \) and \( \delta \). Since the priors for these parameters are conjugate, the full conditionals are normal and inverse Gamma distributions, respectively. These full conditionals can be easily simulated. We omit these formulas and refer readers to Kim, Shephard and Chib (1998).
2.5. Estimation of the MSASV-t model

Sampling the latent states $h_{k,t}, t = 1, ..., T - 1$. The simulation of $h_{k,1}$ and $h_{k,T}$ are similar. The full conditional of $h_{1,t}, t = 2, ..., T - 1$, is

$$
f(h_{1,t}|y, h_{t-1}, h_{t+1}, \theta)
= c_{3t} f(y_t|h_t) f(h_t|h_{t-1}, y_{t-1}, \theta) f(h_t|h_{t+1}, y_t, \theta)
= c_{3t} e^{-h_{1,t}/2} \left(1 + \frac{y_t^2 e^{-h_{1,t}-h_{2,t}}}{v}\right)^{-\frac{v+1}{2}}
\times \exp \left\{-\frac{[(h_{1,t} - \phi h_{1,t-1} - \psi_1 y_{t-1} \exp(-h_{1,t-1}/2 - h_{2,t-1}/2)]^2}{2\tau_1^2}\right\}
\times \exp \left\{-\frac{[(h_{1,t+1} - \phi h_{1,t} - \psi_1 y_t \exp(-h_{1,t}/2 - h_{2,t}/2)]^2}{2\tau_1^2}\right\}
\times \exp \left\{-\frac{[(h_{2,t+1} - \phi h_{2,t} - \psi_1 y_t \exp(-h_{1,t}/2 - h_{2,t}/2)]^2}{2\tau_2^2}\right\}
\leq c_{4t} e^{-h_{1,t}/2} \left(1 + \frac{y_t^2 e^{-h_{1,t}-h_{2,t}}}{v}\right)^{-\frac{v+1}{2}}
\times \exp \left\{-\frac{[(h_{1,t} - \phi h_{1,t-1} - \psi_1 y_{t-1} \exp(-h_{1,t-1}/2 - h_{2,t-1}/2)]^2}{2\tau_1^2}\right\}
$$

where $c_{3t}$ and $c_{4t}$ are the two normalizing constants. The right-hand side of the inequality is a product of three positive functions of $h_{1,t}$ which can be sampled by the slice sampler. This is similar to the simulation of the latent states of the MSASV model, where the proposal distribution is simulated by the slice sampler.

- Sampling $v$. The full conditional of $v$ is

$$
f(v|y, h, \mu, \phi, \sigma^2) \propto f(y|h, v) f(v)
= f(v) \prod_{t=1}^{T} \frac{v^{v/2} \Gamma((v+1)/2)}{\Gamma(v/2)\Gamma(1/2)} \left(v + y_t^2 \exp(-h_{1,t} - h_{2,t})\right)^{-(v+1)/2},
$$
where \( f(v) \) is a prior density of \( v \). In the literature, there are several ways to specify this prior distribution. Jacquier, Polson and Rossi (2004) propose a discrete prior distribution \( U[3, 40] \) 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 the Cauchy prior proportional to \( 1/(1 + v^2) \). In our procedure we adapt the normal prior. Since this full conditional is an unknown distribution, we use a random-walk MH algorithm, in which the proposal density is the standard Gaussian density and the acceptance probability is computed using equation (18).

3. Model Selection and Its Assessment

3.1. Auxiliary particle filter

To perform model comparison, we need to evaluate the model likelihood. For the MSASV and the MSASV-t models proposed in this paper, the likelihood is unfortunately difficult to compute analytically due to its highly non-linear structure. Therefore, to perform this task we employ the auxiliary particle filter (APF) proposed by Pitt and Shephard (1999). This is an efficient recursive algorithm that approximates the filter and the one-step-ahead predictive distributions of the latent states of the model. By successive conditioning, we can represent the sample likelihood of the multiscale model as

\[
f(y|\theta) = f(y_1|\theta) \prod_{t=2}^{T} f(y_t|\mathcal{I}_{t-1}, \theta), \tag{19}
\]
where $\mathcal{I}_t$ represents the information known at time $t$. The conditional density of $y_{t+1}$ given $\theta$ and $\mathcal{I}_t$ has the following representation:

$$f(y_{t+1}|\mathcal{I}_t, \theta) = \int f(y_{t+1}|h_{t+1}, \theta)dF(h_{t+1}|\mathcal{I}_t, \theta)$$

$$= \int f(y_{t+1}|h_{t+1}, \theta)f(h_{t+1}|h_t, \theta)dF(h_t|\mathcal{I}_t, \theta).$$

(20)

Suppose that we have a particle sample $\{h_t^{(i)}, i = 1, \ldots, N\}$ from the filtered distribution of $(h_t|\mathcal{I}_t, \theta)$, with weights $\{\pi_{it}, i = 1, \ldots, N\}$ such that $\sum_{i=1}^{N} \pi_{it} = 1$. Given this sample, the one-step-ahead approximation of the predictive density of $h_{t+1}$ is given by

$$f(h_{t+1}|\mathcal{I}_t, \theta) \approx f_A(h_{t+1}|\mathcal{I}_t, \theta) := \sum_{i=1}^{N} \pi_{it}f(h_{t+1}|h_t^{(i)}, \theta).$$

(21)

If we denote a sample drawn from the distribution of $f(h_{t+1}|\mathcal{I}_t, \theta)$ by $h_t^{(i)}$, $i = 1, 2, \ldots, N$, then the conditional density (20) can be approximated as

$$f(y_{t+1}|\mathcal{I}_t, \theta) \approx \sum_{i=1}^{N} \pi_{it}f(y_{t+1}|h_t^{(i)}, \theta),$$

(22)

For the approximation (21) to be feasible, the predictive density of $h_{t+1}$ must be known. This assumption is satisfied in our context, since the assumed form of the latent process implies that $h_{t+1}$ conditional on $h_t$ has a bivariate normal distribution $\mathcal{N}(\Phi h_t, \Sigma)$ with $\Phi = diag(\phi_1, \phi_2)$ and $\Sigma = diag(\sigma_1^2, \sigma_2^2)$. This fact is also used when we conduct the one-step ahead forecasts of the volatility. A detailed procedure of APF to calculate (20) and (22) is omitted in this paper. We refer readers to Chib, Nardari and Shephard (2002, 2006).
3.2. Diagnostics

There are a number of tools that can be used to assess the goodness-of-fit of our MSASV models. One of them is the Kolmogorov-Smirnov (KS) test, which is designed to examine whether realized observation errors originated from the assumed distribution. Another approach is to use the Probability Integral Transforms (PITs) proposed in Diebold, Guther and Tay (1998).

Suppose that \( \{f(y_t|\mathcal{F}_{t-1})\}_{t=1}^T \) is a sequence of conditional densities of \( y_t \) given the information \( \mathcal{F}_{t-1} \) we have at time \( t - 1 \), and \( \{p(y_t|\mathcal{F}_{t-1})\}_{t=1}^T \) is the corresponding sequence of one-step-ahead density forecasts. The PIT corresponding to an observed value of \( y_t \) is defined as

\[
    u(t) = \int_{-\infty}^{y_t} p(z|\mathcal{F}_{t-1})dz.
\]

(23)

Under a null hypothesis that the sequence \( \{p(y_t|\mathcal{F}_{t-1})\}_{t=1}^T \) coincides with \( \{f(y_t|\mathcal{F}_{t-1})\}_{t=1}^T \), the sequence \( \{u(t)\}_{t=1}^T \) corresponds to independent and identically distributed (i.i.d.) observations from the uniform distribution on the interval \((0, 1)\).

3.3. Model selection

There are several ways to conduct a model selection. The AIC proposed by Akaike (1987) and the BIC proposed by Schwarz (1978) are commonly used to discriminate different versions of the fitted SV models. However both the AIC and BIC require the knowledge of the exact number of independent parameters in the model. This requirement is not satisfied in the estimation approach that we take in this paper, since the latent states are augmented as parameters. Due to the fact that the states are typically highly correlated, strictly speaking, it is not appropriate to treat them as independent parameters. This is an important impediment to using either the AIC or the BIC for model selection in the context of the MSASV models. Motivated by
this concern, a new criterion for model comparison, called DIC, was introduced by Spiegelhalter, Best, Carli, and Van der Linde (2002). This criterion has proved to be particularly useful for hierarchical models such as the SV models. Berg, Meyer and Yu (2004) use this criterion for model comparison of univariate SV models.

The DIC is defined as

$$\text{DIC} = \bar{D} + P_D.$$ 

The first term $\bar{D}$ is a Bayesian measure representing a model fit, which is defined as the posterior mean of the deviance

$$\bar{D}(\theta) = E_{\theta|y}[D(\theta)],$$

where $D(\theta) = -2\log f(y|\theta)$. Larger values of $\bar{D}$ signify deterioration of the model fit. The second term, $P_D$, is defined as

$$P_D = \bar{D} - D(\bar{\theta})$$

$$= E_{\theta|y}[D(\theta)] - D(E_{\theta|y}[\theta]),$$

where $D(\bar{\theta})$ is the deviance of the posterior mean, which measures the complexity of the model. In other words, $P_D$ is the difference between the posterior mean of the deviance and the deviance under the posterior mean of $\theta$. The larger the value of $P_D$, the easier it is for the model to fit the data. The term $P_D$ is called the effective number of parameters. Since the likelihood is analytically intractable in the case of the MSASV models, to compute DIC we resort to numerical methods to evaluate $\bar{D}$ and $D(\bar{\theta})$. In this paper, we use the MCMC outputs to calculate $\bar{D}$ and $D(\bar{\theta})$. As the true value of $\theta$ is unknown, the Bayesian estimate $\hat{\theta}$ of $\theta$ is used instead.
4. Simulation Studies

In this section, we present results of our simulation studies for the MSASV model where the error process of the asset-return equation follows a univariate normal distribution. As the simulation studies for the MSASV-t models produce qualitatively similar results, we do not present these results in this section. After the model has been fitted, we use the KS test to check whether the fitted model agrees with the simulated asset returns. For a given $\theta$, the following equations are used to generate the asset-return time series $y$ and the states $h$:

$$h_{k,t+1} \sim N(\phi_k h_{k,t}, \sigma_k^2), \quad k = 1, 2,$$

$$y_t \sim e^{(h_{1,t}+h_{2,t})/2} \epsilon_t,$$  \hspace{1cm} (25)

where $h_{k,0} \sim N(\sigma_k^2/(1-\phi_k^2))$, $y_T \sim e^{(h_{1,T}+h_{2,T})/2} \epsilon_T$ and $\epsilon_t \sim N(0, \sigma^2)$.

The parameter values used to generate asset returns are presented in the second column of Table 2. We generate 12,000 observations from the MSASV model, in which the first 10,000 observations are fitted by the MSASV model and the other 2,000 observations are reserved for comparison with the one-step-ahead out-of-sample forecasted asset volatilities. Our proposed estimation algorithm is iterated 200,000 iterations and the initial 100,000 sampled points are discarded as the burn-in before Bayesian inference is drawn. In Table 2 we present the estimated parameters together with the Bayesian highest probability density (HPD) confidence intervals and standard deviations. It can be seen that the estimated parameter values are close to the corresponding true values.

The overall model fit is assessed through the analysis of the PITs obtained from the fitted MSASV model. The uniform distribution of $u(t)$ on the interval (0,1) is shown in Figure 1 by means of the scatter plot and the histogram. The two horizontal
Table 2: True and estimated parameters of the MSASV model based on simulated asset returns.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>True</th>
<th>Est.</th>
<th>Std.</th>
<th>HPD CI(95%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi_1$</td>
<td>0.95</td>
<td>0.9799</td>
<td>0.0038</td>
<td>(0.9729, 0.9872)</td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>-0.20</td>
<td>-0.1538</td>
<td>0.0581</td>
<td>(-0.2632, -0.0334)</td>
</tr>
<tr>
<td>$\sigma_1$</td>
<td>0.20</td>
<td>0.1958</td>
<td>0.0177</td>
<td>(0.1630, 0.2260)</td>
</tr>
<tr>
<td>$\phi_2$</td>
<td>0.60</td>
<td>0.5504</td>
<td>0.0344</td>
<td>(0.4834, 0.6168)</td>
</tr>
<tr>
<td>$\rho_2$</td>
<td>0.20</td>
<td>0.2202</td>
<td>0.0273</td>
<td>(0.1672, 0.2736)</td>
</tr>
<tr>
<td>$\sigma_2$</td>
<td>0.80</td>
<td>0.7852</td>
<td>0.0268</td>
<td>(0.7325, 0.8382)</td>
</tr>
<tr>
<td>$\delta$</td>
<td>0.25</td>
<td>0.2828</td>
<td>0.0124</td>
<td>(0.2593, 0.3092)</td>
</tr>
</tbody>
</table>

lines in the histogram plot are the 95% Bayesian confidence bands, whose calculation was detailed in Diebold, Guther and Tay (1998). The KS test statistic is calculated at 0.0091 with the corresponding $p$-value of 0.3805. Based on these values, we do not reject the null hypothesis that the PITs are uniformly distributed over the interval (0,1) at the 5% significance level. In Figure 2 the empirical cumulative distribution function (CDF) of the PITs is depicted together with the theoretical CDF of the uniform distribution $U(0,1)$. The graph confirms our earlier assessment that the fitted MSASV model agrees well with the simulated return data. From the above comparisons and the result of the KS test, we conclude that the proposed MCMC method for the MSASV model fits the simulated return data well.

Once the MSASV model has been estimated, we use the fitted model to perform both the in-sample and out-of-sample one-step ahead volatility forecasts. In Figure 3 we compare the absolute simulated returns with the estimated and one-step-ahead in-sample and out-sample forecasted volatilities, where the latter is separated by a vertical dotted line at $t = 10,000$. The forecasted volatilities resemble closely the true time series of the absolute simulated returns. The time series of the estimated two factors are compared with the absolute simulated returns in Figure 4.

Simulation studies show that our proposed MSASV model and the developed
Figure 1: Analysis of the PITs from the MSASV model based on the simulated return data. The top panel shows the scatter plot of $u(t)$ while the bottom panel shows the histogram of $u(t)$.

Figure 2: Comparison between the CDF of the uniform distribution $U(0, 1)$ and the empirical CDF of the PITs from the MSASV model based on the simulated return data.
Figure 3: Comparison between the absolute returns and the one-step-ahead forecasted volatilities under the MSASV model based on the simulated return data.

Figure 4: Time series of the absolute returns (first panel). Posterior mean of $(h_{1t} + h_{2t})/2$ (second panel). Posterior mean of slow mean reverting of $(h_{1t})/2$ (third panel) and the posterior mean of fast mean reverting of $(h_{2t})/2$ (fourth panel) based on the simulated return data.
MCMC algorithm work well in terms of parameter estimation of the model and are able to capture the two factors that determine the dynamics of the simulated returns.

5. Empirical Analysis

5.1. Data analysis with the MSASV model

In this section, we apply the proposed MSASV model and the developed MCMC algorithm to two benchmark data sets of asset returns, one from the exchange market and another from the equity market. The first data set consists of 945 observations on daily pound/dollar exchange rate from 01/10/1981 to 28/06/1985, called EXC hereafter. This data set is used to enable a more meaningful comparison of our results with the empirical evidence presented in Molina, Han and Fouque (2010), who use returns from the foreign currency markets. This particular data set from the exchange market has been analyzed in Harvey, Ruiz and Shephard (1994), Shephard and Pitt (1997), Meyer and Yu (2000), Skaug and Yu (2008), and Yu (2011), respectively. Since unfortunately there are not many observations contained in this data set, we fit all of the available observations by the proposed MSASV model and compare only in-sample forecasted volatilities with the estimated and the absolute observed returns.

The second data set includes the daily returns of the Australian All Ordinaries stock index, called AUX in short hereafter. The data set contains 1508 observations from January 2, 2000 to December 30, 2005, excluding weekends and holidays. For a comparison purpose, the first 1400 observations are fitted by the proposed MSASV model, and the remaining 108 observations are used for comparison with the estimated and forecasted volatilities.

Table 3 lists the estimated parameters of the MSASV model fitted to the EXC

---

1 We thank Professor Xibin Zhang for kindly providing us this data, which was analyzed in Zhang and King (2008).
data. Bayesian confidence intervals with standard deviations are also provided in this table. With relatively small standard errors, the confidence intervals contain the parameter estimates of the model. It is worth noting that the leverage effect is estimated with an incorrect expected sign for the second factor. However the leverage effect in both factors are estimated very imprecisely. This is consistent with the findings in the one-factor SV literature that the leverage effect is not a prominent feature of the returns in the foreign currency markets.

Table 3: Estimated parameters of the MSASV model based on the EXC data.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Est.</th>
<th>Std.</th>
<th>HPD CI(95%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi_1$</td>
<td>0.9766</td>
<td>0.0126</td>
<td>(0.9529, 0.9987)</td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>-0.1237</td>
<td>0.1868</td>
<td>(-0.5034, 0.2163)</td>
</tr>
<tr>
<td>$\sigma_1$</td>
<td>0.1623</td>
<td>0.0349</td>
<td>(0.0954, 0.2313)</td>
</tr>
<tr>
<td>$\phi_2$</td>
<td>0.1477</td>
<td>0.3676</td>
<td>(-0.6060, 0.7778)</td>
</tr>
<tr>
<td>$\rho_2$</td>
<td>0.1567</td>
<td>0.1996</td>
<td>(-0.2424, 0.5480)</td>
</tr>
<tr>
<td>$\sigma_2$</td>
<td>0.3162</td>
<td>0.0917</td>
<td>(0.1564, 0.4961)</td>
</tr>
<tr>
<td>$\delta$</td>
<td>0.5938</td>
<td>0.0435</td>
<td>(0.5140, 0.6829)</td>
</tr>
</tbody>
</table>

The overall model fit is assessed through the analysis of the PITs obtained from the fitted MSASV model. The uniform distribution of $u(t)$ on the interval (0,1) is visualized in Figure 5 through the scatter plot and the histogram. As the sample size of the PITs is relatively small, the Bayesian confidence bands of the PITs is relatively much wider as expected. The KS test statistic is recorded at 0.0165 with a $p$-value of 0.9557. Based on these values, we do not reject the null hypothesis that the PITs are uniformly distributed over the interval (0,1) at the conventional significance levels. In Figure 6 the empirical CDF of the PITs is depicted together with the theoretical CDF of the $U(0,1)$. The graph is consistent with our earlier finding that the fitted MSASV model compares favorably with the returns from the foreign currency market. From the above comparisons and the result of the KS test, we conclude that the proposed
MCMC method for the MSASV model fits the returns well.

Figure 5: Analysis of the PITs from the MSASV model based on EXC data. The top panel shows the scatter plot of \( u(t) \) while the bottom panel shows the histogram of \( u(t) \).

Figure 6: Comparison between the CDF of the uniform distribution \( U(0, 1) \) and the empirical CDF of the PITs from the MSASV model based on the EXC data.

In Figure 7 we compare the absolute observed returns with the estimated volatilities and the one-step-ahead in-sample forecasted volatilities. The fitted and forecasted volatilities appear to track closely the true time series of the absolute asset returns. The time series of the estimated two factors are compared with the absolute observed returns in Figure 8.
Figure 7: Comparison between the absolute returns and the one-step-ahead forecasted volatilities under the MSASV model based on the EXC data.

Figure 8: Time series of the absolute returns (first panel). Posterior mean of $(h_{1t} + h_{2t})/2$ (second panel). Posterior mean of slow mean reverting of $(h_{1t})/2$ (third panel) and the posterior mean of fast mean reverting of $(h_{2t})/2$ (fourth panel) based on the EXC data.
Next, we perform the same analysis on the AUX returns data. Table 4 lists the estimated parameters of the MSASV model fitted to the AUX data. Bayesian confidence intervals with standard deviations are also provided in this table. Again with relatively small standard errors, the parameter estimates of the model are included in the constructed confidence bands. The leverage effect in both factors is estimated with the correct expected sign. Its estimate for the first factor is quantitatively large and statistically highly significant, while that for the second factor is quantitatively small and statistically not significant.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Est.</th>
<th>Std.</th>
<th>HPD CI(95%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi_1$</td>
<td>0.9659</td>
<td>0.0101</td>
<td>(0.9460, 0.9840)</td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>-0.7171</td>
<td>0.0849</td>
<td>(-0.8636, -0.5437)</td>
</tr>
<tr>
<td>$\sigma_1$</td>
<td>0.1707</td>
<td>0.0287</td>
<td>(0.1195, 0.2271)</td>
</tr>
<tr>
<td>$\phi_2$</td>
<td>0.1785</td>
<td>0.3391</td>
<td>(-0.5074, 0.7824)</td>
</tr>
<tr>
<td>$\rho_2$</td>
<td>-0.0633</td>
<td>0.1682</td>
<td>(-0.4046, 0.2766)</td>
</tr>
<tr>
<td>$\sigma_2$</td>
<td>0.3218</td>
<td>0.0829</td>
<td>(0.1581, 0.4756)</td>
</tr>
<tr>
<td>$\nu$</td>
<td>0.6997</td>
<td>0.0325</td>
<td>(0.6354, 0.7656)</td>
</tr>
</tbody>
</table>

As before the overall model fit is assessed through the analysis of the PITs obtained from the fitted MSASV model. The uniform distribution of $u(t)$ on the interval (0,1) is displayed in Figure 9 through the scatter plot and histogram. The KS test statistic is calculated at 0.0259 with a $p$-value of 0.2979. Based on these values, we do not reject the null hypothesis that the PITs are uniformly distributed over the interval (0,1) at the conventional significance levels. In Figure 10 the empirical CDF of the PITs is depicted together with the theoretical CDF of the Uniform (0,1). The graph supports our earlier conclusion that the fitted MSASV model agrees well with the AUX returns data.

Once the MSASV model has been estimated, as before we use the fitted model to
Figure 9: Analysis of the PITs from the MSASV model based on AUX data. The top panel shows the scatter plot of $u(t)$ while the bottom panel shows the histogram of $u(t)$.

Figure 10: Comparison between the CDF of the uniform distribution $U(0,1)$ and the empirical CDF of the PITs from the MSASV model based on the AUX data.
perform in-sample one-step ahead forecasts. In Figure 11 we compare the absolute observed returns with the estimated and one-step-ahead in-sample and out-sample forecasted volatilities, where the latter is separated by a vertical dotted line at \( t = 1400 \). The forecasted volatilities appear to resemble closely the true time series of the absolute observed returns. The time series of the estimated two factors are compared with the absolute observed returns in Figure 12.

![Time series of the observed returns](image)

![Estimated volatility time series from MCMC](image)

![One−step−ahead forecasted volatility time series](image)

Figure 11: Comparison between the absolute returns and the one-step-ahead forecasted volatilities under the MSASV model based on the AUX data.

Next we compare the proposed MSASV model with the one-factor asymmetric SV (ASV) model where correlation is permitted between the innovation of the asset returns and the innovation of the latent factor process. The two data sets are also fitted by the one-factor ASV model. Table 5 lists the values of \( \bar{D} \), \( P_D \) and DIC calculated based on the fitted MSASV and ASV models. Based on the calculated DIC values, we conclude that the MSASV model fits the two data sets better and suggests evidence of at least two latent factors in the asset returns dynamics studied in this paper.
Figure 12: Time series of the absolute returns (first panel). Posterior mean of \((h_{1t} + h_{2t})/2\) (second panel). Posterior mean of slow mean reverting of \((h_{1t})/2\) (third panel) and the posterior mean of fast mean reverting of \((h_{2t})/2\) (fourth panel) based on the AUX data.

<table>
<thead>
<tr>
<th>Table 5: Model selection for the two data sets.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: MSASV model</strong></td>
</tr>
<tr>
<td>Criterion</td>
</tr>
<tr>
<td>(\bar{D})</td>
</tr>
<tr>
<td>(P_D)</td>
</tr>
<tr>
<td>DIC</td>
</tr>
<tr>
<td><strong>Panel B: ASV model</strong></td>
</tr>
<tr>
<td>Criterion</td>
</tr>
<tr>
<td>(\bar{D})</td>
</tr>
<tr>
<td>(P_D)</td>
</tr>
<tr>
<td>DIC</td>
</tr>
</tbody>
</table>
5.2. Data analysis with the MSASV-t model

We now consider using the heavy-tailed multiscale SV models to analyze the two sets of the asset returns investigated in subsection 5.1. Table 6 includes the estimated parameters of the MSASV-t model to the EXC data set with the standard deviations and the 95% Bayesian confidence intervals. Estimates of the leverage effect in both factors are quantitatively small and statistically not significant. This reinforces the findings in the one-factor SV literature that the leverage effect is not prominent for the returns in the foreign exchange markets.

Table 6: Estimated parameters of the MSASV-t model based on the EXC data.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Est.</th>
<th>Std.</th>
<th>HPD CI(95%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi_1$</td>
<td>0.9958</td>
<td>0.0030</td>
<td>(0.9900, 0.9999)</td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>-0.0254</td>
<td>0.0195</td>
<td>(-0.0619, 0.0141)</td>
</tr>
<tr>
<td>$\sigma_1$</td>
<td>0.1085</td>
<td>0.0267</td>
<td>(0.0639, 0.1624)</td>
</tr>
<tr>
<td>$\phi_2$</td>
<td>0.2621</td>
<td>0.3540</td>
<td>(-0.3741, 0.8909)</td>
</tr>
<tr>
<td>$\rho_2$</td>
<td>0.0498</td>
<td>0.0495</td>
<td>(-0.0495, 0.1437)</td>
</tr>
<tr>
<td>$\sigma_2$</td>
<td>0.2492</td>
<td>0.0847</td>
<td>(0.1113, 0.4137)</td>
</tr>
<tr>
<td>$v$</td>
<td>26.8085</td>
<td>7.5242</td>
<td>(13.8659, 39.9615)</td>
</tr>
</tbody>
</table>

The assessment of the model fit to the data is checked by assessing PITs originated from the fitted MSASV-t model. The uniform distribution of $u(t)$ on the interval $(0,1)$ is again visualized in Figure 13 by means of the scatter plot and the histogram. The KS test statistic is recorded at 0.0283 with a $p$-value of 0.4294. Based on these values, we do not reject the null hypothesis that the PITs are uniformly distributed over the interval $(0,1)$ at the conventional significance levels. In Figure 14 the empirical CDF of the PITs is plotted together with the theoretical CDF of the Uniform $(0,1)$. The graph supports our earlier assessment that the fitted MSASV-t model compares favorably with the simulated return data.

In Figure 15 we compare the absolute observed returns with the estimated and
Figure 13: Analysis of the PITs from the MSASV-t model based on EXC data. The top panel shows the scatter plot of $u(t)$ while the bottom panel shows the histogram of $u(t)$.

Figure 14: Comparison between the CDF of the uniform distribution $U(0,1)$ and the empirical CDF of the PITs from the MSASV-t model based on the EXC data.
forecasted volatilities. The fitted and forecasted volatilities appear to track closely the absolute values of the observed asset returns. The time series of the estimated two factors are compared with the absolute observed returns in Figure 16.

![Graphs showing time series of observed returns, estimated volatility time series from MCMC, and one-step-ahead forecasted volatility time series.](image)

**Figure 15:** Comparison between the absolute returns and the one-step-ahead forecasted volatilities under the MSASV-t model based on the EXC data.

For the AUX return data, the estimated parameters, their confidence intervals and related standard deviations are presented in Table 7. The leverage effect in both factors in the MSASV-t model is estimated with the correct expected sign, and is quantitatively large and statistically highly significant. This suggests that the leverage effect is a prominent feature of the returns in the equity markets, in keeping with much of the findings in the one-factor SV literature.

The overall model fit assessment is conducted by the test of the PITs calculated from the fitted model. The uniform distribution of $u(t)$ on the interval $(0,1)$ is visualized in Figure 17 through the scatter plot and the histogram. The KS test statistic is calculated at 0.0362 with a $p$-value of 0.4867. Based on these values, we do not reject the null hypothesis that the PITs are uniformly distributed over the interval $(0,1)$ at the conventional significance levels. In Figure 18 the empirical CDF of the
Figure 16: Time series of the absolute returns (first panel). Posterior mean of \((h_{1t} + h_{2t})/2\) (second panel). Posterior mean of slow mean reverting of \(h_{1t}/2\) (third panel) and the posterior mean of fast mean reverting of \(h_{2t}/2\) (fourth panel) based on the MSASV-t model for the EXC data.

Table 7: Estimated parameters of the MSASV-t model based on the AUX data.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Est.</th>
<th>Std.</th>
<th>HPD CI(95%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\phi_1)</td>
<td>0.9914</td>
<td>0.0034</td>
<td>(0.9845, 0.9976)</td>
</tr>
<tr>
<td>(\rho_1)</td>
<td>-0.6085</td>
<td>0.0888</td>
<td>(-0.7762, -0.4344)</td>
</tr>
<tr>
<td>(\sigma_1)</td>
<td>0.1220</td>
<td>0.0190</td>
<td>(0.0854, 0.1596)</td>
</tr>
<tr>
<td>(\phi_2)</td>
<td>0.3320</td>
<td>0.3615</td>
<td>(-0.3887, 0.8987)</td>
</tr>
<tr>
<td>(\rho_2)</td>
<td>-0.3162</td>
<td>0.1909</td>
<td>(-0.6933, 0.0468)</td>
</tr>
<tr>
<td>(\sigma_2)</td>
<td>0.2376</td>
<td>0.0759</td>
<td>(0.1057, 0.3859)</td>
</tr>
<tr>
<td>(v)</td>
<td>24.8652</td>
<td>5.9030</td>
<td>(14.5099, 34.9719)</td>
</tr>
</tbody>
</table>
PITs is depicted together with the theoretical CDF of the Uniform (0,1). The graph reinforces our earlier assessment that the fitted MSASV-t model agrees extremely well with the asset returns data.

Figure 17: Analysis of the PITs from the MSASV-t model based on AUX data. The top panel shows the scatter plot of $u(t)$ while the bottom panel shows the histogram of $u(t)$.

Figure 18: Comparison between the CDF of the uniform distribution $U(0,1)$ and the empirical CDF of the PITs from the MSASV-t model based on the AUX data.

Figure 19 compares the absolute observed returns with the estimated and one-step-ahead in-sample and out-of-sample forecasted volatilities. The forecasted volatilities appear to track very closely the true time series of the absolute returns. The time series of the estimated two factors are compared with the absolute observed returns in Figure 20.
Figure 19: Comparison between the absolute returns and the one-step-ahead forecasted volatilities under the MSASV-t model based on the AUX data.

Figure 20: Time series of the absolute returns (first panel). Posterior mean of \((h_{1t} + h_{2t})/2\) (second panel). Posterior mean of slow mean reverting of \(h_{1t}/2\) (third panel) and the posterior mean of fast mean reverting of \(h_{2t}/2\) (fourth panel) based on the MSASV-t model for the AUX data.
As we did previously in section 5.1, we compare the proposed MSASV-t model with the heavy-tailed one-factor asymmetric SV (ASV-t) model where the innovation of asset returns is assumed to have a Student-\(t\) distribution and is correlated with the innovation of the latent factor process. The two data sets are also fitted by the heavy-tailed one-factor ASV-t model. Table 8 lists the values of \(\bar{D}\), \(P_D\) and DIC calculated based on the fitted MSASV-t and heavy-tailed one-factor ASV-t models. It is observed that the MSASV-t model fits the EXC data better than the heavy-tailed one factor ASV-t model, while for the AUX data, the heavy-tailed one-factor ASV-t model does a better job. The estimated degrees of freedom in the fitted MSASV-t and the heavy-tailed ASV-t models are greater than 20, so that the variances of the two Student-\(t\) distributions are very close to unity. Furthermore, the DIC values calculated for the two fitted heavy-tailed models are quite similar in magnitude indicating that the AUX data can be fitted equally well by the two heavy-tailed factor models.

<table>
<thead>
<tr>
<th>Table 8: Model selection for the two sets.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: MSASV-t model</td>
</tr>
<tr>
<td>Criterion</td>
</tr>
<tr>
<td>(\bar{D})</td>
</tr>
<tr>
<td>(P_D)</td>
</tr>
<tr>
<td>DIC</td>
</tr>
<tr>
<td>Panel B: ASV-t model</td>
</tr>
<tr>
<td>Criterion</td>
</tr>
<tr>
<td>(\bar{D})</td>
</tr>
<tr>
<td>(P_D)</td>
</tr>
<tr>
<td>DIC</td>
</tr>
</tbody>
</table>

6. Conclusion

We have proposed a new variant of multiscale SV models to model the dynamics of asset returns. The logarithm of conditional volatilities of asset returns is described
by latent AR(1) processes with different scales. In order for the proposed model to accommodate heavy tails in the marginal distribution of asset returns, the innovation of asset returns is assumed to follow a student-$t$ distribution. Suitable MCMC methods have been developed for parameter estimation of the model. An auxiliary particle filter is employed to approximate the filtering and prediction distributions of the latent states of the model when we calculate model likelihood and perform volatility forecasting. In this paper, we allow for correlation between the innovation of the mean equation and those of the latent factor processes, which loosely defines the leverage effect much in keeping with the one-factor SV literature. But we do not allow for correlation between the innovations of latent AR(1) processes for reason of computational tractability. This is an issue left for future research.

Appendix A. Multiscale Stochastic Volatility Model

For convenience, we reproduce the continuous-time multiscale SV model discussed in Molina, Han and Fouque (2010) which motivates the model in (4) - (6) as its discrete-time approximation.

Let $S(t)$ be the stock price at time $t$. Let $K$ be the number of volatility factor driving the one-dimensional $S(t)$, such that the volatility of $S(t)$ is the exponential sum of $K$ volatilities obeying Ornstein-Uhlenbeck processes:

\[
\begin{align*}
    dS(t) &= \kappa S(t)dt + \sigma(t)S(t)dW^{(0)}(t) \\
    \log(\sigma^2(t)) &= F^{(1)}(t) + \cdots + F^{(K)}(t) \\
    dF^{(j)}(t) &= \alpha_j + (\mu_j - F^{(j)}(t))dt + \beta_j dW^{(j)}(t), \quad j = 1, \cdots, K 
\end{align*}
\]  

(A.1)

where $\kappa$ is the time-invariant rate of return, $\alpha_j$ is the speed of mean reversion of volatility of the $F^{(j)}(t)$ factor toward its long-run mean level $\mu_j$, $\beta_j$ is volatility of
volatility of the $F^{(j)}(t)$ factor, and $W^{(0)}(t)$ and $W^{(j)}(t)$ for $j = 1, \ldots, K$ are possibly mutually correlated Brownian motions.

Note that $1/\alpha_j$ is a typical time scale of the $F^{(j)}(t)$ factor and they are defined as well separated, and ordered by assuming that $0 < \alpha_1 < \alpha_2$ for $K = 2$, so that the first factor can be interpreted as the longest time scale and the second factor as the shortest time scale.

Given a constant time step ($\Delta$), the Euler-Maruyama discretization scheme is applied to $S(t)$ and $F^{(j)}(t)$ for $j = 1, \ldots, K$ at time point $t_k = k\Delta$ to yield:

$$S(t_{k+1}) - S(t_k) = \kappa S(t_k) \Delta + \sigma(t_k) \sqrt{\Delta} \epsilon(t_k)$$

$$\log(\sigma^2(t_k)) = F_1(t_k) + \cdots + F_K(t_k)$$

$$F_j(t_k) - F_j(t_{k-1}) = \alpha_j + (\mu_j - F_j(t_k)) \Delta + \beta_j \sqrt{\Delta} \eta_j(t_k), \quad j = 1, \ldots, K \quad (A.2)$$

where $\epsilon(t_k)$ and $\eta_j(t_k)$ for $j = 1, \cdots, K$ are possibly mutually correlated sequences of i.i.d (initially) normal random variables. In the simplest case, the sequence of $\epsilon(t_k)$’s is initially assumed to be independent of $\eta_j(t_k)$’s but the $\eta_j(t_k)$’s can be pairwise correlated. However, in the equity markets, market participants operate at different time scales (and based on different information sets), but not independently, and where the leverage effect is pronounced.

The above discretely approximated model can be rewritten as

$$\frac{1}{\sqrt{\Delta}} \left[ \frac{S(t_{k+1}) - S(t_k)}{S(t_k)} \right] = \sigma(t_k) \epsilon(t_k)$$

$$\log(\sigma^2(t_k)) = F_1(t_k) + \cdots + F_K(t_k)$$

$$F_j(t_k) - \mu_j = (1 - \alpha_j)(F_j(t_{k-1}) - \mu_j) + \beta_j \sqrt{\Delta} \eta_j(t_k), \quad j = 1, \ldots, K \quad (A.3)$$
Next, we define returns in the usual way as

\[ y_{tk} := \frac{1}{\sqrt{\Delta}} \left[ \frac{S(t_{k+1}) - S(t_k)}{S(t_k)} - \kappa \Delta \right] \]

and write the discrete-time driving vector of volatilities as \( \mathbf{h}(tk) - \mathbf{F}(tk - \mu) \), where \( \mathbf{F}(tk) = (F_1(tk), \ldots, F_K(tk))' \) and \( \mu = (\mu_1, \ldots, \mu_K)' \). The autoregressive parameter is denoted as \( \phi_j = 1 - \alpha_j \Delta, \ j = 1, \ldots, K \). Also we define the standard deviation parameter as \( \sigma_{j,\eta} = \beta_j \sqrt{\Delta}, \ j = 1, \ldots, K \). Since observations are equally spaced, with some abuse of notation, we use \( t \) instead of \( tk/\Delta \) for discrete time indices, such that \( t \in 1, 2, \ldots, T \). This allows us to express a \( K \)-dimensional AR(1) process for log-volatilities in state-space form as

\[
\begin{align*}
    y_t &= \exp \left( (\mathbf{1}' \mathbf{h}_t + \mathbf{1}' \mu) / 2 \right) \epsilon_t, \quad t = 1, \ldots, T \\
    \mathbf{h}_{t+1} &= \Phi \mathbf{h}_t + \Sigma^{1/2} \eta_{t+1}, \quad t = 1, \ldots, T \\
    \mathbf{h}_0 &\sim N(0, \Omega) \quad (A.4)
\end{align*}
\]

where \( \eta_t = (\eta_{1,t}, \ldots, \eta_{K,t})' \) is the vector of standard random Gaussian variates, \( \mathbf{1} \) is a \( K \)-dimensional vector of ones, \( \Phi \) is a \( K \)-dimensional autoregressive (diagonal) matrix with typical elements \( \phi_j \), and \( \Sigma \) is the covariance matrix, assumed to be diagonal since the correlation parameter between the factors is not identifiable. The original version of the model assumes that the Brownian motions driving \( S(t) \) and \( F(t) \) are independent and the components of \( F(t) \) are driven by independent Brownian motions. Our discretized model is based on an extension of the original model in which the components of \( F(t) \) are assumed to be correlated to the Brownian motion driving \( S(t) \). That is, the original model assumes no correlation between the asset return and its volatility, our discretized model is based on the extension of the original model which allows for the so-called leverage effect observed in equity markets. In fact,
Black (1976) and Christie (1982) observed that a decrease of the stock price implies an increase of the associated volatility.

In section 2 of the main text we make precise the parameterization of the discrete-time multiscale stochastic-volatility model used in both the simulation and estimation processes, including the initial state and random variable notation.

References


