Available online at www.sciencedirect.com
ScienceDirect
Journal of Empirical Finance 15 (2008) 332–341
Journal of EMPIRICAL FINANCE
www.elsevier.com/locate/jempfin

Autoregressive stochastic volatility models with heavy-tailed distributions: A comparison with multifactor volatility models

Manabu Asai

Faculty of Economics, Soka University, 1-236 Tangi-cho, Hachioji-shi, Tokyo 192-8577, Japan

Accepted 20 June 2006
Available online 23 May 2007

Abstract

This paper examines two asymmetric stochastic volatility models used to describe the heavy tails and volatility dependencies found in most financial returns. The first is the autoregressive stochastic volatility model with Student's *t*-distribution (ARSV-*t*), and the second is the multifactor stochastic volatility (MFSV) model. In order to estimate these models, the analysis employs the Monte Carlo likelihood (MCL) method proposed by Sandmann and Koopman [Sandmann, G., Koopman, S.J., 1998. Estimation of stochastic volatility models via Monte Carlo maximum likelihood. Journal of Econometrics 87, 271–301.]. To guarantee the positive definiteness of the sampling distribution of the MCL, the nearest covariance matrix in the Frobenius norm is used. The empirical results using returns on the S&P 500 Composite and Tokyo stock price indexes and the Japan–US exchange rate indicate that the ARSV-*t* model provides a better fit than the MFSV model on the basis of Akaike information criterion (AIC) and the Bayes information criterion (BIC).

JEL classification: C22
Keywords: Autoregression; Monte Carlo likelihood method; Multifactor model; Stochastic volatility; Student's *t*-distribution

1. Introduction

It has long been recognized that the return volatility of financial assets tends to change over time with persistence and that there is a negative correlation between returns and this change in volatility. The latter phenomenon is known as the 'leverage' effect. For theoretical developments in continuous time, Hull and White (1987) generalized the Black and Scholes (1973) option pricing formula to analyze stochastic volatility (SV). Hull and White (1987), Scott (1987) and Wiggins (1987) allowed the spot volatility process to follow a diffusion process, and took volatility persistence and the leverage effect into account in their models.

In the framework of discrete time, Taylor (1982, 1986) and Engle (1982) modeled volatility persistence using stochastic variance and autoregressive conditional heteroskedasticity (ARCH) models, respectively. The lognormal SV model in Taylor (1982, 1986) can be thought of as an Euler discretization of the diffusion, such that the log–volatility follows a Gaussian Ornstein–Uhlenbeck process. In empirical research, Wiggins (1987), Chesney and Scott (1989), and Harvey and Shephard (1996) analyzed the leverage effects by the direct correlation between the innovations in both returns and volatility.

E-mail address: m-asai@soka.ac.jp.

0927-5398/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.jempfin.2006.06.006

Although SV models are known to be more appropriate to describe the tail thickness of financial returns than ARCH-type models, extreme movements in returns occur more frequently in the observed data than the model implies. One way to cope with this problem in discrete time is to assume a heavy-tailed distribution, such as Student's *t*-distribution, generalized error distribution (GED) and a mixture of normal distributions, as the conditional distribution of returns. Liesenfeld and Jung (2000) employed the two former distributions in the SV models, while Watanabe and Asai (2003) considered all three in capturing the high kurtosis of returns. Recently, Chernov et al. (2003) considered several two-factor continuous time SV models with a second SV factor dedicated to the exclusive modeling of tail behavior. One of the contributions of Chernov et al. (2003) was to break the link between tail thickness and volatility persistence. This paper considers these two methods to describe the heavy tails of financial returns.

Since it is not easy to derive the exact likelihood function in the framework of SV models, many methods are proposed in the literature. The four major approaches are: (i) the Bayesian Markov chain Monte Carlo (MCMC) technique suggested by Jacquier et al. (1994), (ii) the efficient methods of moments (EMM) proposed by Gallant and Tauchen (1996), (iii) the Monte Carlo likelihood (MCL) method developed by Sandmann and Koopman (1998), and (iv) the efficient importance sampling (EIS) method of Liesenfeld and Richard (2003). From among these, this paper focuses on the MCL method.

For the MCL method, the likelihood function can be approximated arbitrarily precisely by decomposing it into a Gaussian part, constructed with a Kalman filter, and a remainder function, for which the expectation is evaluated through simulation. One of the features of the MCL method is the decomposition. Since only the latter part requires Monte Carlo techniques, the MCL requires a smaller number of simulations than the other importance sampling methods. Monte Carlo experiments conducted by Sandmann and Koopman (1998) show that the MCL method has properties that are very close to those of the MCMC of Jacquier et al. (1994). While the MCMC procedure is computationally more demanding, the Monte Carlo likelihood method proposed by Sandmann and Koopman (1998) is much easier to implement computationally. Regarding the estimation of latent volatility, the MCL and the MCMC can obtain the estimates of model parameters and the values of volatility process simultaneously.

The EMM matches the score of the auxiliary model through simulation. Although the procedures based on the method of moments are known to be suboptimal relative to likelihood-based methods, Gallant and Tauchen (1996) claim that the EMM is as efficient as maximum likelihood if the auxiliary model is an accurate approximation of the distribution of the data. However, there is no method for estimating the instantaneous volatility throughout the sample, $t=1,\ldots,T$, so that an additional form of estimation, such as the Kalman filter based on $\log y_t^2$, is required.

The final EIS method is an extension of the accelerated Gaussian importance sampling (AGIS) method proposed by Danielsson and Richard (1993) and Danielsson (1994). Although the AGIS is specifically designed for models with a latent Gaussian process, the EIS procedure in its general form can be applied to models with arbitrary classes of distributions for the latent variables. While the EIS method considers the estimate of the likelihood function itself, the MCL approach only requires the estimates of the remainder function. In their example, Bauwens and Rombouts (2004) showed that the EIS and MCL produced very similar results.

The purpose of this paper is to consider an autoregressive SV model with Student's *t*-distribution (ARSV-*t*), compare it with the multifactor SV model (MFSV), and develop estimation techniques based on the MCL method proposed by Sandmann and Koopman (1998). As stated above, the MFSV can also cope with the heavy tails of financial data. Since the MFSV model of Chernov et al. (2003) incorporates leverage effects, this paper also considers negative correlation between the innovation terms of returns and volatilities. There are two reasons for including higher-order autoregressive terms. The first is that the ARSV is rarely estimated, in spite of the fact that it is a natural extension of the basic SV model. The second reason is related to the MFSV model. The log–volatility of the MFSV can be interpreted as a linear combination of latent and independent AR(1) processes. As shown by Granger and Morris (1976), the sum of two independent AR(1) processes is an ARMA(p,q) process, where $p \leq 2$ and $q \leq 1$. Since the log–volatility of the MFSV follows the ARMA model, the higher-order autoregressive terms in the ARSV may be fair.

The returns on the S&P 500 Composite index, the Tokyo stock price index (TOPIX), and the exchange rate between Japan and the US are used for the empirical analysis. For these financial returns, the ARSV-*t* and MFSV models considered in the paper can capture the heavy tails and leverage effects. Furthermore, the ARSV-*t* and MFSV models can take account of the higher-order effects of past volatilities.

The remainder of the paper is organized as follows. Section 2 considers the ARSV-*t* model in the framework of the asymmetric SV model proposed by Harvey and Shephard (1996). This section also specifies the discrete time MFSV model. Section 3 explains the MCL method proposed by Sandmann and Koopman (1998), and extends it to the ARSV-*t* and MFSV models. Section 4 reports the empirical results, and Section 5 provides some concluding remarks.

2. Autoregressive SV model with Student's *t*-distribution

This paper considers the p-th order ARSV-t model, ARSV(p)–t, as follows:

$$y_t = \sigma\xi_t \exp(h_t/2), \tag{1}$$

$$h_t = \phi_1 h_{t-1} + \cdots + \phi_p h_{t-p} + \eta_{t-1}, \tag{2}$$

$$\xi_t = \frac{\varepsilon_t}{\sqrt{\kappa_t/(v-2)}}, \quad \kappa_t \sim \chi^2(v), \tag{3}$$

$$\begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho\sigma_\eta \\ \rho\sigma_\eta & \sigma_\eta^2 \end{bmatrix} \right), \tag{4}$$

where κ_t is independent of (ε_t, η_t). Eq. (2) indicates that the log–volatility follows the AR(p) process. By this specification, the conditional distribution, ξ_t, follows the standardized t-distribution with mean zero and variance one, since $\xi_t\sqrt{v/(v-2)}$ has the Student's t-distribution. Since κ_t is independent of (ε_t, η_t), the correlation coefficient between ξ_t and η_t is also ρ. When $p=1$ and $v\rightarrow\infty$, the ARSV-t model reduces to the asymmetric SV model of Harvey and Shephard (1996), implying that there is a leverage effect when $\rho<0$. By this specification, a negative shock in returns increases the one-step-ahead volatility.

Taking the logarithm of the squared return, y_t^2, in the ARSV-t model Eqs. (1)–(4) gives a state space form,

$$\ln y_t^2 = \ln\sigma^2 + \ln\left(\frac{v-2}{v}\right) + Z\alpha_t + \zeta_t, \quad \zeta_t = \ln\left(\frac{\varepsilon_t^2}{\kappa_t/v}\right), \tag{5}$$

$$\alpha_{t+1} = \Phi\alpha_t + Z'\eta_t, \quad \eta_t \sim N(0,\sigma_\eta^2), \quad E(\zeta_t\eta_t) = 0, \tag{6}$$

where $\alpha_t=(h_t, \ldots,h_{t-p+1})'$,

$$\Phi = \left(\begin{array}{ccc|c} \phi_1 & \cdots & \phi_{p-1} & \phi_p \\ \hline & & & 0 \\ & I_{p-1} & & \vdots \\ & & & 0 \end{array} \right) : p \times p,$$

and Z is a $1\times p$ vector with one in the first element and zeros in the other elements. The disturbance of measurement Eq. (5), ζ_t, has the logarithmic F distribution with degrees of freedom (1, v). The density function of ζ_t is given by

$$p(\zeta) = C_F(1+e^\zeta/v)^{-(v+1)/2}e^{\zeta/2}, \quad C_F = \frac{\Gamma((v+1)/2)}{\Gamma(v/2)\sqrt{v\pi}}. \tag{7}$$

It should be noted that, once the transformation is accomplished, the information about the correlation coefficient, ρ, is lost, but can be recovered by conditioning the signs of y_t (Harvey and Shephard, 1996).

Let s_t denote the signs of y_t so that s_t takes one when y_t is positive, and zero otherwise. Harvey and Shephard (1996) showed that, conditionally on s_t, $\eta_t|s_t\sim N(as_t, \sigma_\eta^2-a^2)$ and $E(\varepsilon_t \ln \varepsilon_t^2|s_t)=1.1061s_t$, where $a=\rho\sigma_\eta\sqrt{2/\pi}=0.7979\rho\sigma_\eta$. Turning to the state space form of the ARSV-t model, the transition Eq. (6) given s_t is

$$\alpha_{t+1} = as_tZ' + \Phi\alpha_t + Z'\eta_t^*, \quad \eta_t^* \sim N(0,\sigma_\eta^2 - a^2), \ E(\zeta_t\eta_t^*) = bs_t, \tag{8}$$

where $b=1.1061\rho\sigma_\eta$. The last equality is obtained by the independence of ε_t and κ_t, and $E(\zeta_t\eta_t*)=E(\zeta_t\eta_t|s_t)=E[\zeta_tE(\eta_t|\varepsilon_t)|s_t]=\rho\sigma_\eta E(\varepsilon_t\zeta_t|s_t)$. Based on the state space forms (5) and (8), Section 3 develops the MCL method to estimate the parameters of the ARSV-t model.

Before explaining the MCL method, this paper introduces a K-factor SV model as follows:

$$y_t = \xi_t \sigma(h_{1t}, \ldots, h_{Kt}), \tag{9}$$

$$\xi_t = \sum_{i=0}^{K} \psi_i \varepsilon_{it}, \tag{10}$$

$$h_{i,t+1} = \phi_i h_{it} + \sigma_{\varepsilon,i}\varepsilon_{it}, \; i = 1, \ldots, K, \tag{11}$$

$$\varepsilon_{it} \sim N(0,1), \quad i = 0, 1, \ldots, K, \tag{12}$$

where $\varepsilon_{0t}, \ldots, \varepsilon_{Kt}$ are mutually independent, and volatility, $\sigma(\cdot)$, is defined by

$$\sigma(h_1, \ldots, h_K) = \sigma \exp\left(\frac{1}{2}\sum_{i=1}^{K} h_i\right).$$

Namely, volatility $\sigma(\cdot)$ is a log-linear function of the sum of K stochastic volatility factors. It is assumed that $\psi_0 = \sqrt{1-\psi_1^2-\cdots-\psi_K^2}$, and that $\psi_1, \ldots, \psi_K$ are correlation coefficients to describe the leverage effects. By the specification, ξ_t follows the standard normal distribution. In other words, the restriction on ψ_0 is an identification restriction setting the innovation variance equal to one. Furthermore, the correlation coefficient between ξ_t and ε_{it} is given by ψ_i for $i=1, \ldots, K$. When $\psi_i<0$ for all i, a negative shock in return increases one-step-ahead volatility via $h_{i,t+1}$. Thus, the restriction that $\psi_i<0$ for all i is the sufficient condition for leverage.

If $K=2$, the MFSV model is equivalent to Durham's (2006) two-factor model. Furthermore, the MFSV model is reduced to the two-component model of Liesenfeld and Richard (2003) when $K=2$ and $\psi_1=\psi_2=0$.

The models (9)–(12) can be interpreted as an Euler discretization of the continuous time multifactor SV model proposed by Gallant et al. (1999) and Chernov et al. (2003). These analyses introduced additional SV factors to behave as factors dedicated to the exclusive modeling of tail behavior. For $K=2$ the sum of the two factors follows an ARMA process. This result comes from Granger and Morris (1976); the sum of two independent AR processes, AR(p_1) and AR(p_2), is expressed as ARMA(p, q), where $p \le p_1+p_2$ and $q \le \max(p_1, p_2)$. Thus, the multifactor model captures not only extreme values but also the ARMA process of log–volatility. In the extreme case of $K=1$, the models (9)–(12) reduce to the asymmetric SV model of Harvey and Shephard (1996).

Given s_t, $\ln y_t^2$ in the MFSV models (9)–(12) has a state space form

$$\ln y_t^2 = \ln \sigma^2 + Z^+\alpha_t + \zeta_t^+, \quad \zeta_t^+ \sim \ln\chi_1^2, \tag{13}$$

$$\alpha_{t+1} = c_t + \Phi^+\alpha_t + e_t, \quad e_t \sim N(0,Q), \; E(e_t\zeta_t) = G_t, \tag{14}$$

where $\alpha_t=(h_{1t}, \ldots, h_{Kt})'$, $\Phi^+=\mathrm{diag}(\phi_1, \ldots, \phi_K)$, Z^+ is a $1\times K$ vector of ones, $c_t=0.7979 s_t\,(\psi_1\sigma_{\varepsilon,1}, \ldots, \psi_K\sigma_{\varepsilon,K})'$, $G_t=1.1061 s_t\,(\psi_1\sigma_{\varepsilon,1}, \ldots, \psi_K\sigma_{\varepsilon,K})'$, and

$$Q = \left(I_K - \frac{2}{\pi}\mathrm{diag}(\psi_1^2, \ldots, \psi_K^2)\right) \odot \mathrm{diag}(\sigma_{\varepsilon,1}^2, \ldots, \sigma_{\varepsilon,K}^2).$$

The following section also develops the MCL estimation technique based on the state space form.

3. Monte Carlo likelihood method

The Monte Carlo likelihood (MCL) approach for non-Gaussian models is based on importance sampling techniques. Durbin and Koopman (1997) showed that the log-likelihood function of state space models with non-Gaussian measurement disturbances can be simply expressed as

$$\ln L(Y|\theta) = \ln L_{\mathrm{G}}(Y|\theta) + \ln E_{\mathrm{G}}\left[\frac{p_{\mathrm{true}}(\zeta|\theta)}{p_{\mathrm{G}}(\zeta|Y,\theta)}\right], \tag{15}$$

where $Y=(\ln y_1^2,\ldots, y_T^2, s_1,\ldots, s_T)'$, $\ln L_G(Y|\theta)$ is the log-likelihood function of the approximating Gaussian model, $p_{\text{true}}(\zeta|\theta)$ is the true density function of the measurement equation, $p_G(\zeta|Y, \theta)$ is the Gaussian density of the measurement disturbances of the approximating model, and E_G refers to the expectation with respect to the 'so-called' importance density $p_G(\zeta|Y, \theta)$ associated with an approximating model. Eq. (15) shows that the non-Gaussian log-likelihood function can be expressed as the log-likelihood function of the Gaussian approximating model plus a correction for the departure from Gaussian assumptions in relation to the true model.

The estimator of $\ln L(Y|\theta)$ is given by

$$\ln \hat{L}(Y|\theta) = \ln L_G(Y|\theta) + \ln \overline{w} + \frac{s_w^2}{2M\,\overline{w}^2}, \tag{16}$$

where

$$\overline{w} = \frac{1}{M}\sum_{i=1}^{M} w_i, \quad s_w^2 = \frac{1}{M-1}\sum_{i=1}^{M}(w_i - \overline{w})^2, \quad w_i = \frac{p_{\text{true}}(\zeta^i|\theta)}{p_G(\zeta^i|Y,\theta)},$$

and ζ^i denotes a draw from the importance density $p_G(\zeta|Y, \theta)$. The accuracy of this estimator depends on the properties of the so-called weights w_i. Since the simulation samples are independent, it immediately follows that the variance due to simulation decreases as M increases. It should be noted that the last term in Eq. (16) is the correction term for the bias due to the log-linear transformation of the likelihood function, and that the estimated asymptotic variance of $\ln \hat{L}(Y|\theta)$ is $s_w^2/(M\,w^{-2})$ (Durbin and Koopman, 1997). The current paper sets $M=10$, as the Monte Carlo results of Sandmann and Koopman (1998) showed that $M=5$ is sufficient for practical purposes. As shown in Section 4, the standard error of the estimated log likelihood is very small.

The approximating Gaussian density, $p_G(\zeta|Y, \theta)$, considered in the paper is based on $\zeta_t \sim N(0, H_t)$ $(t=1,\ldots,T)$. In the approximating density, H_t is used as the variance of the measurement equation. The scalar variance H_t are chosen so as to make the differences between the log densities $p_{\text{true}}(\zeta|\theta)$ and $p_G(\zeta|Y, \theta)$ as constant as possible in the neighborhood of $\hat{\zeta}=E(\zeta|Y, \theta)$. Equalizing the first derivatives of the true and approximating densities at $\hat{\zeta}$ yields

$$H_t = \frac{2\hat{\zeta}_t(v+\hat{\zeta}_t)}{v(e^{\hat{\zeta}_t}-1)} \tag{17}$$

for the ARSV-t model, and

$$H_t = \frac{2\hat{\zeta}_t}{e^{\hat{\zeta}_t}-1} \tag{18}$$

for the MFSV model. By the specification, H_t is always positive. In order to construct $p_G(\zeta|Y, \theta)$, the joint density of ζ_t and the disturbance of the transition equation must have the positive definite covariance matrix, Ω_t. Although the above H_t produce an efficient approximation of the density of ζ_t, they invoke another problem: that is, the nonpositive definiteness of some of Ω_t, which is defined by

$$\Omega_t = \begin{pmatrix} \sigma_\eta^2 - a^2 & bs_t \\ bs_t & H_t \end{pmatrix},$$

for the ARSV-t model, and

$$\Omega_t = \begin{pmatrix} Q & G_t \\ G_t' & H_t \end{pmatrix},$$

for the MFSV model. Here, the elements of Ω_t except for H_t are obtained from the transition Eqs. of (8) and (14). If there are no asymmetric effects, i.e., $\rho=0$ for the ARSV-t and $\psi_1=\cdots=\psi_K=0$ for the MFSV, then Ω_t is always positive definite due to the construction of H_t. This paper, however, assumes asymmetry, and hence the off-diagonal elements of Ω_t cause the problem. For instance, H_t must be larger than $b^2/(\sigma_\eta^2-a^2)$ in order to be positive definite, but there are situations where the condition is in conflict with Eq. (17). This is crucial since the nonpositive definiteness of some Ω_t leads to a stall in the Kalman filtering and smoothing.

Table 1
MCL estimates of asymmetric SV models

	S&P	TOPIX	YEN/USD
ϕ	0.9540 (0.0081)	0.9609 (0.0088)	0.9471 (0.0139)
σ_η	0.2269 (0.0223)	0.2420 (0.0243)	0.2318 (0.0312)
ρ	−0.3117 (0.0505)	−0.4561 (0.0491)	−0.1796 (0.0288)
σ	0.7965 (0.0366)	1.0221 (0.0596)	0.6282 (0.0288)
Log like	−7872.5	−5163.1	−5307.9
AIC	15753.0	10334.2	10623.7
BIC	15777.7	10357.3	10646.9

Note: Standard errors are in parenthesis. Maximized log likelihoods are based on ln y_t^2.

To cope with this problem, this paper proposes to use the nearest covariance matrix of Higham (1988). Higham (1988) proved that the nearest positive symmetric matrix in the Frobenius norm to any real symmetric matrix C is $(C+P)/2$, where P is the symmetric polar factor of C. When Ω_t is nonpositive semidefinite in the above approximation, this approach replaces it by its nearest covariance matrix. For the models and data sets used in this paper, the Kalman filtering and smoothing employed the nearest covariance matrix with at most two Ω_t in each iteration.

The smoother error vector $\hat{\zeta}=(\hat{\zeta},\ldots,\hat{\zeta}_T)'$ is calculated by the Kalman filter smoother. For each trial H_t, a new value of $\hat{\zeta}$ is obtained by routine Kalman filtering and smoothing and this value is used to obtain new H_t. Convergence is quick, and requires 8–10 iterations for the ARSV-t and MFSV models.

The simulation smoother proposed by de Jong and Shephard (1995) is used to generate sample ζ^i from the approximating Gaussian density p_G $(\zeta|Y,\theta)$. The weights for the Monte Carlo estimates are given by

$$w_i = \prod_{t=1}^{t} C_F\sqrt{2\pi H_t}\exp\Big(0.5\{\zeta_t^i + \zeta_t^{i2}/H_t - (v+1)\ln(1+e^{\zeta_t^i}/v)\}\Big),$$

for the ARSV-t model, and

$$w_i = \prod_{t=1}^{t} \exp\Big(0.5\{\ln H_t + \zeta_t^i + \zeta_t^{i2}/H_t - e^{\zeta_t^i}\}\Big),$$

for the MFSV model.

The MCL estimates of model parameters, θ, are obtained by numerical optimization of the unbiased estimate of Eq. (15). The log-likelihood function of the approximating model, ln $L_G(Y|\theta)$, can be used to obtain starting values. See Sandmann and Koopman (1998) and Durbin and Koopman (1997) for more details regarding the MCL method.

The performance of the MCL estimator is investigated to a certain extent. For the basic SV model, which can be obtained by setting $\rho=0$ in the asymmetric SV model of Harvey and Shephard (1996), the Monte Carlo results conducted by Sandmann and Koopman (1998) showed that the MCL method has properties that are very close to the Bayesian Markov chain Monte Carlo method of Jacquier et al. (1994). As for the asymmetric SV models, the Monte Carlo results of Asai and McAleer (2005a,b) showed that the MCL estimator has a small bias in finite samples, and that the root mean squared errors are quite reasonable.

Table 2
MCL estimates of SV-t models

	S&P	TOPIX	YEN/USD
ϕ	0.9933 (0.0026)	0.9692 (0.0080)	0.9820 (0.0066)
σ_η	0.0833 (0.0127)	0.1971 (0.0236)	0.1258 (0.0221)
σ	0.7053 (0.0724)	0.9599 (0.0664)	0.5576 (0.0409)
v	5.9647 (0.6389)	10.130 (2.3062)	7.6078 (1.2242)
Log like	−7835.9	−5183.0	−5294.3
AIC	15679.9	10373.9	10596.6
BIC	15704.6	10397.1	10619.8

Note: Standard errors are in parenthesis. Maximized log likelihoods are based on ln y_t^2.

Table 3
AIC and BIC based on $\ln y_t^2$ for fitted K-factor SV models

K	S&P		TOPIX		YEN/USD	
	AIC	BIC	AIC	BIC	AIC	BIC
1	15753.0	15777.7	10334.2	10357.4	10623.7	10646.9
2	15651.9*	15695.3*	10316.8*	10357.3*	10598.5*	10639.2*
3	15719.1	15781.6	10324.8	10382.7	10599.9	10658.0
4	15653.7	15734.2	10327.1	10402.3	10599.2	10674.7

Note: '*' denotes the optimal model.

Smoothed estimates of volatility can be obtained using Sandmann and Koopman's (1998) method, while filtered estimates can be calculated based on the efficient importance sampling (EIS) approach of Liesenfeld and Richard (2003). For the purpose of diagnostic checking, the latter estimates can be used to construct standardized residuals $\hat{z}_t = y_t/\hat{\sigma}_t$, where $\hat{\sigma}_t^2$ is the EIS estimate of the conditional expectation of volatility given the past observations of returns. For instance, the conditional expectation of volatility of the ARSV-t model is given by $E[\exp(h_t)|Y_{t-1}]$, where $Y_{t-1} = (y_1,\ldots, y_{t-1})$. In order to check the distributional assumptions, the current paper applies a similar approach to the method of Kim et al. (1998). Under the assumption of the ARSV-t model, $z_t = y_t \exp(-0.5h_t)$ follows a standardized t-distribution. Let $F_t()$ denotes the cumulative distribution function of standardized t-distribution. If the model is correctly specified, $u_t = F_t(\hat{z}_t)$ is a serially independent random variable following a uniform distribution on [0, 1]. Using the inverse of a standard normal distribution function, the variable u_t can be mapped into a standard normal distribution: $z_t^* = F_N^{-1}(u_t)$. Under the hypothesis of a correctly specified model, these normalized residuals are serially independent random variables following a standard normal distribution.

4. Empirical evidence

This section examines the MCL estimates of the two SV models, i.e., the multifactor SV (MFSV) and ARSV(p)–t models, for three sets of empirical data, namely, Standard and Poor's 500 Composite Index (S&P), the Tokyo stock price index (TOPIX), and the Japanese yen/US dollar exchange rate (YEN/USD). The sample period for S&P is 1/6/1986 to 12/4/2000, giving 3604 observations, TOPIX is from 1/4/1990 to 9/30/1999, providing 2404 observations, and the YEN/USD exchange rate is from 1/4/1990 to 12/28/1999, yielding 2467 observations. These same data sets have been previously analyzed by Asai and McAleer (2005a).

As a preliminary result, Table 1 shows the MCL estimates of the basic asymmetric SV model, which can be obtained by setting $K=1$ in the MFSV, or by letting $p=1$ and $\nu\to\infty$ in the ARSV(p)–t model. The estimates of ϕ are between 0.947 and 0.961, while those of σ_η are between 0.226 and 0.242. The estimates of ρ are negative, and this implies that, for each data set, the innovations in the mean and volatility are negatively correlated. Although it is not shown in the table, the Monte Carlo standard error for the S&P is 0.140. Table 2 presents the MCL estimates of the SV-t model, which is a special case of the ARSV(1) model restricted to $\rho=0$. The estimates of ν are between 5.9 and 10.2, and the results indicate that there are extreme values that are difficult to capture by the normal distribution. Another interesting

Table 4
MCL estimates of 2-factor SV models

	S&P	TOPIX	YEN/USD
ϕ_1	0.9931 (0.0028)	0.9781 (0.0065)	0.9860 (0.0062)
$\sigma_{\eta,1}$	0.0813 (0.0152)	0.1781 (0.0232)	0.0970 (0.0234)
ψ_1	−0.4342 (0.1231)	−0.5417 (0.0733)	−0.1267 (0.1297)
ϕ_2	0.3539 (0.1367)	0.3882 (0.1806)	0.5695 (0.1500)
$\sigma_{\eta,2}$	0.5572 (0.0437)	0.4024 (0.0552)	0.4183(0.0576)
ψ_2	−0.3621 (0.0664)	−0.2407 (0.0935)	−0.2389 (0.0771)
σ	0.7691 (0.0676)	0.9915 (0.0712)	0.6084(0.0428)
Log like	−7819.0	−5151.4	−5292.3
AIC	15651.9	10316.8	10598.5
BIC	15695.3	10357.3	10639.2

Note: Maximized log-likelihoods are based on $\ln r_t^2$.

Table 5
AIC and BIC based on ln y_t^2 for fitted ARSV(p)–t models

p	S&P		TOPIX		YEN/USD	
	AIC	BIC	AIC	BIC	AIC	BIC
1	15650.1	15681.0	10313.0*	10341.9*	10590.8*	10619.9*
2	15649.0	15686.2	10314.7	10349.4	10592.3	10627.2
3	15636.2	15679.6*	10314.3	10354.7	10593.7	10634.4
4	15634.4	15684.0	10313.9	10360.2	10593.6	10640.1
5	15634.1*	15689.8	10315.8	10367.8	10592.1	10644.4

Note: '*' denotes the optimal model.

feature is that the value of ϕ (σ_η) is larger (smaller) than the value for the asymmetric SV model. In other words, the heavy tails for the mean innovation term make volatility persistence high.

Instead of standardized Student's *t*-distribution, the MFSV model includes additional factors to capture extreme values. Table 3 shows the Akaike information criterion (AIC) and Bayes information criterion (BIC) for the MFSV model to help choose the optimal number of factors. The results indicate that AIC and BIC selected $K=2$ for all data sets. Table 4 reports the MCL estimates for the two-factor SV model. The estimates of $(\phi_1,\sigma_{\eta,1},\psi_1)$, which represents the first factor, are close to those of Table 2, indicating that high persistence is accomplished without using a heavy-tailed distribution.
Estimates of ϕ_2 for the second factor of S&P and TOPIX are smaller than 0.4. Furthermore, ϕ_2 for S&P is insignificant. The first volatility factors have high persistence, while the small values of ϕ_2 indicate the low persistence of the second volatility factors. All values of ψ_i are negative, indicating a negative correlation between returns and changes in volatilities. Regarding the correlation coefficient of the second factor, the estimates of ψ_2 are all smaller than -0.2 and significant. These results imply that, although the persistence of the second factor is low, the second factor plays an important role in the sense that it captures extreme values, which may produce the leverage effect.

Compared to Table 2, Table 3 shows that AIC and BIC select the two-factor model for S&P and TOPIX, while they choose the SV-*t* for YEN/USD. Since the SV-*t* model neglects asymmetric effects, further investigation is desirable.

The ARSV(p)–t model is an extension of the SV-*t*, and can capture not only heavy tails but also the higher-order volatilities and leverage effects. Table 5 shows the values of AIC and BIC based on ln y_t^2 for fitted ARSV(p)–t models. While AIC and BIC selected $p=1$ for TOPIX and YEN/USD, they favor higher-order models for S&P. As for S&P, AIC and BIC selected $p=5$ and $p=3$, respectively. Compared to Table 4, Table 5 shows that the selected ARSV(p)–t model is superior to the two-factor SV model with respect to AIC and BIC. Table 6 reports parts of the MCL estimates of ARSV(p)-*t* models. Regarding the case of $p=1$, the results of Table 5 (except for ρ) are close to those of Table 2. A more interesting feature for S&P is that, as the lag-length p increases, the degree of freedom v grows. Similar results are found for the TOPIX and YEN/USD, but these are omitted to save space. This relation implies that introducing an additional lagged term may capture extreme values to a certain extent, but not in all cases.

Table 6
MCL estimates of ARSV-*t* models

	S&P			TOPIX	YEN/USD
	$p=1$	$p=3$	$p=5$	$p=1$	$p=1$
ϕ_1	0.9890 (0.0036)	0.9914 (0.2578)	0.9756 (0.1984)	0.9745(0.0068)	0.9803 (0.0069)
ϕ_2		−0.6680 (0.3520)	−0.9282(0.2939)		
ϕ_3		0.6574 (0.1204)	0.9600 (0.2796)		
ϕ_4			−0.3999 (0.2580)		
ϕ_5			0.3692(0.1169)		
σ_η	0.1156 (0.0160)	0.1996 (0.0317)	0.2529 (0.0399)	0.1946 (0.0215)	0.1274 (0.0227)
ρ	−0.5315 (0.0683)	−0.5615 (0.0601)	−0.5363 (0.0650)	−0.5544 (0.0571)	−0.2697 (0.0894)
σ	0.8482 (0.0639)	0.8350 (0.0608)	0.8333 (0.0643)	1.0347 (0.0707)	0.6521 (0.0425)
v	6.4031 (0.7023)	6.9173 (0.7931)	7.3261 (1.0028)	10.378 (1.5683)	7.7564 (1.3009)
Log like	−7820.0	−7811.12	−7808.1	−5151.5	−5290.4
AIC	15650.1	15636.2	15634.1	10313.0	10590.8
BIC	15681.0	15679.6	15689.8	10341.9	10619.9

Note: Maximized log likelihoods are based on ln $y_t{}^2$.

Table 7
Diagnostics for ARSV(p)–t models

	S&P			TOPIX	YEN/USD
	$p=1$	$p=3$	$p=5$	$p=1$	$p=1$
$Q_{30}(\hat{z}_t)$	0.078	0.052	0.072	0.195	0.155
$Q_{30}(\ln \hat{z}_t^2)$	0.436	0.125	0.073	0.807	0.077
J–B	0.037	0.172	0.465	0.196	0.718

Note: p-values are shown in the table. '$Q_{30}(\hat{z}_t)$' and '$Q_{30}(\ln \hat{z}_t^2)$' denote the Ljung–Box statistics including 30 lags for $\hat{z}t$ and $\ln \hat{z}_t^2$, respectively, where $\hat{z}_t$ are the standardized residuals. 'J–B' is the Jarque–Bera test for z_t^*, which are the converted series based on the method explained in the text.

The EIS method of Liesenfeld and Richard (2003) is used to obtain the filtered volatility estimates, $\hat{\sigma}_t$. By using the standardized residuals, $\hat{z}_t = y_t/\hat{\sigma}_t$, three diagnostic statistics were calculated, and are shown in Table 7. The Ljung–Box statistic $Q_{30}(\cdot)$ for $\hat{z}_t$ and $\ln \hat{z}_t^2$ including 30 lags indicates that the model successfully accounts for the serial correlation in the volatility of all three series based on the models listed in Table 5. As for S&P, the kurtosis of the converted series z_t^* are 3.6497 and 3.1062 for $p=1$ and $p=3$, respectively. These values are very close, but the Jarque–Bera test shows a difference. While the Jarque–Bera test rejects the former, it does not reject the latter. This result indicates that ARSV (1)-*t* is inappropriate to describe tail behavior for the data set. The ARSV(5) also pass the Jarque–Bera test. The test does not reject the assumption that the ARSV(1)-*t* is correctly specified for TOPIX and YEN/USD.

For TOPIX and YEN/USD, the ARSV(1)-*t* model, which can cope with heavy tails and the leverage effect, is preferred to the two-factor SV model. As the two-factor SV model has ARMA structure volatility, the results indicate that AR(1) is sufficient for volatilities of TOPIX and YEN/USD. Regarding S&P, the ARSV(p)–t and ARSV(5)–*t* models are also preferred to the two-factor SV model. Although the MFSV models have several interesting features, the empirical results show that the ARSV(p)–t better fits the S&P, TOPIX and YEN/USD.

5. Concluding remarks

This paper investigated two kinds of asymmetric SV models to describe heavy tails for financial returns. One is the autoregressive SV model with Student's *t*-distribution (ARSV-*t*), and the other is the multifactor stochastic volatility (MFSV) model. The analysis employed Monte Carlo likelihood methods to estimate these models. To guarantee the positive definiteness of the importance sampling distribution, this paper proposed the use of the nearest covariance matrix in the Frobenius norm. The ARSV-*t* models were compared with MFSV models using S&P 500, TOPIX and YEN/USD returns. The empirical results show that a two-factor SV model can describe extreme values to a certain extent, without using any heavy-tailed distribution, but there still remains an unexplained component. Both AIC and BIC favor the ARSV-*t* model over the two-factor SV model for all three series.

This paper has made certain contributions, but several extensions are still possible. First, the paper focuses on leptokurtic distribution, but it is also worthwhile fitting skewed distributions, including the skewed *t*-distribution (see Hansen (1994) and Fernández and Steel (1998)). Secondly, the paper only considered autoregression, but there is ample room to consider more elaborate models such as ARMA and long-memory models. Thirdly, it may be worthwhile to include jump diffusions.

Acknowledgements

The author would like to acknowledge helpful discussions with Yoshihisa Baba of Soka University. The author also gratefully acknowledges the financial support of a grant-in-aid for the 21st century COE program "Microstructure and Mechanism Design in Financial Markets" from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

Asai, M., McAleer, M., 2005a. Dynamic asymmetric leverage in stochastic volatility models. Econometric Reviews 24, 317–332.
Asai, M., McAleer, M., 2005b. Alternative leverage and threshold effects in stochastic volatility models. Unpublished manuscript, Faculty of Economics, Soka University.

Bauwens, L., Rombouts, J.V.K., 2004. Econometrics. In: Gentle, J.E., Hardle, W., Mori, Y. (Eds.), Handbook of Computational Statistics, vol. 1. Springer-Verlag, New York, pp. 951–979.

Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. Journal of Political Economy 81, 637–654.

Chernov, M., Gallant, A.R., Ghysels, E., Tauchen, G., 2003. Alternative models for stock price dynamics. Journal of Econometrics 116, 225–257.

Chesney, M., Scott, L.O., 1989. Pricing European currency options: a comparison of the modified Black–Scholes model and a random variance model. Journal of Financial and Quantitative Analysis 24, 267–284.

Danielsson, J., 1994. Stochastic volatility in asset prices: estimation with simulated maximum likelihood. Journal of Econometrics 61, 375–400.

Danielsson, J., Richard, J.-F., 1993. Quadratic acceleration for simulated maximum likelihood evaluation. Journal of Applied Econometrics 8, 153–173.

Durbin, J., Koopman, S.J., 1997. Monte Carlo maximum likelihood estimation for non-Gaussian state space models. Biometrika 84, 669–684.

Durham, G., 2006. Monte Carlo methods for estimating, smoothing, filtering one- and two-factor stochastic volatility models. Journal of Econometrics 135, 273–305.

Engle, R.F., 1982. Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. Econometrica 50, 987–1007.

Fernández, C., Steel, M.F.J., 1998. On Bayesian modeling of fat tails and skewness. Journal of the American Statistical Association 93, 359–371.

Gallant, A.R., Tauchen, G., 1996. Which moments to match? Econometric Theory 12, 657–681.

Gallant, A.R., Hsu, C.-T., Tauchen, G., 1999. Using daily range data to calibrate volatility diffusions and extract the forward integrated variance. Review of Economics and Statistics 81, 617–631.

Granger, C.W.J., Morris, M., 1976. Time series modeling and interpretation. Journal of the Royal Statistical Society. A 139, 246–257.

Hansen, B.E., 1994. Autoregressive conditional density estimation. International Economic Review 35, 705–730.

Harvey, A.C., Shephard, N., 1996. Estimation of an asymmetric stochastic volatility model for asset returns. Journal of Business and Economic Statistics 14, 429–434.

Higham, N.J., 1988. Computing a nearest symmetric positive semidefinite matrix. Linear Algebra and its Applications 103, 103–118.

Hull, J., White, A., 1987. The pricing of options on assets with stochastic volatility. Journal of Finance 42, 281–300.

Jacquier, E., Polson, N.G., Rossi, P.E., 1994. Bayesian analysis of stochastic volatility models. Journal of Business and Economic Statistics 12, 371–389.

de Jong, P., Shephard, N., 1995. The simulation smoother for time series models. Biometrika 82, 339–350.

Kim, S., Shephard, N., Chib, S., 1998. Stochastic volatility: likelihood inference and comparison with ARCH models. Review of Economic Studies 65, 361–393.

Liesenfeld, R., Jung, R.C., 2000. Stochastic volatility models: conditional normality versus heavy-tailed distributions. Journal of Applied Econometrics 15, 137–160.

Liesenfeld, R., Richard, J.-F., 2003. Univariate and multivariate stochastic volatility models: estimation and diagnostics. Journal of Empirical Finance 10, 505–531.

Sandmann, G., Koopman, S.J., 1998. Estimation of stochastic volatility models via Monte Carlo maximum likelihood. Journal of Econometrics 87, 271–301.

Scott, L., 1987. Option pricing when the variance changes randomly: theory, estimation and an application. Journal of Financial and Quantitative Analysis 22, 419–438.

Taylor, S.J., 1982. Financial returns modeled by the product of two stochastic processes — a study of daily sugar prices 1961–79. In: Anderson, O.D. (Ed.), Time Series Analysis: Theory and Practice, vol. 1. North-Holland, Amsterdam, pp. 203–226.

Taylor, S.J., 1986. Modelling Financial Time Series. John Wiley, Chichester.

Watanabe, T., Asai, M., 2003. Stochastic volatility models with heavy-tailed distributions: a Bayesian analysis. Unpublished manuscript, Tokyo Metropolitan University.

Wiggins, J.B., 1987. Option values under stochastic volatility: theory and empirical estimates. Journal of Financial Economics 19, 351–372.