Bayesian analysis of GARCH and stochastic volatility: modeling leverage, jumps and heavy-tails for financial time series*

Jouchi Nakajima

Department of Statistical Science, Duke University, Durham 27705, USA.

`jouchi.nakajima@stat.duke.edu`

September, 2009

Abstract

This paper develops a Bayesian model comparison for two broad major classes of varying volatility model, GARCH and stochastic volatility (SV) models on financial time series. The leverage effect, jumps and heavy-tailed errors are incorporated into the two models. For estimation, the efficient Markov chain Monte Carlo methods are developed and the model comparisons are examined based on the marginal likelihood. The empirical analyses are illustrated using the daily return data of US stock indices, individual securities and exchange rates of UK Sterling and Japanese Yen against US Dollar. The estimation results indicate that the SV model with leverage and Student-t errors yields the best performance among the competing models on the return data.

Key words: GARCH, Heavy-tailed errors, Jumps, Leverage effect, Markov chain Monte Carlo, Stochastic volatility.

---

1 Introduction

The time-varying volatility models have been widely used in various contexts of a time series analysis. Two main streams of modeling a changing variance, the GARCH (generalized autoregressive conditional heteroskedasticity) and the stochastic volatility (SV) model, are well established in financial econometrics. Numerous papers develop their extensions, and these specifications are more and more often applied for empirical analyses in financial economics and macroeconomics. Among them, this paper focuses on three common factors, namely leverage effects, jumps and heavy-tailed errors, which are stressed in literature as important elements to capture the behaviors of financial time series such as stock price and exchange rates.

The GARCH specification, proposed by Bollerslev (1986), formulates the serial dependence of volatility and incorporates the past observations into the future volatility (e.g., Bollerslev et al. (1994)). Nelson (1991) proposes the EGARCH (exponential GARCH) specification, modeling the leverage effect, which refers to the increase in volatility following a previous drop in stock returns (Black (1976)). Glosten et al. (1993) extends the GARCH model with leverage effect in another way, called the GJR model. These models came to be standards of the so-called asymmetric GARCH model. As for the jump specification, Jorion (1988) firstly introduces the GARCH model with jumps, and later, more complicated jump dynamics are developed by several papers (e.g., Chan and Maheu (2002), Maheu and McCurdy (2004)).

In the other stream, the SV models, based on the continuous-time probability process, have also been well studied in financial econometrics (see e.g., Ghysels et al. (2002), Shephard (2005)). Among their generalizations, the leverage effect, jump components and heavy-tailed errors in stock returns are well-known to be important for financial time series in the recent literature (Chib et al. (2002), Jacquier et al. (2004), Berg et al. (2004), Yu (2005), Omori et al. (2007), Nakajima and Omori (2009)). The SV model with Student-t errors is one of the most popular models to account for heavier tailed returns. However, it has been found insufficient to express the tail fatness of returns to some extents. The jump components have recently been introduced to explain the tail behavior (Eraker et al. (2003), Nakajima and Omori (2009)). Various specifications of the SV-jump models are compared in empirical studies (Chernov et al. (2003), Raggi and Bordignon (2006), Li et al. (2008)).

The purpose of this paper is to compare the fit of the models in the class of the GARCH and the SV model with leverage, jumps and heavy-tails. The GARCH and the SV models have not been compared directly, especially in the classes with the assumptions of these three

components. Several studies (Kim et al. (1998), Giot and Laurent (2004)) examine the model comparisons among the models in the two classes. Lehar et al. (2002) provides a model comparison between the GARCH and the SV models from an option pricing point of view.

The major reason which makes it difficult to compare the GARCH and the SV class is that the likelihood function of the SV model is not easily available. It is possible to compute the likelihood using a simulation-based method for a given set of parameters, but it requires a computational burden since we need to repeat the filtering procedure for many times to evaluate the likelihood function for each set of parameters until it reaches the maximum. To overcome this difficulty, we take a Bayesian estimation approach with the MCMC methods (e.g., Chib and Greenberg (1996)) for a precise and efficient estimation of the SV model. In the SV literature, Kim et al. (1998) develop a fast and reliable MCMC algorithm, called *mixture sampler*. Using this method, the jumps and heavy-tails (Chib et al. (2002)), the leverage and heavy-tails (Omori et al. (2007)), and the leverage, jumps and heavy-tails (Nakajima and Omori (2009)) are incorporated into the SV model.

On the other hand, Bauwens and Lubrano (1998), Vrontos et al. (2000), Nakatsuma (2000) develop the MCMC estimation method for the models in the GARCH class. In this paper the MCMC algorithms for the GARCH and the SV model with leverage, jumps and heavy-tails are developed. This paper adopts the Bayesian model comparison for both the GARCH and the SV models based on the marginal likelihood, which can be computed by the technique of Chib (1995), Chib and Jeliazkov (2001, 2005).

The rest of paper is organized as follows. In Section 2, the MCMC estimation method for the GARCH model with leverage, jumps and heavy-tails is developed. Section 3 reviews the MCMC estimation scheme for the SV model with leverage, jumps and heavy-tails. In Section 4, we show the estimation results of the Bayesian model comparison among competing models using daily US stock returns. Section 5 provides the model comparison of daily exchange rates of UK Sterling and Japanese Yen against US Dollar. In Section 6, the robustness of the model comparison is examined with respect to sample period and prior sensitivity. Finally, Section 7 concludes.

2 Bayesian inference for the GARCH model with leverage, jumps and heavy-tails

2.1 The model

We first consider a standard GARCH$(1,1)$ model with jumps and heavy-tails formulated as

$$
\begin{aligned}
y_t &= E_{t-1}(y_t) + e_t, & (1)\\
e_t &= k_t\gamma_t + \sqrt{\lambda_t}\varepsilon_t\sigma_t, \quad \varepsilon_t \sim N(0,1), \quad t = 1,\ldots,n, & (2)\\
\sigma_t^2 &= \omega + \beta\sigma_{t-1}^2 + \alpha e_{t-1}^2, \quad t = 2,\ldots,n, & (3)
\end{aligned}
$$

where y_t is a stock return, σ_t^2 is a conditional variance, $\omega > 0$, $\beta, \alpha \geq 0$, and $\beta + \alpha < 1$. The $k_t\gamma_t$ represents a jump component in the equation (2). Following Jorion (1988), the γ_t is a jump flag defined as a Bernoulli random variable such that

$$
\pi(\gamma_t = 1) = \kappa, \quad \pi(\gamma_t = 0) = 1 - \kappa, \quad 0 < \kappa < 1,
$$

and the k_t is a jump size specified as $k_t \sim N(0, \delta^2)$, where the jump parameters, κ and δ, are unknown and to be estimated. The measurement error $\sqrt{\lambda_t}\varepsilon_t$ is assumed to follow a Student-t distribution, which is a standard heavy-tailed distribution, with unknown degrees of freedom ν by letting

$$
\lambda_t^{-1} \sim \text{Gamma}(\nu/2, \nu/2).
$$

We label the model (1)–(3) the GARCHJt model. When $\lambda_t \equiv 1$ for all t, the model reduces to the GARCH model with normal errors, namely the GARCH (without jumps), or the GARCHJ model (with jumps). The GARCHt model is the one with the Student-t errors without jumps, which omits the $k_t\gamma_t$ from the equation (2). For simplicity, we assume $\log\sigma_1^2 = (\omega + \alpha\kappa\delta^2)/(1 - \alpha - \beta)$.

Next, we introduce an EGARCH$(1,1)$ model with jumps, heavy-tails formulated as

$$
\begin{aligned}
y_t &= E_{t-1}(y_t) + e_t, & (4)\\
e_t &= k_t\gamma_t + \sqrt{\lambda_t}\sigma_t\varepsilon_t, \quad \varepsilon_t \sim N(0,1), \quad t = 1,\ldots,n, & (5)\\
\log\sigma_t^2 &= \omega + \beta\log\sigma_{t-1}^2 + \theta\frac{e_{t-1}}{\sigma_{t-1}} + \alpha\left(\left|\frac{e_{t-1}}{\sigma_{t-1}}\right| - \zeta\right), \quad t = 2,\ldots,n, & (6)
\end{aligned}
$$

where $0 < \beta < 1$, $\zeta = E[|z|]$ where z is a random variable which follows a Student-t distribution with degrees of freedom ν. The EGARCH model can be interpreted as the GARCH model incorporated the leverage effect. If the coefficient θ is negative, it measures the leverage effect, which implies the increase in volatility following a previous drop in the stock return. We label the model (4)–(6) the EGARCHJt model. Similarly to the reduced models in the GARCHJt class, we consider the EGARCHJ (with normal errors and jumps), EGARCHt (with Student-t errors but without jumps), and the EGARCH (with normal errors but without jumps) model. Finally, we assume $\log \sigma_1^2 = \omega/(1 - \beta)$ for the EGARCHJt class.

2.2 MCMC algorithm

As shown in many studies, the parameter estimates of the GARCH and the EGARCH model can be obtained by the maximum likelihood estimation. In this paper, alternatively, a Bayesian inference using the MCMC algorithm is applied to provide the model comparison including the SV class whose likelihood is not easily available.

Let $y = \{y_t\}_{t=1}^n$, $\gamma = \{\gamma_t\}_{t=1}^n$, $\lambda = \{\lambda_t\}_{t=1}^n$, and $\vartheta = (\omega, \beta, \theta, \alpha)$ (θ is omitted in the case of the GARCHJt class). We set the prior probability density, $\pi(\vartheta)$, $\pi(\kappa)$, $\pi(\delta)$, and $\pi(\nu)$ for ϑ, κ, δ and ν. Deriving the posterior distribution of the GARCHJt and the EGARCHJt model,

$$\pi(\vartheta, \kappa, \delta, \nu, \gamma, \lambda | y),$$

we develop the procedure to sample from this posterior distribution by the MCMC technique as follows:

Algorithm 1: MCMC algorithm for the GARCHJt and the EGARCHJt model

1. Initialize $\vartheta, \kappa, \delta, \nu, \gamma$ and λ.

2. Sample $\vartheta | \delta, \gamma, \lambda, y$.

3. Sample $(\delta, \gamma) | \vartheta, \kappa, \lambda, y$ by

   (a) Sampling $\delta | \vartheta, \kappa, \lambda, y$,

   (b) Sampling $\gamma | \vartheta, \kappa, \delta, \lambda, y$.

4. Sample $\kappa | \vartheta, \delta, \gamma, \lambda, y$.

5. Sample $(\lambda, \nu)|\vartheta, \delta, \gamma, y$ by

    (a) Sampling $\lambda|\vartheta, \delta, \nu, \gamma, y$,

    (b) Sampling $\nu|\lambda$.

6. Go to 2.

We show the details of the MCMC algorithm in Appendix. We note that a marginalization of the conditional posterior density for some parameters enables us to accelerate the convergence of the MCMC sampling. In the algorithm, the likelihood function of the GARCHJt and the EGARCHJt model can be marginalized on the state variable for the jumps, namely k_t. In addition, the conditional posterior density of δ can be marginalized on γ in step 3(a). The performance of the algorithm is examined with simulated data below.

2.3 Simulation study for the EGARCHJt model

For simulation study of the proposed MCMC algorithm, 3,000 observations from the EGARCHJt model are generated with the parameters $\omega = -0.2$, $\beta = 0.98$, $\theta = -0.05$, $\alpha = 0.15$, $\kappa = 0.01$, $\delta = 0.03$, and $\nu = 10$. The following prior distributions are assumed:

$$\omega \sim N(0,1), \quad \beta \sim \text{Beta}(8,1), \quad \theta \sim N(0,1), \quad \alpha \sim N(0,1),$$
$$\kappa \sim \text{Beta}(2,100), \quad \delta \sim N(5,0.05), \quad \nu \sim \text{Gamma}(16,0.8).$$

These prior distributions and the parameters for simulated data reflect the values obtained in the past literature to some extents.

We draw $M = 5,000$ sample after the initial 10,000 sample are discarded. The computational results are generated using Ox version 4.02 (Doornik (2006)). Figure 1 shows the sample autocorrelation functions, the sample paths and the posterior densities for each parameter. After discarding sample in burn-in period, the sample paths look stable and the sample autocorrelations drop quickly. This indicates that our sampling method produces the uncorrelated sample efficiently.

Table 1 gives the estimates for posterior means, standard deviations and the 95% credible intervals. All estimated posterior means are close to the true values and the true values are contained in the 95% credible intervals. The inefficiency factors are also reported to check the performance of our sampling efficiency. The inefficiency factor is defined as $1+2\sum_{s=1}^{\infty}\rho_s$ where ρ_s is the sample autocorrelation function at lag s. It is the ratio of variance of the posterior
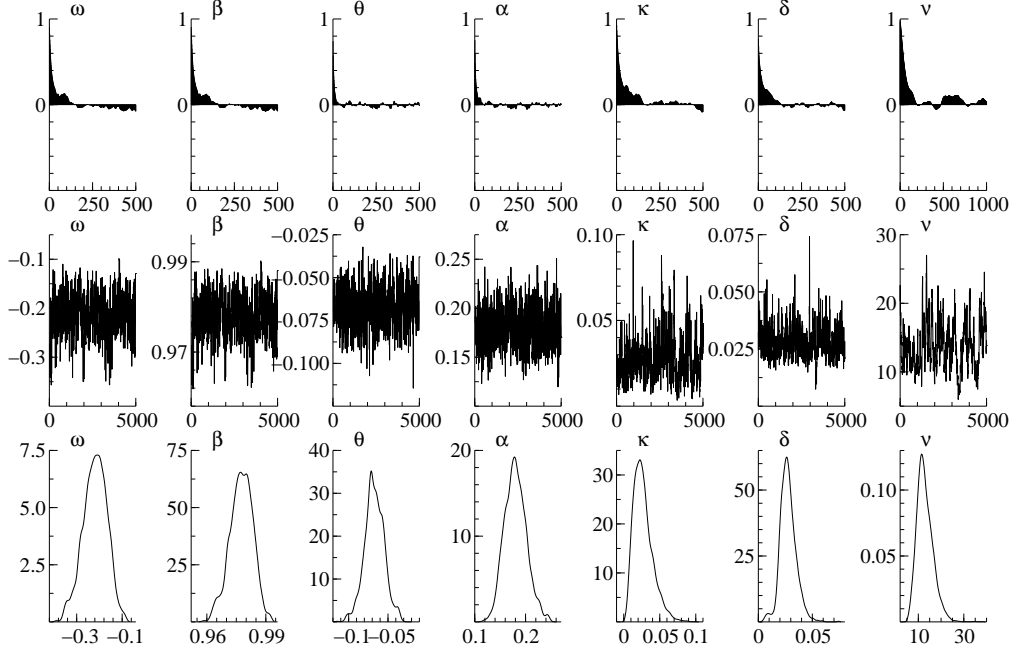
Figure 1: Estimation result of the EGARCHJt model for the simulated data. Sample autocorrelations (top), sample paths (middle) and posterior densities (bottom).

Parameter	True	Mean	Stdev.	95% interval	Inefficiency
ω	-0.2	-0.2194	0.0524	[-0.3338, -0.1244]	34.17
β	0.98	0.9775	0.0058	[0.9649, 0.9881]	34.72
θ	-0.05	-0.0673	0.0125	[-0.0924, -0.0411]	7.57
α	0.15	0.1801	0.0215	[0.1411, 0.2261]	12.11
κ	0.01	0.0270	0.0134	[0.0077, 0.0595]	64.97
δ	0.03	0.0276	0.0074	[0.0137, 0.0436]	38.63
ν	10.0	12.9531	3.5755	[7.2511, 20.921]	122.87

Table 1: Estimation result of the EGARCHJt model for the simulated data.

7

mean from the correlated draws to the one from the hypothetical uncorrelated sample, which measures the loss of sampling efficiency in our correlated MCMC draws (see e.g., Chib (2001)). In the estimation, it is computed with a bandwidth 500. The estimated inefficiency factors in Table 1 are low enough, which assures the successful sampling without loss of efficiency.

2.4 Alternative jump specification

As an alternative model to incorporate the jump components into the EGARCH model, it would be possible to formulate the model as

$$
\begin{aligned}
y_t &= E_{t-1}(y_t) + e_t, \\
e_t &= k_t \gamma_t + \sigma_t \varepsilon_t, \quad \varepsilon_t \sim N(0,1), \quad t = 1, \ldots, n, \\
\log \sigma_t^2 &= \omega + \beta \log \sigma_{t-1}^2 + \theta \varepsilon_{t-1} + \alpha(|\varepsilon_{t-1}| - E(|\varepsilon_{t-1}|)), \quad t = 2, \ldots, n,
\end{aligned}
$$

where $\varepsilon_{t-1} = (e_{t-1} - k_{t-1}\gamma_{t-1})/\sigma_{t-1}$. Here, the jump component does not affect the volatility process. In this specification, however, the conditional posterior distributions for jump variables are not easily computed because the value of the jump variables, k_t and γ_t, are state variables to be sampled in the MCMC algorithm, while the k_t and γ_t affect all the volatility from time t to n. The draw of $\{k_t\}_{t=1}^n$ and $\{\gamma_t\}_{t=1}^n$ requires so much time that it would be almost unfeasible to implement the MCMC procedure. Because of this difficulty, we choose the specification of equation (5) and (6) for the EGARCHJt model in this paper.

3 Bayesian inference for the SV model with leverage, jumps and heavy-tails

3.1 The model

We consider a discrete-time SV model formulated as

$$
\begin{aligned}
y_t &= E_{t-1}(y_t) + e_t, & (7) \\
e_t &= k_t \gamma_t + \sqrt{\lambda_t} \varepsilon_t \exp(h_t/2), \quad t = 1, \ldots, n, & (8) \\
h_{t+1} &= \mu + \phi(h_t - \mu) + \eta_t, \quad t = 1, \ldots, n-1, & (9)
\end{aligned}
$$

where h_t is an unobserved log-volatility, $|\phi| < 1$, $h_1 \sim N(0, \sigma^2/(1 - \phi^2))$,

$$\begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix} \sim N(0, \Sigma), \quad \text{and} \quad \Sigma = \begin{pmatrix} 1 & \rho\sigma \\ \rho\sigma & \sigma^2 \end{pmatrix}.$$

The correlation coefficient ρ measures the leverage effect and $\rho = 0$ implies the SV model without leverage effect. We introduce a jump component $k_t\gamma_t$ in the measurement equation (8). The γ_t is a jump flag defined as a Bernoulli random variable defined in the previous section, and the k_t is a jump size specified by

$$\psi_t \equiv \log(1 + k_t) \sim N(-0.5\delta^2, \delta^2), \tag{10}$$

following Andersen et al. (2002), Chib et al. (2002). Though this specification of the jump size is different from the one incorporated into the GARCHJt and the EGARCHJt model in the previous section, the distribution (10) is derived from a discretization of a Lévy process, which is used in the continuous time modeling of financial asset pricing. We label the model defined by equation (8) and (9) as the SVLJt model and consider the following reduced models; SVLJ (with normal errors and jumps), SVLt (with Student-t errors but without jumps), the SVL (with normal errors but without jumps) model and the model without leverage for each specification, namely the SVJt, the SVJ, SVt and SV model, respectively. For simplicity, we compute y_t by a log-return of the stock price, demean it by the sample mean, and assume that $E_{t-1}(y_t) \equiv 0$, throughout the paper.

3.2 Auxiliary mixture sampler

Following Omori et al. (2007), we define $y_t^* = \log(y_t - k_t\gamma_t)^2 - \log \lambda_t$, $d_t = \text{sign}(y_t - k_t\gamma_t) = I(\varepsilon_t > 0) - I(\varepsilon_t \leq 0)$, which rewrite the equation (7) and (8) as

$$y_t^* = h_t + \xi_t, \tag{11}$$

where $\xi_t = \log \varepsilon_t^2$. Omori et al. (2007) propose to approximate the bivariate conditional density of $(\xi_t, \eta_t)|d_t$ by a ten-component mixture of bivariate normal distribution, which is an exhaustive extension of Kim et al. (1998) approach. The key essence of their approach is that the model (11) and (9) can be approximated to a linear Gaussian state space model

i	p_i	m_i	v_i^2	a_i	b_i
1	0.00609	1.92677	0.11265	1.01418	0.50710
2	0.04775	1.34744	0.17788	1.02248	0.51124
3	0.13057	0.73504	0.26768	1.03403	0.51701
4	0.20674	0.02266	0.40611	1.05207	0.52604
5	0.22715	-0.85173	0.62699	1.08153	0.54076
6	0.18842	-1.97278	0.98583	1.13114	0.56557
7	0.12047	-3.46788	1.57469	1.21754	0.60877
8	0.05591	-5.55246	2.54498	1.37454	0.68728
9	0.01575	-8.68384	4.16591	1.68327	0.84163
10	0.00115	-14.65000	7.33342	2.50097	1.25049

Table 2: Selection of $(p_i, m_i, v_i^2, a_i, b_i)$ proposed by Omori et al. (2007).

conditioned on the mixture component indicator $s_t \in \{1, 2, \ldots, K\}$ as

$$\begin{pmatrix} y_t^* \\ h_{t+1} \end{pmatrix} = \begin{pmatrix} h_t \\ \mu + \phi(h_t - \mu) \end{pmatrix} + \begin{pmatrix} \xi_t \\ \eta_t \end{pmatrix}, \qquad (12)$$

where

$$\left\{ \begin{pmatrix} \xi_t \\ \eta_t \end{pmatrix} \Big| d_t, (s_t = i) \right\} \overset{L}{=} \begin{pmatrix} m_i + v_i z_{1t} \\ d_t \rho \sigma (a_i + b_i v_i z_{1t}) \exp(m_i/2) + \sigma \sqrt{1 - \rho^2} z_{2t} \end{pmatrix},$$

for $i = 1, 2, \ldots, K$, and $z_t = (z_{1t}, z_{2t})' \sim N(0, I_2)$. Given $s = \{s_1, \ldots, s_n\}$, we can sample the latent variable $h = \{h_1, \ldots, h_n\}$ in one block from its joint distribution using the simulation smoother for a linear Gaussian state space model (de Jong and Shephard (1995), Durbin and Koopman (2002)). The mixture component parameters are provided by Omori et al. (2007) in the case of $K = 10$ (reproduced in Table 2). Note that (m_i, v_i, a_i, b_i) do not depend on model parameters, $\theta \equiv (\phi, \sigma, \rho)$ and μ.

3.3 MCMC algorithm

Let $y^* = \{y_t^*\}_{t=1}^n$, $d = \{d_t\}_{t=1}^n$, $k = \{k_t\}_{t=1}^n$, and we set the prior probability density $\pi(\theta)$, $\pi(\mu)$, $\pi(\kappa)$, $\pi(\delta)$, and $\pi(\nu)$ for θ, μ, κ, δ, and ν. Then, we draw sample from the posterior distribution

$$\pi(\theta, \mu, \kappa, \delta, \nu, s, h, k, \gamma, \lambda | y),$$

10

by the MCMC algorithm. Let us reparameterize k_t by $\psi_t \equiv \log(1 + k_t)$ and denote $\psi = \{\psi_t\}_{t=1}^n$, $\psi^{(0)} = \{\psi_t | t = 1, \ldots, n, \text{ s.t. } \gamma_t = 0\}$, $\psi^{(1)} = \{\psi_t | t = 1, \ldots, n, \text{ s.t. } \gamma_t = 1\}$. Following Omori et al. (2007), we use the following sampling algorithm.

Algorithm 2: MCMC algorithm for the SVLJt model

1. Initialize $\theta, \mu, \kappa, \delta, \nu, s, h, \psi, \gamma$ and λ.

2. Sample $(\theta, \mu, h) | s, y^*, d$ by

    (a) Sampling $\theta | s, y^*, d$,

    (b) Sampling $(\mu, h) | \theta, s, y^*, d$.

3. Sample $\psi^{(1)} | \theta, \mu, \delta, h, \gamma, \lambda, y$.

4. Sample $(\delta, \psi^{(0)}) | \psi^{(1)}, \gamma$ by

    (a) Sampling $\delta | \psi^{(1)}, \gamma$,

    (b) Sampling $\psi^{(0)} | \delta, \gamma$.

5. Sample $(\gamma, s) | \theta, \mu, \kappa, h, \psi, \lambda, y$ by

    (a) Sampling $\gamma | \theta, \mu, \kappa, h, \psi, \lambda, y$,

    (b) Sampling $s | \theta, \mu, h, y^*, d$.

6. Sample $\kappa | \gamma$.

7. Sample $(\lambda, \nu) | \theta, \mu, s, h, \psi, \gamma, y$ by

    (a) Sampling $\lambda | \theta, \mu, \nu, s, h, \psi, \gamma, y$,

    (b) Sampling $\nu | \lambda$.

8. Go to 2.

The details of the algorithm is developed by Nakajima and Omori (2009). They provide a simulation study, which shows an efficient performance of the MCMC algorithm for the SVLJt model.

3.4 Alternative jump specification

The EGARCHJt model defined by equation (5) and (6) has the jumps which affect the volatility process, while the SVLJt model defined by equation (8) and (9) has the jumps which do not affect the volatility. As mentioned in section 2.4, it is unfeasible to estimate the EGARCHJt model with jumps which do not affect the volatility. Alternatively, we consider another SV model with jumps which do affect the volatility in order to compare with the EGARCHJt class.

We consider the SV model with correlated jumps (we SVLCJ model) given by

$$
\begin{aligned}
y_t &= E_{t-1}(y_t) + e_t, \\
e_t &= k_t \gamma_t + \varepsilon_t \exp(h_t/2), \quad t = 1, \ldots, n, \\
h_{t+1} &= \mu + \phi(h_t - \mu) + j_t \gamma_t + \eta_t, \quad t = 1, \ldots, n-1.
\end{aligned}
$$

(13)

(14)

The equations (13) and (14) have a common jump indicator variable, γ_t, to model the jumps that occur concurrently both in return and in volatility so that the jumps affect the volatility process. The joint distribution of jump sizes is assumed to be

$$
\begin{aligned}
j_t &\sim \text{Exp}(\mu_J), \\
k_t | j_t &\sim N(\mu_k + \beta_J j_t, \sigma_k^2),
\end{aligned}
$$

where Exp denotes the exponential distribution. The correlation between jump sizes in return and in volatility is considered by the parameter β_J. This type of jumps in the SV model is studied in the recent literature (e.g., Eraker et al. (2003), Kobayashi (2006)). Nakajima and Omori (2009) compare the SVLCJ model with the models in the SVLJt class. We also include the SVLCJ model for the model comparison in this paper.

4 Application to stock return data

4.1 Data

We estimate the models in the EGARCHJt and the SVLJt class for daily stock returns. The series are two US stock price indices; S&P500 and NASDAQ; and two US individual stocks; GM (General Motors) and IBM (International Business Machines). The sample period is from January 1992 to December 2006. The log-difference returns are computed as $y_t = \log P_t -$

$\log P_{t-1}$, and demeaned, where P_t is the closing price on the business day t. The sample size is 3,781 for each series. Table 3 summarizes the descriptive statistics and Figure 2 plots the four series of daily return. Note that the statistics are based on the return data before demeaned. Regarding the higher order moments, in Table 3, the skewness of the S&P500 series is negative, while two individual stocks have slightly positive skewnesses. The kurtosis of each series is around seven to nine, which is clearly larger than the one of a normal distribution.

	Mean	Stdev.	Skewness	Kurtosis	Max.	Min.
S&P500	0.0003	0.010	-0.109	7.200	0.056	-0.071
NASDAQ	0.0004	0.016	0.002	8.616	0.133	-0.102
GM	0.0002	0.021	0.083	6.739	0.166	-0.151
IBM	0.0004	0.020	0.017	9.688	0.124	-0.169

Table 3: Summary statistics for the daily stock return data (1992/Jan – 2006/Dec, $n = 3,781$).
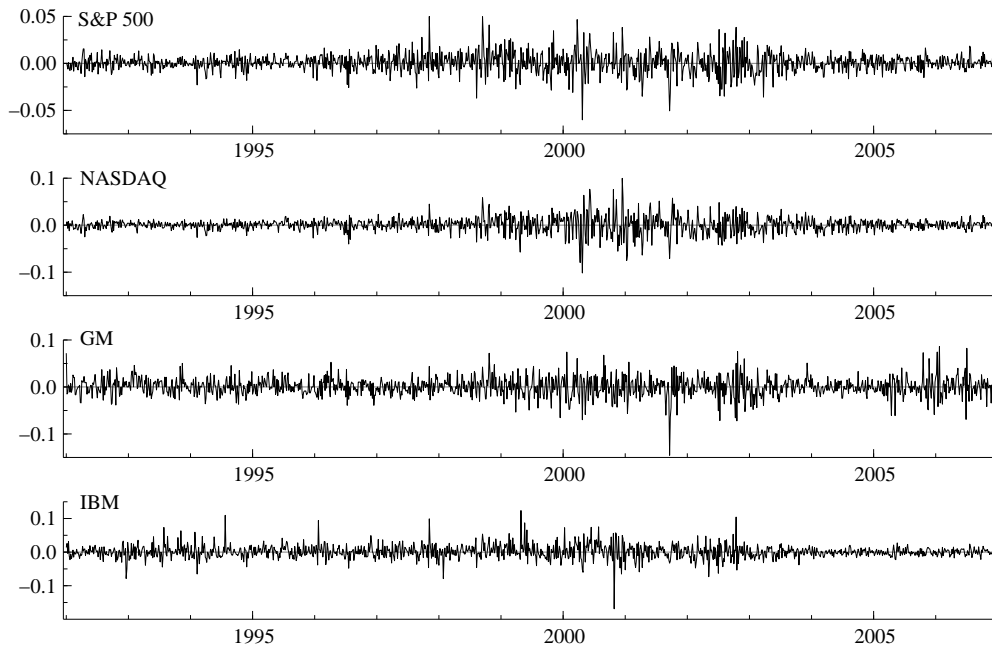

Figure 2: The time-series plots for four stock returns (1992/Jan – 2006/Dec).

4.2 Parameter estimates

We report the results for the parameter estimation of the EGARCHJt and the SVLJt model for the S&P500 series. The priors for the GARCHJt and the EGARCHJt class, the same settings are used as Section 2.3. For the SVLJt class, we assume the following prior:

$$\frac{\phi + 1}{2} \ \sim \ \text{Beta}(20, 1.5), \quad \sigma^{-2} \sim \text{Gamma}(2.5, 0.025),$$
$$\rho \ \sim \ U(-1, 1), \quad \mu \sim N(-10, 1), \quad \kappa \sim \text{Beta}(2, 100),$$
$$\log(\delta) \ \sim \ N(-2.5, 0.15), \quad \nu \sim \text{Gamma}(16, 0.8).$$

As suggested by Kim et al. (1998) for the estimation of the SV model using the mixture sampler, we take $y_t^* = \log((y_t - k_t \gamma_t)^2 + c)$, where c is an offset for the case where $(y_t - k_t \gamma_t)^2$ is too small. We set $c = 10^{-7}$ in this paper. The number of MCMC iterations is same as the simulation study.

Table 4 reports the parameter estimates of the EGARCHJt and the SVLJt model for the S&P500 returns. Figure 3 and 4 plot the sampling results for the EGARCHJt and the SVLJt model respectively. The estimates of the volatility parameters $(\omega, \beta, \theta, \alpha)$ and $(\phi, \sigma, \rho, \exp(\mu/2))$ are consistent with the results of the previous literature (e.g., Vrontos et al. (2000), Nakajima and Omori (2009)). For both two models, the posterior means of β and ϕ are close to one, which implies a well-known high persistence of volatility on stock returns. The parameters θ and ρ are estimated negative and the 95% credible intervals do not contain zero. This indicates that there exists the leverage effect in our stock return data.

We find that there are specific differences for the estimates of the jump parameters between the EGARCHJt and the SVLJt model. The posterior mean of κ for the EGARCHJt model is about 3%, while the one for the SVLJt model is much smaller, 0.09%. This indicates that the EGARCHJt model more often captures the excess returns by the jump component than the SVLJt model. Though we can not compare the jump sizes between the two models directly because the specifications of the jump size are different, if we calculate the standard deviation of the jump size for the SVLJt model using the posterior mean of δ, it is 0.0996 for the SVLJt model and 0.0098 for the EGARCHJt model. The empirical results show that the SVLJt model captures larger excess returns by the jump component with the smaller probability than the EGARCHJt model. This would be caused from the difference of the specification of the volatility process and the jump component between the two models. The SV models have the disturbance for their volatility process, while the EGARCH models do not have it and the

Parameter	Mean	Stdev.	95% interval	Inefficiency
EGARCHJt model				
ω	-0.1299	0.0223	[-0.1740, -0.0895]	19.20
β	0.9879	0.0022	[0.9834, 0.9920]	19.14
θ	-0.0880	0.0103	[-0.1097, -0.0693]	10.79
α	0.1051	0.0116	[0.0836, 0.1279]	9.74
κ	0.0307	0.0182	[0.0064, 0.0730]	49.33
δ	0.0098	0.0027	[0.0058, 0.0166]	22.36
ν	13.9744	2.8954	[9.7847, 20.816]	59.79
SVLJt model				
ϕ	0.9864	0.0029	[0.9800, 0.9916]	12.54
σ	0.1373	0.0134	[0.1119, 0.1648]	13.91
ρ	-0.7327	0.0412	[-0.8044, -0.6450]	10.54
$\exp(\mu/2)$	0.0090	0.0006	[0.0079, 0.0103]	1.77
κ	0.0009	0.0007	[0.0001, 0.0026]	3.80
δ	0.0993	0.0525	[0.0334, 0.2192]	62.60
ν	20.8137	4.1579	[14.505, 30.931]	175.81

Table 4: Estimation results for S&P500 returns.

volatility on the next day is determined by the return and the volatility on the current day. In other words, the volatility process of the SV models can move more flexibly than the EGARCH models. In short, when the return marks a certain excess shock, if the effect of the shock is not persistent, the volatility process of the SVLJt model would capture the shock, while the EGARCHJt would capture it by the jump component.

In addition, the posterior mean of the parameter ν for the EGARCHJt model is smaller than the one for the SVLJt model. As discussed by Nakajima and Omori (2009), the model whose jump probability is estimated to be smaller tends to have a heavier-tailness of the errors. It is considered that the jump component captures less excess returns, when the errors have the heavier-tails.

4.3 Model comparisons

In a Bayesian framework, we can compare a model fit based on a marginal likelihood or a Bayes factor. When the prior probabilities are assumed to be equal, we choose the model which yields the largest marginal likelihood. In order to compare the competing models in the GARCHJt, the EGARCHJt and the SVLJt class, we estimate their marginal likelihood for the return data.

The marginal likelihood is defined as the integral of the likelihood with respect to the prior density of the parameter. Following Chib (1995), we estimate the log of marginal likelihood,
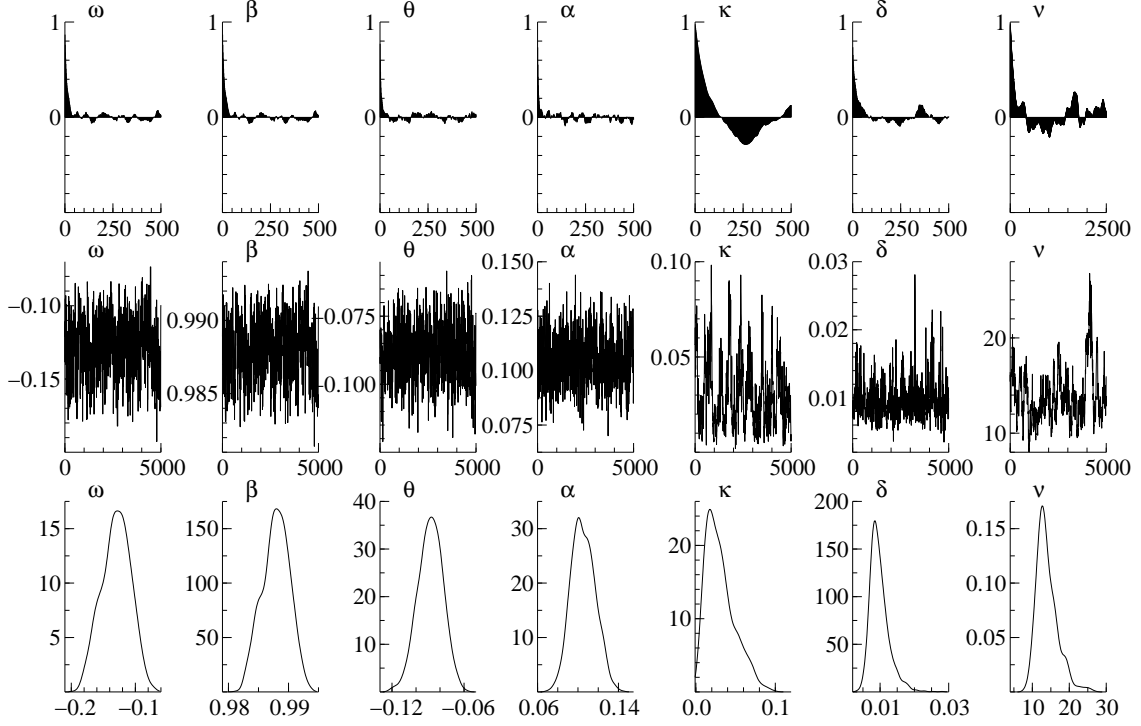
Figure 3: Estimation result of the S&P500 returns (EGARCHJt model). Sample autocorrelations (top), sample paths (middle) and posterior densities (bottom).

denoted by $m(y)$, as

$$\log m(y) = \log f(y|\Theta) + \log \pi(\Theta) - \log \pi(\Theta|y),$$

where Θ is a parameter set in the model, $f(y|\Theta)$ is a likelihood, $\pi(\Theta)$ is a prior probability density and $\pi(\Theta|y)$ is a posterior density. This equality holds for any Θ, but we usually use the posterior mean of Θ to obtain a stable estimate of $m(y)$. The prior probability density is easily calculated, though the likelihood and posterior part requires a simulation evaluation. For the SVLJt class, the likelihood can be estimated by the particle filter (e.g., Pitt and Shephard (1999), Chib et al. (2002), Omori et al. (2007)). For the posterior part, we use the method of Chib (1995), Chib and Jeliazkov (2001, 2005) to compute $\pi(\Theta|y)$ using the sample obtained through the reduced iteration of the MCMC algorithm.

We show the results of the model comparison for the 17 competing models of two classes including the SVLCJ model, for four stock return data. For the GARCHJt class, we assume
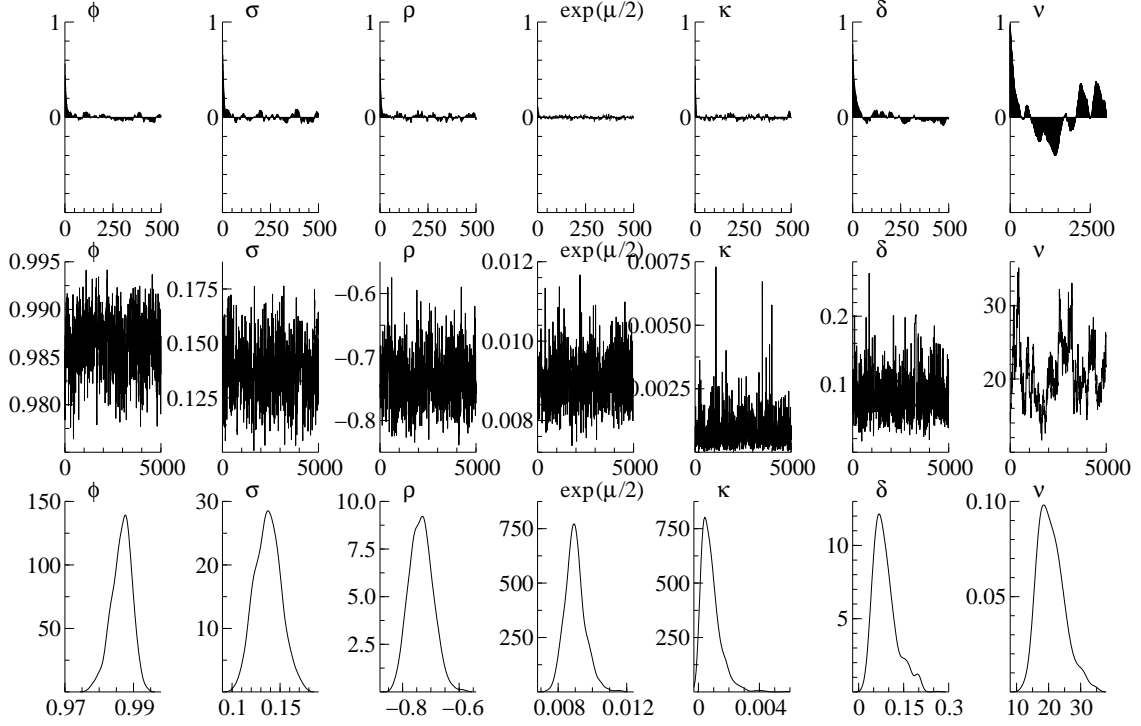
Figure 4: Estimation result of the S&P500 returns (SVLJt model). Sample autocorrelations (top), sample paths (middle) and posterior densities (bottom).

the priors as following:

$$
\begin{aligned}
\omega^{-1} &\sim \text{Gamma}(5, 5 \times 10^{-4}), \quad \alpha \sim \text{Beta}(3, 3), \quad \beta(1 - \alpha)|\alpha \sim \text{Beta}(8, 1), \\
\theta &\sim U(0, 1 - \alpha - \beta), \qquad \kappa \sim \text{Beta}(2, 100), \\
\delta &\sim N(5, 0.05), \qquad \nu \sim \text{Gamma}(16, 0.8).
\end{aligned}
$$

For the SVLCJ model, we assume the following priors:

$$
\begin{aligned}
\mu_J &\sim \text{Exp}(0.2), \quad \beta_J \sim N(0, 1), \\
\mu_k &\sim N(0, 1), \quad \sigma_k^{-2} \sim \text{Gamma}(2.5, 0.025).
\end{aligned}
$$

We evaluate the posterior density at the posterior mean for Θ through the reduced MCMC sampling, which is iterated for 5,000 draws. We run ten replications of the particle filter to estimate the standard error of the likelihood.

17

	S&P500		NASDAQ	
Model	Log-ML	Ranking	Log-ML	Ranking
GARCH	12578.42 (0.01)	17	11300.17 (0.01)	17
GARCHt	12621.00 (0.26)	14	11321.13 (0.18)	14
GARCHJ	12606.61 (0.02)	16	11317.49 (0.04)	16
GARCHJt	12618.57 (0.22)	15	11319.88 (0.19)	15
EGARCH	12661.70 (0.09)	9	11345.23 (0.06)	11
EGARCHt	12702.52 (0.68)	3	11367.52 (0.93)	5
EGARCHJ	12692.14 (0.12)	7	11369.65 (0.11)	3
EGARCHJt	12703.09 (0.46)	2	11368.57 (0.61)	4
SV	12650.05 (0.43)	12	11349.39 (0.76)	10
SVt	12660.25 (0.80)	10	11350.26 (1.18)	9
SVJ	12644.53 (0.78)	13	11334.23 (2.51)	13
SVJt	12653.52 (0.80)	11	11344.76 (0.94)	12
SVL	12702.21 (0.43)	4	11374.48 (0.83)	2
SVLt	12708.90 (0.66)	1	11375.05 (0.98)	1
SVLJ	12697.26 (1.14)	6	11358.97 (2.42)	7
SVLJt	12702.03 (0.79)	5	11363.84 (0.91)	6
SVLCJ	12691.27 (1.02)	8	11358.49 (0.78)	8

	GM		IBM	
Model	Log-ML	Ranking	Log-ML	Ranking
GARCH	9385.76 (0.01)	17	9937.35 (0.01)	17
GARCHt	9492.06 (0.81)	5	10035.42 (0.57)	9
GARCHJ	9477.39 (0.06)	11	10026.50 (0.03)	12
GARCHJt	9481.35 (0.59)	8	10030.32 (0.45)	11
EGARCH	9401.47 (0.02)	16	9974.36 (0.01)	16
EGARCHt	9493.83 (0.66)	2	10046.06 (0.55)	6
EGARCHJ	9479.09 (0.08)	10	10037.06 (0.03)	8
EGARCHJt	9492.36 (0.70)	3	10038.00 (0.27)	7
SV	9471.54 (0.40)	13	10022.87 (0.45)	14
SVt	9475.38 (0.54)	12	10060.31 (1.24)	4
SVJ	9468.12 (1.02)	15	10019.30 (1.38)	15
SVJt	9469.80 (0.92)	14	10049.99 (1.21)	5
SVL	9490.23 (0.22)	6	10034.99 (0.43)	10
SVLt	9506.58 (0.62)	1	10074.95 (1.64)	1
SVLJ	9479.82 (1.19)	9	10025.63 (1.45)	13
SVLJt	9492.25 (1.00)	4	10066.63 (1.86)	2
SVLCJ	9487.94 (1.32)	7	10063.01 (1.97)	3

* The values are based on log scale and standard error in parentheses.

Table 5: Marginal likelihood (ML) for stock return data.

Table 5 reports the estimated marginal likelihoods, standard errors and rankings for all the competing models. The estimates show that the marginal likelihoods of the EGARCH models are higher than the GARCH models, and the ones of the SV models with leverage are higher than the SV models without leverage. This indicates that the leverage effect is important to analyze the stock returns as discussed by Omori et al. (2007).

The models in the EGARCHJt class perform well. For example, the EGARCHt and the EGARCHJt models outperform the SVLJ and SVLJt models for the S&P500, the NASDAQ and the GM return data. Compared with the SVLCJ model, whose jumps do affect the volatility as the EGARCHJt class, the EGARCHJ and the EGARCHJt models outperform the SVLCJ model for the S&P500 and the NASDAQ return data.

Overall, the marginal likelihood of the SVLt model is the highest among the competing models. As pointed out by Nakajima and Omori (2009), the SVLt model performs better than the other models in the SVLJt class, and we find that it is also favored over the EGARCHJt models.

The heavy-tails contribute most of the models for their marginal likelihoods. Most of the EGARCH models with Student-t errors are favored over the ones without them. On the other hand, the jumps do not always contribute these models as discussed by Nakajima and Omori (2009). Overall, we find that the jumps and the heavy-tails have the large contributions for the EGARCH models. The ratio of the marginal likelihood of the SVL model to the one of the EGARCH model is quite large, while the one of the SVLJt model to the EGARCHJt model is less than one, which implies the EGARCHJt model outperforms the SVLJt model by incorporating the jumps and the heavy-tails for the S&P500, the NASDAQ and the GM return data.

5   Application to exchange rate data

5.1   Data

In this section, we estimate the EGARCHJt and the SVLJt classes for daily exchange rate returns. We use two exchange rates series; UK Sterling and Japanese Yen against US Dollar. The sample period is from October 1986 to August 1996. The returns are computed as $y_t = P_t/P_{t-1} - 1$, and demeaned, where P_t is the closing price on the business day t. The sample size is 2,566 for each series. Table 6 summarizes the descriptive statistics (based on data before demeaned) and Figure 5 plots the two series of daily exchange rate return. It is interesting that UK Sterling has a negative skewness, while Japanese Yen has a positive one over the

sample period. In literature, it is often argued that empirical return distribution of stock price is negatively skewed because of the heavier tail on its left side, which theoretically is connected to the market player's preference. On the other hand, the exchange rate is dealt by two-side players, therefore the skewness of the return distribution is considered to depend on the currency and sample period. The kurtosis of each series is around five to six, which is larger than the one of a normal distribution.

	Mean	Stdev.	Skewness	Kurtosis	Max.	Min.
UK Sterling	0.0001	0.006	-0.181	5.327	0.031	-0.030
Japanese Yen	0.0002	0.007	0.217	6.236	0.034	-0.033

Table 6: Summary statistics for the daily exchange rate return data (1986/Oct – 1996/Aug, $n = 2,566$).
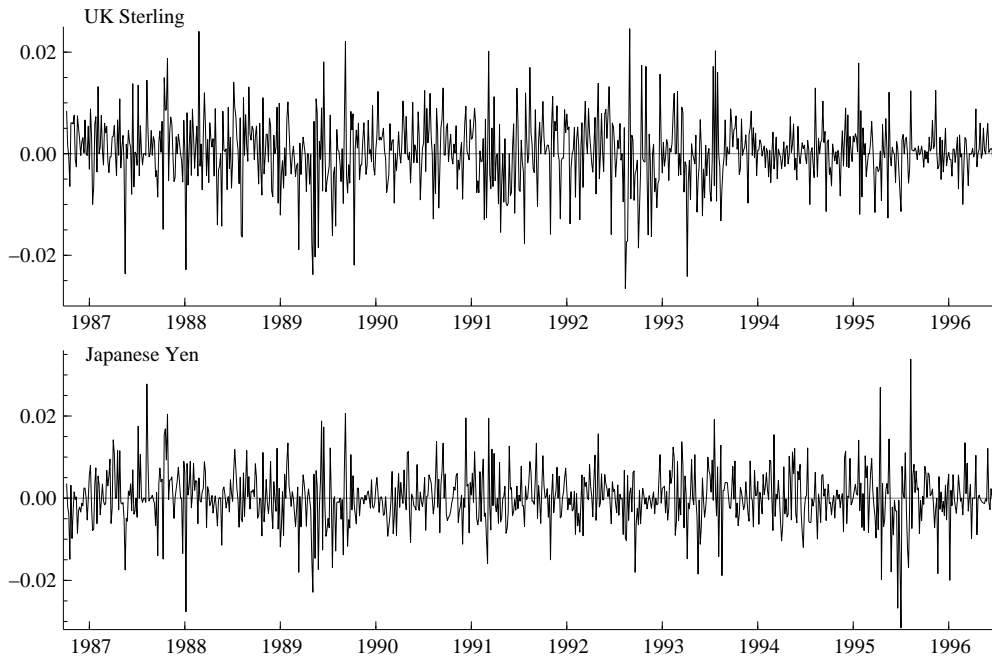

Figure 5: The time-series plots for returns of two exchange rates against US Dollar (1986/Oct – 1996/Aug).

5.2 Model Comparison

We estimate the marginal likelihood of the GARCH and the SV classes for the daily return of exchange rate series. The computational settings are same as the previous section.

Model	UK Sterling Log-ML	Ranking	Japanese Yen Log-ML	Ranking
GARCH	9495.02 (0.01)	17	9425.15 (0.01)	17
GARCHt	9497.09 (0.21)	16	9435.99 (0.25)	14
GARCHJ	9498.73 (0.95)	14	9438.23 (1.02)	10
GARCHJt	9498.20 (0.48)	15	9436.84 (1.32)	12
EGARCH	9502.19 (0.04)	13	9430.55 (0.85)	16
EGARCHt	9509.22 (0.20)	11	9436.24 (0.90)	13
EGARCHJ	9507.15 (1.43)	12	9435.30 (1.42)	15
EGARCHJt	9510.24 (1.14)	10	9439.95 (1.44)	8
SV	9525.02 (0.32)	6	9445.37 (1.24)	5
SVt	9548.70 (1.06)	2	9454.27 (1.13)	2
SVJ	9519.02 (0.39)	8	9439.00 (0.98)	9
SVJt	9543.17 (1.91)	4	9449.66 (1.35)	4
SVL	9524.73 (0.51)	7	9441.26 (0.79)	6
SVLt	9549.73 (1.16)	1	9454.65 (0.57)	1
SVLJ	9518.35 (0.36)	9	9437.35 (1.08)	11
SVLJt	9544.53 (1.15)	3	9453.97 (1.48)	3
SVLCJ	9538.12 (1.95)	5	9440.48 (0.99)	7

* The values are based on log scale and standard error in parentheses.

Table 7: Marginal likelihood (ML) for exchange rate return data.

Table 7 reports the estimated marginal likelihoods for 17 competing models. Clearly, the SV models are favoured over the GARCH models for both return data. In the SV class, the SVt model yields the best performance and the next is interestingly not the SV models with leverage but the SVt model (without leverage). For the stock return data examined in the previous section, the non-leverage SV models have less performance at all compared to the leverage SV models. However, for the exchange rate return data, non-leverage SV models are favoured almost in the same level as the leverage SV models. It is probably because there are two-side participants in the currency market. The asymmetry between the disturbances of observation equation and volatility process, which is refeed as leverage effect here, would depend on the difference of dealer's preference, economic condition and price level between two countries, and moreover it would totally depend on the sample period to estimate. Table 8 shows the posterior estimates of the SVLt model for the UK Sterling series. In fact, the posterior mean for the parameter ρ is -0.1227 and the 95% credible intervals contain zero,

Parameter	Mean	Stdev.	95% interval	Inefficiency
ϕ	0.9831	0.0062	[0.9676, 0.9928]	15.66
σ	0.1238	0.0221	[0.0852, 0.1686]	21.96
ρ	-0.1227	0.0996	[-0.3043, 0.0819]	1.31
$\exp(\mu/2)$	0.0048	0.0004	[0.0040, 0.0056]	12.36
ν	5.9712	0.7682	[4.7057, 7.4455]	109.17

Table 8: Estimation results of the SVLt model for UK Sterling returns against US Dollar.

which indicates little evidence for the leverage effect. This result implies that the leverage effect plays the important role for stock return, while not so much for exchange rate return.

The jump models are not favoured over the no-jump models overall for both the GARCH and the SV models. The kurtosis of the empirical return distribution is higher than the one of the normal, while the estimation results indicate that it is enough to incorporate the heavy-tailed error distribution on our dataset.

6 Robustness check

In this section, additional model comparisons for the subsample periods and different priors are examined as a robustness check. First, the sample period of the S&P500 return data is divided into two subsample periods; the first-half ($n = 1,891$) and the second-half ($n = 1,890$) period. From the discussion in the previous section, the leverage effect is found to contribute the model fit clearly. Thus, for the estimation of the robustness checks, the models without leverage effect are omitted and the marginal likelihoods of the models with leverage effect are estimated for the two subsample periods.

Table 9 reports the estimated marginal likelihoods for the subsample periods of the S&P500 return data. Still, the marginal likelihood of the SVLt model is the highest for both subsample periods. The second (EGARCHJt) and the third (EGARCHt) best models are unchanged, although the following ranking is slightly changed. For the first-half subsample period, the SV models with jumps (SVLJ and SVLJt) are relatively highly favored, although they are outperformed by the SVL model for the second-half subsample period.

Second, a Bayesian model comparison requires a prior sensitivity analysis to check a robustness of estimation results. Two different priors are examined for the model comparison using the S&P500 return data. Let the priors used in the previous sections denoted the Prior1. The different priors are specified as follows:

	1st.half		2nd.half	
Model	Log-ML	Ranking	Log-ML	Ranking
EGARCH	6588.46 (0.03)	9	6088.73 (0.02)	9
EGARCHt	6602.91 (0.33)	3	6105.43 (0.87)	3
EGARCHJ	6595.31 (0.06)	6	6094.79 (0.31)	7
EGARCHJt	6603.43 (0.43)	2	6105.94 (0.43)	2
SVL	6593.46 (0.38)	8	6102.44 (0.98)	4
SVLt	6608.52 (0.68)	1	6110.17 (0.72)	1
SVLJ	6597.01 (0.96)	5	6096.19 (1.34)	6
SVLJt	6597.43 (0.75)	4	6099.16 (1.03)	5
SVLCJ	6595.20 (0.76)	7	6092.08 (1.21)	8

* The values are based on log scale and standard error in parentheses.

Table 9: Marginal likelihood (ML) for the S&P500 return data (subsample period).

Prior2

$$\beta \sim \text{Beta}(4,1), \quad \theta \sim N(-0.08, 0.5), \quad \kappa \sim \text{Beta}(2, 100),$$

$$\delta \sim \text{Gamma}(5, 0.05), \quad \nu \sim \text{Gamma}(16, 0.8), \text{ for EGARCHJt},$$

$$\sigma^{-2} \sim \text{Gamma}(5, 0.05), \quad \mu \sim N(-10, 2), \quad \kappa \sim \text{Beta}(2, 100),$$

$$\log(\delta) \sim N(-2.5, 0.15), \quad \nu \sim \text{Gamma}(16, 0.8), \text{ for SVLJt}.$$

Prior3

$$\beta \sim \text{Beta}(4,1), \quad \theta \sim N(0, 1), \quad \kappa \sim \text{Beta}(1, 100),$$

$$\delta \sim \text{Gamma}(10, 0.2), \quad \nu \sim \text{Gamma}(20, 0.5), \text{ for EGARCHJt},$$

$$\sigma^{-2} \sim \text{Gamma}(5, 0.1), \quad \mu \sim N(-10, 1), \quad \kappa \sim \text{Beta}(1, 100),$$

$$\log(\delta) \sim N(-2.5, 0.4), \quad \nu \sim \text{Gamma}(20, 0.5), \text{ for SVLJt}.$$

The priors not mentioned here are specified same as the previous sections. The marginal likelihoods of the competing models are estimated for the full sample period of the S&P500 return data.

Table 10 reports the estimated marginal likelihoods for the two different priors. Again, the SVLt model best fits the data under both priors. Overall, the heavy-tailed model such as the EGARCHt, the EGARCHJt and the SVLJt models are ranked high. Under the Prior2, the EGARCHt and the EGARCHJt models outperform the SV models except the SVLt model. On the other hand, the SVL and SVLJt model outperform the EGARCH models under the

| | **Prior2** | | **Prior3** | |
Model	Log-ML	Ranking	Log-ML	Ranking
EGARCH	12671.20 (0.09)	9	12669.93 (0.04)	9
EGARCHt	12704.10 (0.39)	2	12702.77 (0.38)	5
EGARCHJ	12694.48 (0.22)	7	12692.58 (0.32)	8
EGARCHJt	12702.12 (0.43)	3	12702.90 (0.64)	4
SVL	12701.09 (0.39)	4	12708.12 (0.94)	2
SVLt	12719.13 (0.92)	1	12720.91 (0.58)	1
SVLJ	12695.83 (1.43)	6	12699.89 (1.23)	6
SVLJt	12701.02 (0.91)	5	12703.28 (0.97)	3
SVLCJ	12693.20 (1.62)	8	12695.58 (1.83)	7

* The values are based on log scale and standard error in parentheses.

Table 10: Marginal likelihood (ML) for the S&P500 data (prior sensitivity analysis).

Prior3. Though the ranking changes slightly between the two priors, the estimation results indicate that the best performance of the SVLt model is quite robust.

7 Conclusion

This paper compares the empirical performance of the fit among the GARCH and the SV models with leverage, jumps and heavy-tailed errors. The estimation methodology for these models is developed using the Markov chain Monte Carlo estimation methods and the model comparison is examined based on the marginal likelihood. The empirical results show the SV model with leverage and Student-t distribution errors best fits the daily returns of stock return and exchange rate series among the competing models. The estimation results indicates that the leverage effect plays the important role for stock return, while not so much for exchange rate return probably because of two-side participants in the currency market. The results of model comparison are found robust for the subsample periods and different priors.

Appendix. MCMC algorithm for GARCHJt and EGARCHJt model

We illustrate the MCMC procedure for the GARCHJt and the EGARCHJt model in this appendix. The proposed algorithm is as follows:

1. Initialize $\vartheta, \kappa, \delta, \nu, \gamma$ and λ.

2. Sample $\vartheta|\delta, \gamma, \lambda, y$.

   To sample ϑ from its conditional posterior distribution $\pi(\vartheta|\delta, \gamma, \lambda, y) \propto \pi(\vartheta) f(y|\vartheta, \delta, \gamma, \lambda)$, we use the Metropolis-Hasting (M-H) algorithm (see e.g., Chib and Greenberg (1995)), because the posterior distribution is not available in the form of an usual distribution such as a normal distribution. We construct the proposal density in the form of a normal distribution for the M-H algorithm by fitting the mean and variance on the target posterior density from the product of second-order Taylor expansion. For the restriction for the parameters in the GARCH models, we consider the transformation of $\vartheta \to \tilde{\vartheta} = (\tilde{\omega}, \tilde{\alpha}, \tilde{\beta})$ such that $\tilde{\omega} = \log \omega$, $\tilde{\alpha} = \log(\alpha/(1-\alpha))$, and $\tilde{\beta} = \log(\beta/(1-\alpha-\beta))$. The posterior density is transformed over the full space of R^3, where we easily implement the proposal density. To draw a candidate of the M-H algorithm, we find $\tilde{\vartheta}^*$ which maximizes (or approximately maximizes) the posterior density, $\pi(\tilde{\vartheta}|\delta, \gamma, \lambda, y)$, and generate the candidate $\tilde{\vartheta}_*$ from the normal distribution $N(\mu_*, \Sigma_*)$, where

   $$\mu_* = \tilde{\vartheta}^* + \Sigma_* \left.\frac{\partial \log \pi(\tilde{\vartheta}|\delta, \gamma, \lambda, y)}{\partial \vartheta}\right|_{\tilde{\vartheta}=\tilde{\vartheta}^*}, \quad \Sigma_*^{-1} = -\left.\frac{\partial \log \pi(\tilde{\vartheta}|\delta, \gamma, \lambda, y)}{\partial \tilde{\vartheta} \partial \tilde{\vartheta}'}\right|_{\tilde{\vartheta}=\tilde{\vartheta}^*}.$$

   which is obtained from the second-order Taylor expansion around $\tilde{\theta}$. Let ϑ_0 denote the current point of ϑ and we reversely transform $\tilde{\vartheta}^*$ to ϑ^*. We accept the candidate ϑ^* with probability

   $$\alpha(\theta_0, \theta^*|\delta, \gamma, \lambda, y) = \min\left\{\frac{\pi(\theta^*|\delta, \gamma, \lambda, y) q(\theta_0|\delta, \gamma, \lambda, y)}{\pi(\theta_0|\delta, \gamma, \lambda, y) q(\theta^*|\delta, \gamma, \lambda, y)}, 1\right\},$$

   where q denotes the proposal density. If the candidate θ^* is rejected, we take the current value θ_0.

3. Sample $(\delta, \gamma)|\vartheta, \kappa, \lambda, y$.

   (a) Sample $\delta|\vartheta, \kappa, \lambda, y$.

The joint posterior distribution of (δ, γ) is given by

$$\pi(\delta, \gamma | \vartheta, \kappa, \lambda, y) \propto \pi(\delta) \prod_{t=1}^{n} \kappa^{\gamma_t} (1 - \kappa)^{1-\gamma_t} \frac{1}{\sqrt{\sigma_t^2 \lambda_t + \gamma_t \delta^2}} \exp \left\{ -\frac{y_t^2}{2(\sigma_t^2 \lambda_t + \gamma_t \delta^2)} \right\},$$

where k_t is marginalized in the likelihood function. To sample δ, we further marginalize this joint posterior distribution over γ. The marginalized conditional posterior distribution is formed as

$$\pi(\delta | \vartheta, \kappa, \lambda, y) \propto \pi(\delta) \prod_{t=1}^{n} \left[\frac{\kappa}{\sqrt{\sigma_t^2 \lambda_t + \delta^2}} \exp \left\{ -\frac{y_t^2}{2(\sigma_t^2 \lambda_t + \delta^2)} \right\} + \frac{1 - \kappa}{\sigma_t \sqrt{\lambda_t}} \exp \left(-\frac{y_t^2}{2\sigma_t^2 \lambda_t} \right) \right].$$

We can sample δ using the M-H algorithm. We note that this marginalization enables us to accelerate the convergence of the MCMC sampling.

(b) Sample $\gamma | \vartheta, \kappa, \delta, \lambda, y$.

Sampling γ from its posterior distribution requires only to evaluate the Bernoulli distribution $\pi(\gamma_t | \vartheta, \kappa, \delta, \lambda, y)$ where $\gamma_t = 0, 1$. We sample γ_t using the probability mass function of its posterior density as

$$\pi(\gamma_t = 1 | \vartheta, \kappa, \delta, \lambda, y) \propto \frac{\kappa}{\sqrt{\sigma_t^2 \lambda_t + \delta^2}} \exp \left\{ -\frac{y_t^2}{2(\sigma_t^2 \lambda_t + \delta^2)} \right\},$$

$$\pi(\gamma_t = 0 | \vartheta, \kappa, \delta, \lambda, y) \propto \frac{1 - \kappa}{\sigma \sqrt{\lambda_t}} \exp \left(-\frac{y_t^2}{2\sigma_t^2 \lambda_t} \right),$$

for $t = 1, \ldots, n$.

4. Sample $\kappa | \vartheta, \delta, \gamma, \lambda, y$.

When we specify the prior as $\kappa \sim \text{Beta}(n_{\alpha 0}, n_{\beta 0})$ for the EGARCHJt model, we sample κ from $\text{Beta}(n_{\alpha 0} + n_1, n_{\beta 0} + n_0)$, where n_1 and n_0 denote the count of $\gamma_t = 1$ and $\gamma_t = 0$ respectively. On the other hand, for the GARCHJt model, we consider the posterior distribution of κ conditional on $(\vartheta, \delta, \gamma, \lambda)$, because we assume $\sigma_1^2 = (\omega + \alpha \kappa \delta^2)/(1 - \alpha - \beta - \rho/2)$. The posterior distribution of κ is not written in the form of a beta distribution. We sample κ using the M-H algorithm with the candidate drawn from $\text{Beta}(n_{\alpha 0} + n_1, n_{\beta 0} + n_0)$.

5. Sample $(\lambda, \nu) | \vartheta, \delta, \gamma, y$.

(a) Sample $\lambda | \vartheta, \delta, \nu, \gamma, y$.

The joint posterior distribution of (λ, ν) is given by

$$\pi(\lambda, \nu | \vartheta, \delta, \gamma, y) \quad \propto \quad \pi(\nu) \prod_{t=1}^{n} \frac{\left(\frac{\nu}{2}\right)^{\frac{\nu}{2}}}{\Gamma\left(\frac{\nu}{2}\right)} \lambda_t^{-\left(\frac{\nu}{2}+1\right)} \exp\left(-\frac{\nu}{2\lambda_t}\right)$$
$$\times \frac{1}{\sqrt{\sigma_t^2 \lambda_t + \gamma_t \delta^2}} \exp\left\{-\frac{y_t^2}{2(\sigma_t^2 \lambda_t + \gamma_t \delta^2)}\right\}.$$

We sample λ_t from its conditional posterior density,

$$\pi(\lambda_t | \vartheta, \delta, \nu, \gamma, y) \propto \lambda_t^{-\left(\frac{\nu}{2}+1\right)} \exp\left(-\frac{\nu}{2\lambda_t}\right) \frac{1}{\sqrt{\sigma_t^2 \lambda_t + \gamma_t \delta^2}} \exp\left\{-\frac{y_t^2}{2(\sigma_t^2 \lambda_t + \gamma_t \delta^2)}\right\},$$

by the M-H algorithm with the candidate drawn as $(\lambda_t^*)^{-1} \sim \text{Gamma}(\nu/2, \nu/2)$, for $t = 1, \ldots, n$.

(b) Sample $\nu | \lambda$.

Finally, the conditional posterior distribution for ν is given by

$$\pi(\nu | \lambda) \propto \pi(\nu) \frac{\left(\frac{\nu}{2}\right)^{\frac{n\nu}{2}}}{\Gamma\left(\frac{\nu}{2}\right)^n} \prod_{t=1}^{n} \lambda_t^{-\frac{\nu}{2}} \exp\left(-\frac{\nu}{2} \sum_{t=1}^{n} \lambda_t^{-1}\right).$$

We sample ν by the M-H algorithm with the normal proposal density.

References

Andersen, T., L. Benzoni, and J. Lund (2002). An empirical investigation of continuous-time models for equity returns. *Journal of Finance 57*, 1239–1284.

Bauwens, L. and M. Lubrano (1998). Bayesian inference on GARCH models using Gibbs sampler. *Econometrics Journal 1*, c23–c46.

Berg, A., R. Meyer, and J. Yu (2004). DIC as a model comparison criterion for stochastic volatility models. *Journal of Business and Economic Statistics 22*, 107–120.

Black, F. (1976). Studies of stock market volatility changes. In *Proceedings of the American Statistical Association, Business and Economic Statistics Section*, pp. 177–181.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics 31*, 307–327.

Bollerslev, T., R. F. Engle, and D. B. Nelson (1994). ARCH models. In R. F. Endgle and D. McFadden (Eds.), *The Handbook of Econometrics 4*, pp. 2959–3038. Amsterdam: North-Holland.

Chan, W. H. and J. M. Maheu (2002). Conditional jump dynamics in stock market returns. *Journal of Business and Economic Statistics 20*, 377–389.

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

Chib, S. (1995). Marginal likelihood from the Gibbs output. *Journal of the American Statistical Association 90*, 1313–1321.

Chib, S. (2001). Markov chain Monte Carlo methods: computation and inference. In J. J. Heckman and E. Leamer (Eds.), *Handbook of Econometrics*, Volume 5, pp. 3569–3649. Amsterdam: North-Holland.

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

Chib, S. and E. Greenberg (1996). Markov chain Monte Carlo simulation methods in econometrics. *Econometric Theory 12*, 409–431.

Chib, S. and I. Jeliazkov (2001). Marginal likelihood from the Metropolis-Hastings output. *Journal of the American Statistical Association 96*, 270–291.

Chib, S. and I. Jeliazkov (2005). Accept-Reject Metropolis-Hastings sampling and marginal likelihood estimation. *Statistica Neerlandica 59*, 30–44.

Chib, S., F. Nardari, and N. Shephard (2002). Markov chain Monte Carlo methods for stochastic volatility models. *Journal of Econometrics 108*, 281–316.

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

Doornik, J. (2006). *Ox: Object Oriented Matrix Programming*. London: Timberlake Consultants Press.

Durbin, J. and S. J. Koopman (2002). Simple and efficient simulation smoother for state space time series analysis. *Biometrika 89*, 603–616.

Eraker, B., M. Johanners, and N. G. Polson (2003). The impact of jumps in returns and volatility. *Journal of Finance 53*, 1269–1330.

Ghysels, E., A. C. Harvey, and E. Renault (2002). Stochastic volatility. In C. R. Rao and G. S. Maddala (Eds.), *Statistical Methods in Finance*, pp. 119–191. Amsterdam: North-Holland.

Giot, P. and S. Laurent (2004). Modelling daily Value-at-Risk using realized volatility and ARCH type models. *Journal of Empirical Finance 11*, 379–398.

Glosten, L. R., R. Jagannathan, and D. Runkle (1993). On the relation between the expected value and the volatility of nominal ecxess returns on stocks. *Journal of Finance 48*, 1779–1801.

Jacquier, E., N. Polson, and P. Rossi (2004). Bayesian analysis of stochastic volatility with fat-tails and correlated errors. *Journal of Econometrics 122*(1), 185–212.

Jorion, P. (1988). On jump processes in the foreign exchange and stock markets. *Review of Financial Studies 1*, 427–445.

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

Kobayashi, M. (2006). Testing for volatility jumps in the stochastic volatility process. *Asia-Pacific Financial Markets 12*, 143–157.

Lehar, A., M. Scheicher, and C. Schittenkopf (2002). GARCH vs. stochatic volatility: Option pricing and risk management. *Journal of Banking and Finance 26*, 323–345.

Li, H., M. T. Wells, and C. L. Yu (2008). A Bayesian analysis of return dynamics with Lévy jumps. *Review of Financial Studies 21*, 2345–2378.

Maheu, J. M. and T. H. McCurdy (2004). News arrival, jump dynamics, and volatility components for individual stock returns. *Journal of Finance 59*, 755–793.

Nakajima, J. and Y. Omori (2009). Leverage, heavy-tails and correlated jumps in stochastic volatility models. *Computational Statistics and Data Analysis 53*, 2535–2553.

Nakatsuma, T. (2000). Bayesian analysis of ARMA-GARCH models: A Markov chain sampling approach. *Journal of Econometrics 95*, 57–69.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: a new approach. *Econometrica 59*, 347–370.

Omori, Y., S. Chib, N. Shephard, and J. Nakajima (2007). Stochastic volatility with leverage: fast likelihood inference. *Journal of Econometrics 140*, 425–449.

Pitt, M. and N. Shephard (1999). Filtering via simulation: auxiliary particle filter. *Journal of the American Statistical Association 94*, 590–599.

Raggi, D. and S. Bordignon (2006). Comparing stochastic volatility models through Monte Carlo simulations. *Computational Statistics and Data Analysis 50*, 1678–1699.

Shephard, N. (Ed.) (2005). *Stochastic Volatility: Selected Readings.* Oxford: Oxford University Press.

Vrontos, I. D., P. Dellaportas, and D. N. Politis (2000). Full Bayesian inference for GARCH and EGARCH models. *Journal of Business and Economic Statistics 18*, 187–198.

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