ISSN 1440-771X

MONASH UNIVERSITY
AUSTRALIA

DEPARTMENT OF ECONOMETRICS AND BUSINESS STATISTICS

Estimation of Asymmetric Box-Cox Stochastic Volatility Models Using MCMC Simulation

Xibin Zhang and Maxwell L. King

Working Paper 10/2003

Estimation of Asymmetric Box-Cox Stochastic Volatility Models Using MCMC Simulation

Xibin Zhang and **Maxwell L. King** [1]

Department of Econometrics and Business Statistics
Monash University, Clayton
Victoria 3800, Australia

April 2003

Abstract: The stochastic volatility model enjoys great success in modeling the time-varying volatility of asset returns. There are several specifications for volatility including the most popular one which allows logarithmic volatility to follow an autoregressive Gaussian process, known as log-normal stochastic volatility. However, from an econometric viewpoint, we lack a procedure to choose an appropriate functional form for volatility. Instead of the log-normal specification, Yu, Yang and Zhang (2002) assumed Box-Cox transformed volatility follows an autoregressive Gaussian process. However, the empirical evidence they found from currency markets is not strong enough to support the Box-Cox transformation against the alternatives, and it is necessary to seek further empirical evidence from the equity market. This paper develops a sampling algorithm for the Box-Cox stochastic volatility model with a leverage effect incorporated. When the model and the sampling algorithm are applied to the equity market, we find strong empirical evidence to support the Box-Cox transformation of volatility. In addition, the empirical study shows that it is important to incorporate the leverage effect into stochastic volatility models when the volatility of returns on a stock index is under investigation.

Key words: Box-Cox transformation, leverage effect, sampling algorithm.

JEL Classification: C6, C22 and C52

[1]The email addresses of the authors are, respectively, `Xibin.Zhang@BusEco.Monash.edu.au` and `Max.King@BusEco.Monash.edu.au`.

1 Introduction

The volatility of asset returns often exhibits a time-varying feature. One way of modeling volatility is to let it be a function of previous squared returns and lagged volatilities. This leads to the autoregressive conditional heteroskedasticity (ARCH) model developed by Engle (1982) and the generalized ARCH (GARCH) model by Bollerslev (1986). An alternative is the stochastic volatility (SV) model in which the volatility follows a latent stochastic process. The SV model has received more and more attention in the finance literature, because it provides an alternative approach to the Black-Scholes option pricing formula (Hull and White (1987)), and pricing an option based on the SV model is more accurate than that based on the Black-Scholes model (see, for example, Melino and Turnbull (1990)). Taylor (1982, 1986) showed that the SV model is often formulated in terms of stochastic differential equations,

$$\begin{cases} d(\ln p_t) = \alpha \; dt + \sigma_t \; dw_{1t} \\ d(\ln \sigma_t^2) = \lambda(\xi - \ln \sigma_t^2) \; dt + \sigma_w \; dw_{2t} \end{cases}, \tag{1}$$

where p_t is the price of an asset at time t and $(w_{1t}, w_{2t})'$ is a bivariate standard Brownian motion. The correlation between dw_{1t} and dw_{2t}, denoted by $\rho = corr(dw_{1t}, dw_{2t})$, captures the leverage effect [2]. The parameter ξ represents the long-run mean of the log-volatility, λ represents the adjustment rate and σ_w captures the variation in the log-volatility. Model (1) is the continuous-time log-normal stochastic volatility model. The empirical version of the SV model is typically formulated in discrete time as

$$\begin{cases} y_t = \sigma_t \; \varepsilon_t \\ \ln \sigma_{t+1}^2 = \mu + \phi \, (\ln \sigma_t^2 - \mu) + \sigma_\eta \eta_{t+1} \end{cases}, \tag{2}$$

[2] The leverage effect refers to the phenomenon that price movements are negatively correlated with volatility. This kind of asymmetric behaviour is often observed in stock price movements. Empirical evidence on leverage effects can be found in Nelson (1991), Gallant, Rossi and Tauchen (1992, 1993), Campbell and Kyle (1993) and Engle and Ng (1993) among others.

where y_t is the observed continuously compounded return, $\varepsilon_t \sim N(0,1)$, $\eta_t \sim N(0,1)$, $\ln \sigma_1^2 \sim N(0, \sigma_\eta^2/(1-\phi^2))$ and the correlation between ε_t and η_{t+1}, denoted by $\rho = corr(\varepsilon_t, \eta_{t+1})$, captures the leverage effect [3]. To reflect the asymmetric feature of the error terms in the mean and volatility equations, this model is often termed the asymmetric log-normal SV model, which was set up based on models of Clark (1973) and Tauchen and Pitts (1983) and was first documented by Taylor (1982).

A notable feature of (1) is that the logarithmic volatility is assumed to follow an Ornstein-Uhlenbeck (OU) process (while in the discrete-time context, the logarithmic volatility forms an autoregressive process with Gaussian errors as specified in (2)). Actually, there are some other specifications for the volatility process. In terms of continuous-time SV models, Johnson and Shanno (1987) assumed that the square root of the volatility follows a geometric Brownian motion, Stein and Stein (1991) specified the square root of the volatility as an OU process, Hull and White (1987) formulated the volatility as a geometric Brownian motion process, and Heston (1993) assumed that the volatility follows a square-root process which is similar to that of Cox, Ingersoll and Ross (1985). In terms of discrete-time SV models, Andersen (1994) introduced a class of polynomial SV models which encompasses most of the discrete-time SV models in the literature, and Barndorff-Nielsen and Shephard (2001) presented non-Gaussian OU based SV models.

From an econometric viewpoint, given all these specifications for the volatility process, we lack a procedure to select an appropriate functional form for the volatility process.

[3]Taylor (1994) suggested that the correlation between ε_t and η_{t+1} captures the leverage effect, and empirical evidence can be found in Ghysels, Harvey and Renault (1996), Harvey and Shephard (1996) and Meyer and Yu (2000) among others. Jacquier, Polson and Rossi (2002) and Eraker, Johannes and Polson (2002) allowed the correlation between ε_t and η_t to capture the leverage effect. Though the difference is marginal in empirical studies when the time interval between two successive observations is very small, strictly speaking, the correlation between ε_t and η_{t+1} is more accurate in capturing the asymmetric feedback between error terms in mean and volatility equations. See, for example, Yu (2002) for a discussion on the leverage effect.

The correct specification for stochastic volatility is very important, because different functional forms lead to different formulae for option pricing, and any misspecification of the functional form may result in incorrect option prices.

Yu, Yang and Zhang (2002) presented a generalization to the specification of log-normal volatility which allows the Box-Cox transformed volatility to follow an autoregressive Gaussian process,

$$\begin{cases} y_t = \sigma_t \, \varepsilon_t \\ h\left(\sigma_t^2, \delta\right) = \mu + \phi\left[h\left(\sigma_{t-1}^2, \delta\right) - \mu\right] + \sigma_u u_t \end{cases}, \tag{3}$$

where $\varepsilon_t \sim N(0,1)$, $u_t \sim N(0,1)$, the correlation between ε_t and u_{t+1} is ρ and $h\left(\sigma_t^2, \delta\right)$ is the Box-Cox transformation of σ_t^2 (Box and Cox, 1964), with $h\left(\cdot, \delta\right)$ being defined by,

$$h\left(x, \delta\right) = \begin{cases} (x^\delta - 1)/\delta & \text{if } \delta \neq 0 \\ \ln x & \text{if } \delta = 0 \end{cases}. \tag{4}$$

This model is called the Box-Cox transformed stochastic volatility (BCSV) model. In the case of $\rho = 0$, Yu, Yang and Zhang (2002) developed a Markov chain Monte Carlo (MCMC) algorithm to sample parameters and latent volatilities. When applying the BCSV model to daily returns of the dollar/pound exchange rate, they found that the 90% Bayesian confidence interval does not cover 0 or 0.5, which represents the logarithmic and square-root transformations of volatility, respectively. They concluded that the Box-Cox transformation is more appropriate than the alternatives of the logarithmic and the square-root transformation. When the BCSV model was applied to daily returns of the other exchange rates, which are, respectively, the Canadian dollar, French franc, Deutsche mark and Japanese yen, the estimated δ was not statistically different from zero, because the Bayesian confidence intervals always cover zero. Hence the empirical evidence obtained from these currency markets is not strong enough to support the Box-Cox transformation to the volatility. While the BCSV model is meaningful in theory, it is necessary to seek strong empirical evidence from the equity market to support this kind of specification.

In the equity market, returns on equity prices often exhibit a strong leverage effect (see, for example, Eraker, Johannes and Polson (2002) for an empirical evidence). Jacquier, Polson and Rossi (2002) pointed out that the leverage effect often induces skewness in the marginal distribution of returns on asset prices. Their finding is consistent with the non-parametric evidence found by Gallant, Hsieh and Tauchen (1997). Yu, Yang and Zhang (2002) did not incorporate the leverage effect into the BCSV model in their empirical study, because leverage effects seem to be relatively unimportant in currency markets. If the BCSV model is employed to model returns on equity prices, leverage effects cannot be ignored and should be incorporated into the BCSV model. Hence it is very important to develop a relevant sampling algorithm to estimate the BCSV model with leverage effects being incorporated, or equivalently the asymmetric BCSV model.

This paper develops a MCMC algorithm for the BCSV model based on the fully specified posterior density of parameters and latent volatilities. The paper is organized as follows. Section 2 presents the description of the asymmetric BCSV model, the fully specified posterior density, conditional densities, and sampling algorithm designed to sample parameters and volatilities. In Section 3, we apply the asymmetric BCSV model and the sampling algorithm to a generated dataset so that the performance of the sampling algorithm can be examined. Section 4 presents an application of the asymmetric BCSV model and the sampling algorithm to daily returns on six major stock indexes. We find strong empirical evidence to support the Box-Cox transformation against the alternatives of the logarithmic and square-root transformations of volatility. Section 5 concludes the paper.

2 MCMC in the BCSV Model

2.1 BCSV Model with Leverage Effects

The discrete-time BCSV model can be expressed as,

$$\begin{cases} y_t = \sigma_t \ \varepsilon_t \\ h\left(\sigma_{t+1}^2, \delta\right) = \mu + \phi\left[h\left(\sigma_t^2, \delta\right) - \mu\right] + \sigma_u u_{t+1} \end{cases}, \tag{5}$$

where $h(\cdot, \delta)$ is defined in (4), $(\varepsilon_t, u_{t+1})'$ follows a bivariate normal distribution with mean zero and covariance matrix

$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \tag{6}$$

for $t = 1, 2, \cdots, n-1$, $h(\sigma_1^2, \delta) \sim N\left(\mu, (1-\rho^2)\sigma_u^2/(1-\phi^2)\right)$ and $y_n \sim N(0, \sigma_n^2)$. This model can be equivalently represented by

$$\begin{cases} y_t = \sqrt{g(\alpha_t, \delta)} \ \varepsilon_t \\ \alpha_{t+1} = \mu + \phi(\alpha_t - \mu) + \sigma_u u_{t+1} \end{cases}, \tag{7}$$

where $\alpha_t = h\left(\sigma_t^2, \delta\right)$ and

$$g(\alpha_t, \delta) = \begin{cases} (1 + \delta\alpha_t)^{1/\delta} & \text{if } \delta \neq 0 \\ \exp(\alpha_t) & \text{if } \delta = 0 \end{cases},$$

which is denoted hereafter by g_t. Define [4]

$$u_{t+1} = \rho\varepsilon_t + \sqrt{1-\rho^2}\eta_{t+1}, \tag{8}$$

for $t = 1, 2, \cdots, n-1$, where η_{t+1} is assumed to follow $N(0,1)$ and to be uncorrelated with ε_t. Equation (8) shows that $var(u_{t+1}) = 1$ and $cov(u_{t+1}, \varepsilon_t) = \rho$ which satisfies the model specification in (5) and (7). Substituting (8) into (7), we obtain

$$\alpha_{t+1} = \mu + \phi(\alpha_t - \mu) + \rho\sigma_u g_t^{-1/2} y_t + \sqrt{1-\rho^2}\sigma_u \eta_{t+1}.$$

[4] The transformation made here is a common practice in the finance literature to deal with the leverage effect in a SV model. See Section 2.3.1 in Fouque, Papanicolaou and Sircar (2000) for more details.

When incorporating the leverage effect into log-normal SV models, Jacquier, Polson and Rossi (2002) re-parameterized ρ and σ as $\varphi = \rho\sigma_u$ and $\tau^2 = (1-\rho^2)\sigma_u^2$, respectively. We follow this re-parameterization and obtain

$$\begin{cases} y_t = \sqrt{g_t}\ \varepsilon_t \\ \\ \alpha_{t+1} = \mu + \phi(\alpha_t - \mu) + \varphi g_t^{-1/2} y_t + \tau\eta_{t+1} \end{cases}, \tag{9}$$

where $\alpha_1 \sim N(\mu, \tau^2/(1-\phi^2))$ and $y_n \sim N(0, g_n)$. As ε_t and η_{t+1} are uncorrelated, we can easily obtain the joint likelihood of y_t given parameters and latent volatilities. Hereafter we refer to (9) as the asymmetric BCSV model.

2.2 MCMC

Bayesian inference concerning a parameter vector θ conditional on data $\mathbf{y}$ is made through the posterior density $\pi(\theta|\mathbf{y})$ which takes the form

$$\pi(\theta|\mathbf{y}) = c\, L(\mathbf{y}|\theta)\, \pi(\theta),$$

where c is a normalizing constant, $L(\mathbf{y}|\theta)$ is the likelihood of $\mathbf{y}$ conditional upon θ, and $\pi(\theta)$ is the prior density of θ. The Bayesian approach requires that statistical inference be based on the posterior. However, dealing with the posterior is often analytically intractable, because the normalizing constant is typically unknown. The MCMC method aims to provide a general mechanism to sample the parameter vector from its posterior density. While simulating directly from the posterior distribution is typically very difficult, the MCMC method sets up a Markov chain so that its stationary distribution is the same as the posterior density. When the Markov chain converges, the simulated values may be regarded as a posterior sample of the parameter vector.

The MCMC approach to inference in SV models requires a number of components: the likelihood, the latent volatility dynamics and prior parameter distributions. Let $\mathbf{y}$

denote the vector of observed returns, α denote the vector of the Box-Cox transformed volatilities and θ the parameter vector. By the Bayes theorem, the posterior of (θ, α) is

$$\pi(\theta, \alpha|\mathbf{y}) \propto p(\mathbf{y}|\theta, \alpha)p(\alpha|\theta)p(\theta), \tag{10}$$

where $p(\mathbf{y}|\theta, \alpha)$ is the likelihood, $p(\alpha|\theta)$ is the distribution of the transformed volatility, arising from the parametric model specification, and $p(\theta)$ is the prior distribution of the parameter vector. As discussed in Kim, Shephard and Chib (1998), the key issue in estimating a log-normal SV model is that the likelihood, which is expressed as

$$f(\mathbf{y}|\theta) = \int p(\mathbf{y}|\theta, \alpha)p(\alpha|\theta)d\alpha, \tag{11}$$

is intractable. This problem can be overcome by focusing instead on $\pi(\theta, \alpha|\mathbf{y})$, and MCMC algorithms can be developed to sample θ and α from $\pi(\theta, \alpha|\mathbf{y})$ without directly computing the likelihood function $f(\theta|\mathbf{y})$. One characteristic of the SV model is that latent volatilities are highly correlated, and they can be sampled as a vector. However, the highly correlated nature of latent volatilities adds many difficulties in developing a MCMC algorithm to sample parameters and latent volatilities. MCMC algorithms for log-normal SV models can be found in Shephard (1993), Jacquier, Polson and Rossi (1994), Shephard and Pitt (1997), Kim, Shephard and Chib (1998), Chib, Nardari and Shephard (2002), Eraker, Johannes and Polson (2002), and Jacquier, Polson and Rossi (2002) among many others. Chib (2001) provided a recent survey.

2.3 Joint Posterior of Parameters and Latent Volatilities

Assume that $(\phi + 1)/2 \sim Beta(\omega, \gamma)$ and $\tau^2 \sim IG(\nu/2, S_\sigma/2)$, which are, respectively, expressed explicitly as

$$p(\phi) \propto \left(\frac{\phi+1}{2}\right)^{\omega-1} \left(1 - \frac{\phi+1}{2}\right)^{\gamma-1},$$

$$p(\tau^2) \sim \left(\frac{1}{\tau^2}\right)^{\nu/2+1} \exp\left\{-\frac{S_\tau/2}{\tau^2}\right\},$$

where ω, γ, ν and S_τ are hyperparameters to be defined by the investigator. The priors of the other parameters are, respectively, $\varphi|\tau^2 \sim N(\varphi_0, \tau^2/p_0)$, $\mu|\tau^2 \sim N(\mu_0, \tau^2/q_0)$ and $\delta \sim N(\mu_\delta, \sigma_\delta^2)$ with φ_0, μ_0, p_0 and q_0 being hyperparameters. The joint prior of θ is

$$\begin{aligned} p(\phi, \delta, \mu, \rho, \sigma) &= p(\delta) \times (1+\phi)^{(\omega-1/2)}(1-\phi)^{(\gamma-1/2)} \times \left(\frac{1}{\tau^2}\right)^{\nu/2+1} \exp\left\{-\frac{S_\tau/2}{\tau^2}\right\} \times \\ &\quad \left(\frac{1}{\tau^2/p_0}\right)^{1/2} \exp\left\{-\frac{(\varphi-\varphi_0)^2}{2\tau^2/p_0}\right\} \times \left(\frac{1}{\tau^2/q_0}\right)^{1/2} \exp\left\{-\frac{(\mu-\mu_0)^2}{2\tau^2/q_0}\right\}. \end{aligned}$$

According to (9) and (10), the posterior of (θ, α) is

$$\begin{aligned} \pi(\theta, \alpha|\mathbf{y}) &\propto \prod_{t=1}^{n-1} p(y_t|\alpha_t, \theta) \times p(y_n|\theta) \times p(\alpha_1|\theta) \times \prod_{t=1}^{n-1} p(\alpha_{t+1}|\alpha_t, \theta) \times p(\phi, \delta, \mu, \rho, \sigma) \\ &= \left(\prod_{t=1}^{n} g_t^{-1/2}\right) \exp\left\{-\frac{1}{2}\sum_{t=1}^{n}\frac{y_t^2}{g_t}\right\} \times \left(\frac{1}{\tau^2}\right)^{(n+\nu+2)/2+1} \exp\left\{-\frac{\kappa}{2\tau^2}\right\} \\ &\quad \times p(\delta)(1+\phi)^{\omega-1/2}(1-\phi)^{\gamma-1/2} \end{aligned} \tag{12}$$

where

$$\begin{aligned} \kappa &= (1-\phi^2)(\alpha_1-\mu)^2 + \sum_{t=1}^{n-1}\left(\alpha_{t+1} - \mu - \phi(\alpha_t - \mu) - \varphi g_t^{-1/2} y_t\right)^2 \\ &\quad + p_0(\varphi-\varphi_0)^2 + q_0(\mu-\mu_0)^2 + S_\tau. \end{aligned}$$

After integrating out τ^2 from the joint posterior (12), we obtain the log-posterior

$$\begin{aligned} \log\ p(\phi, \delta, \mu, \varphi, \alpha|\mathbf{y}) &= \log\ p(\delta) + (\omega - 1/2)\log(1+\phi) + (\gamma-1/2)\log(1-\phi) \\ &\quad -\frac{1}{2}\sum_{t=1}^{n}\log(g_t) - \frac{1}{2}\sum_{t=1}^{n}\frac{y_t^2}{g_t} - \frac{n+\nu+2}{2}\log(\kappa/2). \end{aligned} \tag{13}$$

Hence we obtain the posterior density of $(\phi, \delta, \mu, \varphi, \alpha')'$, while the conditional posterior of τ^2 is the inverted gamma density. In the appendix, we present a different method to obtain the joint posterior of $(\theta', \alpha')'$ given $\mathbf{y}$. These two approaches result in exactly the same posterior density.

2.4 Conditional Posteriors

Once we obtain the joint posterior of $(\theta', \alpha')'$, we can use the Gibbs sampler to sample each component of $(\theta', \alpha')'$ conditional on the other components. However, the mixing speed will generally be slow. If conditional posteriors of some parameters can be obtained, these parameters can be sampled, respectively, from their conditional posteriors independently. As a consequence, the overall mixing performance will be greatly improved [5].

2.4.1 Conditional Posterior of τ^2

As τ^2 can be integrated out of the joint posterior (12), the conditional posterior of τ^2 is,

$$\tau^2 \sim IG\left(\frac{n+\nu+2}{2}, \frac{\kappa}{2}\right). \tag{14}$$

Hence τ^2 can be sampled directly from its conditional posterior, given the other parameters and latent volatilities.

2.4.2 Sampling φ

When φ is to be sampled based on the joint posterior (12), the relevant part for φ in the joint posterior is

$$\begin{aligned}
& p(\varphi|\tau^2, \phi, \delta, \mu, \alpha, \mathbf{y}) \\
\propto\ & \exp\left\{-\frac{1}{2\tau^2}\left[\sum_{t=1}^{n-1}\left((\alpha_{t+1}-\mu)-\phi(\alpha_t-\mu)-\varphi g_t^{-1/2} y_t\right)^2 + p_0(\varphi-\varphi_0)^2\right]\right\} \\
=\ & \exp\left\{-\frac{1}{2\tau^2}\left[a_{11}\varphi^2 - 2a_{12}\varphi + a_{22} + p_0(\varphi^2 - 2\varphi_0\varphi + \varphi_0^2)\right]\right\} \\
\propto\ & \exp\left\{-\frac{1}{2\tau^2/(a_{11}+p_0)}\left(\varphi^2 - 2\frac{a_{12}+\varphi_0 p_0}{a_{11}+p_0}\varphi\right)\right\},
\end{aligned}$$

where

$$a_{11} = \sum_{t=1}^{n-1} y_t^2/g_t, \quad a_{12} = \sum_{t=1}^{n-1}[\alpha_{t+1}-\mu-\phi(\alpha_t-\mu)]y_t/\sqrt{g_t},$$

[5]See, for example, Johannes and Polson (2003) for a discussion on sampling techniques based on conditional posteriors.

$$a_{22} = \sum_{t=1}^{n-1} [\alpha_{t+1} - \mu - \phi(\alpha_t - \mu)]^2.$$

Hence the conditional posterior of φ is the Gaussian distribution,

$$\varphi \sim N\left(\frac{a_{12} + \varphi_0 p_0}{a_{11} + p_0}, \frac{\tau^2}{a_{11} + p_0}\right),$$

based on which φ can be sampled directly, given the other parameters and latent volatilities. Once τ^2 and φ are sampled, respectively, from their conditional posteriors, we can calculate ρ and σ_u through $\sigma_u^2 = \varphi^2 + \tau^2$ and $\rho = \varphi/\sigma_u$.

2.4.3 Conditional Posterior of μ

The Gibbs sampler allows us to update μ according to the joint posterior (12), given the other parameters and vector of transformed latent volatilities. As far as μ is concerned, the relevant part in the joint posterior is

$$\begin{aligned} & p(\mu|\tau^2, \phi, \delta, \varphi, \alpha, \mathbf{y}) \\ \propto \; & \exp\left\{-\frac{1}{2\tau^2}\left[(1-\phi^2)(\alpha_1 - \mu)^2 + \sum_{t=1}^{n-1}(b_{t+1} - (1-\phi)\mu)^2 + q_0(\mu - \mu_0)^2\right]\right\} \\ \propto \; & \exp\left\{-\frac{(n-1)(1-\phi)^2 + (1-\phi^2) + q_0}{2\tau^2}\left(\mu^2 - 2\frac{(1-\phi^2)\alpha_1 + (1-\phi)\sum b_{t+1} + q_0\mu_0}{(n-1)(1-\phi)^2 + (1-\phi^2) + q_0}\,\mu\right)\right\}, \end{aligned}$$

where $b_{t+1} = \alpha_{t+1} - \phi\alpha_t - \varphi g_t^{-1/2} y_t$ for $t = 1, 2, \cdots, n-1$. Then the conditional posterior of μ is the Gaussian distribution with mean and variance being defined, respectively, by

$$\begin{aligned} \mu_* &= \frac{(1-\phi^2)\alpha_1 + (1-\phi)\sum_{t=1}^{n-1} b_{t+1} + q_0\mu_0}{(n-1)(1-\phi)^2 + (1-\phi^2) + q_0}, \\ \sigma_*^2 &= \frac{\tau^2}{(n-1)(1-\phi)^2 + (1-\phi^2) + q_0}. \end{aligned}$$

Hence μ can be sampled directly from $N(\mu_*, \sigma_*^2)$, given the other parameters and latent volatilities.

2.4.4 Sampling ϕ and δ

When sampling ϕ and δ, we use the random-walk Metropolis algorithm. We can update (ϕ, δ) simultaneously on an elliptical contour and accept or reject the updated values according to the Metropolis-Hastings rule, while the acceptance probability is calculated based on the joint posterior (12) or (13). As the other parameters and latent volatilities are given, it does not matter which form of the joint posterior is used.

2.4.5 Sampling α

The Gibbs sampler allows us to update each component of α at a time and accept or reject the updated value according to the Metropolis-Hastings rule, where the acceptance probability is calculated based on the joint posterior (12) or (13). A disadvantage of such an approach is that we need to compute the full joint posterior when updating each component of α. The extensive computation usually results in a relatively low mixing rate. The relevant part in the joint posterior (12) for computing the acceptance probability is

$$\begin{aligned}\log \pi(\alpha|\theta, \mathbf{y}) \propto\ & \frac{1}{2}\sum_{t=1}^{n} \log\ g_t - \frac{1}{2}\sum_{t=1}^{n} \frac{y_t^2}{g_t} - \frac{1}{2\tau^2} \\ & \left\{(1-\phi^2)(\alpha_1-\mu)^2 + \sum_{t=1}^{n-1}\left(\alpha_{t+1} - \mu - \phi(\alpha_t - \mu) - \varphi g_t^{-1/2} y_t\right)^2\right\}. \quad (15)\end{aligned}$$

Moreover, when updating a component of α, say α_k, we only need to calculate the related terms, denoted by $\pi(\alpha_k)$, in (15) with the other terms unchanged. We have the following expressions,

$$\begin{aligned}\pi(\alpha_1) &= -\frac{\log g_1}{2} - \frac{y_1^2}{2g_1} - \frac{1}{2\tau^2}\left[(1-\phi^2)(\alpha_1-\mu)^2 + \left(\alpha_2 - \mu - \phi(\alpha_1-\mu) - \varphi g_1^{-1/2} y_1\right)^2\right], \\ \pi(\alpha_k) &= -\frac{\log g_k}{2} - \frac{y_k^2}{2g_k} - \frac{1}{2\tau^2}\sum_{t=k-1}^{k}\left[\alpha_{t+1} - \mu - \phi(\alpha_t - \mu) - \varphi g_t^{-1/2} y_t\right]^2, \text{for } k = 2, \cdots, n-1, \\ \pi(\alpha_n) &= -\frac{\log g_n}{2} - \frac{y_n^2}{2g_n} - \frac{1}{2\tau^2}\left[\alpha_n - \mu - \phi(\alpha_{n-1} - \mu) - \varphi g_{n-1}^{-1/2} y_{n-1}\right]^2.\end{aligned}$$

When updating α_k ($k = 1, 2, \cdots, n$), we use $\pi(\alpha_k)$ to calculate the acceptance probability, with which the updated value is accepted.

It is important to note that the proposed sampling algorithm is very flexible. If we fix $\rho = 0$ in the above algorithm during the MCMC simulation, the algorithm is exactly the same as that of Yu, Yang and Zhang (2002). If we fix $\delta = 0$, the algorithm is similar in spirit as that of Jacquier, Polson and Rossi (2002). [6]

3 Application to Artificially Generated Data

In order to examine the accuracy and reliability of the proposed MCMC algorithm for sampling parameters and latent volatilities in the asymmetric BCSV model, we apply the algorithm to a dataset which is generated through the asymmetric BCSV model by using the following parameters: $\phi = 0.98$, $\delta = 0.3$, $\mu = 1.0$, $\rho = -0.25$ and $\sigma = 0.2$. These values are chosen to represent typical daily returns of financial assets (see, for example, Shephard and Pitt (1997) and Yu, Yang and Zhang (2002)). The generated dataset contains the returns and transformed volatility, which provides an opportunity to compare the estimated parameters with true parameters, as well as the estimated volatility and true volatility.

3.1 Data Generation

Instead of making a transformation to u_{t+1} as expressed in (8), we make a transformation to ε_t as

$$\varepsilon_t = \rho\, u_{t+1} + \sqrt{1-\rho^2} w_t,$$

[6] In terms of the log-normal SV model, Kim, Shephard and Chib (1998) presented a single-move accept/reject algorithm to sample latent volatilities. Jacquier, Polson and Rossi (2002) obtained an approximate "blanket" for the posterior density of α_i ($i = 1, 2, \cdots, n$) and used the accept/reject algorithm to sample latent volatilities. Both methods are similar in spirit and are efficient for sampling the latent volatility. However, both methods should be properly modified to meet the specific features of the BCSV model. The sampling method presented here is easy to implement and is eligible for all SV models.

where $w_t \sim N(0,1)$, $E(u_{t+1}w_t) = 0$ and u_{t+1} is the error term in the volatility equation defined in (7), which can be expressed as

$$u_{t+1} = \frac{\alpha_{t+1} - \mu - \phi(\alpha_t - \mu)}{\sigma}.$$

Then (7) can be equivalently represented as

$$\begin{cases} y_t = \sqrt{g_t}\rho\left[\alpha_{t+1} - \mu - \phi(\alpha_t - \mu)\right]/\sigma + \sqrt{g_t(1-\rho^2)}w_t, \\ \alpha_{t+1} = \mu + \phi(\alpha_t - \mu) + \sigma u_{t+1}, \end{cases} \tag{16}$$

for $t = 1, 2, \cdots, n-1$, where $\alpha_1 \sim N(\mu, \sigma^2/(1-\phi^2))$ and $y_n \sim N(0, g_n(1-\rho^2))$.

Given θ, we can generate α and $\mathbf{y}$ through the following equations,

$$\alpha_{t+1}|(\alpha_t, \theta) \sim N\left(\mu + \phi(\alpha_t - \mu), \sigma^2\right), \tag{17}$$

$$y_t|(\alpha_{t+1}, \alpha_t, \theta) \sim N\left(\frac{\rho}{\sigma}\sqrt{g_t}\left[\alpha_{t+1} - \mu - \phi(\alpha_t - \mu)\right], g_t(1-\rho^2)\right), \tag{18}$$

where $\alpha_1 \sim N(\mu, \sigma^2/(1-\phi^2))$ and $y_n \sim N(0, g_n(1-\rho^2))$.

3.2 Assessment of Sampling Accuracy

When the MCMC iteration procedure has converged, the recorded draws, denoted by $\{(\theta^{(i)}, \alpha^{(i)})' : i = 1, 2, \cdots, N\}$, form a Markov chain whose stationary transition density is the posterior $\pi(\theta, \alpha|\mathbf{y})$ defined in (12). The MCMC output is often summarized in terms of the ergodic average or the posterior mean in the form of

$$\overline{f}_N = \frac{1}{N}\sum_{i=1}^{N} f(\theta^{(i)}), \tag{19}$$

where $f(\cdot)$ is a real-valued function to be estimated [7]. Roberts (1996) pointed out that most Markov chains produced in MCMC simulations converge geometrically to the stationary distribution $\pi(\theta, \alpha|\mathbf{y})$, and one of the consequences of the geometric convergence

[7] Under the circumstance that the parameter vector itself is of interest, the function $f(\cdot)$ is the identity function, that is, $f(x) = x$.

is that

$$\sqrt{N}\left(\overline{f}_N - E_\pi[f(\theta)]\right) \xrightarrow{D} N(0, \sigma_f^2), \tag{20}$$

where $E_\pi[\cdot]$ denotes the expectation operator under $\pi(\theta, \alpha|\mathbf{y})$, and the convergence is in distribution. In order to assess the accuracy of the ergodic average as an estimate of $E_\pi[f(\theta)]$, it is necessary to estimate σ_f^2, and one of the most commonly used methods is to estimate σ_f^2 using the batch-mean method stated below.

Let the number of recorded draws be $N = m \times n$, where n is sufficiently large so that

$$y_k = \frac{1}{n}\sum_{i=(k-1)n+1}^{kn} f(\theta^{(i)}), \tag{21}$$

for $k = 1, 2, \cdots, m$, are approximately independently distributed as $N(E_\pi[f(\theta)], \sigma_f^2/n)$. Therefore σ_f^2 can be estimated by

$$\hat{\sigma}_f^2 = \frac{n}{m-1}\sum_{k=1}^{m}(y_k - \overline{f}_N)^2, \tag{22}$$

where $\overline{f}_N$ is defined in equation (19). Thus, the standard error of $\overline{f}_N$ can be estimated by $\sqrt{\hat{\sigma}_f^2/N}$, which is called the batch-mean standard error (BMSE) and is commonly used for checking the mixing performance.

In addition to the BMSE, one may also compute the standard deviation $\tilde{\sigma}_f$ directly based on the sampled path using the formula

$$\tilde{\sigma}_f = \left\{\frac{1}{N-1}\sum_{i=1}^{N}\left[f(\theta^{(i)}) - \overline{f}_N\right]^2\right\}^{1/2}. \tag{23}$$

It is important to note that the computation of standard deviation via (23) should be based on an independent posterior sample. As the draws from a MCMC procedure form a Markov chain, one simple procedure for obtaining independent draws is to retain one draw for every ℓ draws where ℓ is typically between 5 and 100. In this paper, we report both the BMSE and standard deviation defined, respectively, by (22) and (23).

3.3 MCMC Results

The hyperparameters are, respectively, $\omega = 20.0$, $\gamma = 1.5$, $\nu = 10.0$, $S_\tau = 0.1$, $\varphi_0 = 0.0$, $\mu_0 = 1.0$, $p_0 = 2.0$, $q_0 = 2.0$, and the prior of δ is assumed to be $N(-0.25, 2)$. These values indicate that the prior of each parameter is very flat. Also our experience shows that the outcome of a MCMC simulation does not rely on different chioces of these hyperparameters. The burn-in period of the sampling algorithm consists of 50,000 iterations, and the posterior sample of the parameter vector consists of $N = 500,000$ iterations.

We apply the asymmetric BCSV model and the proposed sampling algorithm to the generated data. In order to remove the effect of possible serial correlation in the posterior sample (as the posterior sample forms a Markov chain), we retain one draw for every 50 draws during MCMC iterations. Table 1 summarizes the MCMC output, including the posterior mean, the 95% confidence interval, the BMSE and the standard deviation. The retained draws for each parameter are plotted in the left-hand panel of Figure 1, where the right-hand panel is a column of histograms obtained through the retained posterior samples of parameters. Figure 2 plots the generated volatility and the sampled volatility.

We can obtain the following evidence from the MCMC simulation. First, both the BMSE and Figure 1 indicate that the MCMC simulation has been mixing very well. Second, the posterior means of parameters are very close to their corresponding true values, and the posterior mean of the volatility approximates the true volatility very well. Third, the 95% confidence interval of δ does not cover either 0 or 0.5, indicating strong evidence to support the Box-Cox transformation of volatility against the alternative of the logarithmic or square-root transformation.

4 Application to the Equity Market

4.1 Data

This section will explore the application of the asymmetric BCSV model to daily returns of major stock indexes, which are the Dow Jones Industrial Average (DJIA), S&P 500, New York Stock Exchange (NYSE) composite, Nasdaq 100, Nikkei 225 and Hang Seng indexes. The historical data on these indexes were downloaded from Data Stream. Let p_t denote the asset price at time t and $x_t = \ln p_t - \ln p_{t-1}$ represent the continuously compounded return. As required by the construction of SV models, the return series should be mean-corrected and variance-scaled which is defined by

$$y_t = \frac{x_t - \overline{x}}{\sqrt{(n-1)^{-1}\sum_{i=1}^{n}(x_i - \overline{x})^2}},$$

where $\overline{x}$ is the mean of observed return series $\{x_t : t = 1, 2, \cdots, n\}$. The observed returns of Nasdaq 100 index are adopted for the period from the 1st January 1983 to 31st December 2002, while the other datasets are from the 1st January 1988 to 31st December 1998. All datasets exclude weekends and holidays. The purpose of the empirical study in this section is to seek empirical evidence to support the Box-Cox transformation of volatility in preference to the logarithmic or the square-root transformations, as well as to address the importance of the incorporation of leverage effects into the BCSV model.

4.2 Empirical Results

The hyperparameters required in the joint prior density are set, respectively, to $\omega = 20.0$, $\gamma = 1.5$, $\nu = 2.0$, $S_\tau = 0.01$, $\varphi_0 = 0.0$, $\mu_0 = -0.7$, $p_0 = 2.0$, $q_0 = 2.0$, and the prior of δ is assumed to be $N(-0.25, 2)$. We apply the asymmetric BCSV model and the sampling algorithm to daily returns of the DJIA index. During the implementation of MCMC iterations, we retain one draw for every 40 draws so as to remove the effect of possible serial correlation in the posterior sample. The retained draws for each parameter are plotted in the left-hand panel of Figure 2, while its right-hand panel is a column of histograms obtained based on the retained posterior samples, respectively. Table 2 summarizes the empirical output, including the posterior mean and the 95% confidence interval, the BMSE and the standard deviation.

We obtained the following evidence from the empirical example. First, both the BMSE and Figure 1 indicate that the proposed sampling algorithm has been mixing very well. Second, the posterior mean of δ is statistically different from zero, because its 95% confidence interval does not cover zero. The significance of δ is strong evidence to support the Box-Cox transformation of volatility against the alternative of the logarithmic transformation. As a consequence, the logarithmic transformation of volatility should be rejected when modeling the volatility of daily returns of DJIA index. Moreover, the posterior mean of δ is negative and the 95% confidence interval of δ does not cover 0.5. Therefore, the square-root transformation for volatility should also be excluded. Third, the 95% confidence interval of ρ does not cover zero. This is strong empirical evidence to support the incorporation of leverage effects into the asymmetric BCSV model.

Then we applied the asymmetric BCSV model and the sampling algorithm to the remaining datasets containing daily return series of Nasdaq 100, NYSE composite, S&P 500, Nikkei 225 and Hang Seng indexes. Table 2 presents a summary of the retained posterior samples of the parameter vector for each dataset. The sampling algorithm achieves very good mixing performance for each dataset. To save space, we shall not present the plot and histogram of the retained posterior sample for each dataset. We found in each dataset that the posterior mean of δ is statistically different from zero, indicating that the Box-Cox transformation of volatility is more favorable than the logarithmic transformation. Moreover, the 95% confidence interval of δ does not cover 0.5, indicating that the square-root transformation of volatility should also be excluded. In all these datasets, the 95% confidence interval of ρ does not cover zero, showing the importance of incorporating the leverage effect into the asymmetric BCSV model.

5 Conclusion

This paper developed a sampling algorithm for the asymmetric BCSV model where the leverage effect is incorporated to capture the dynamics of returns of asset prices. The Box-Cox transformation of volatility encompasses the logarithmic and square-root transformations as special cases by setting δ to 0 and 0.5, respectivly. Hence the specification of the BCSV model provide a possibility for model selection through a Bayesian approach. By applying the asymmetric BCSV model and the proposed sampling algorithm to the equity market, we found strong evidence to support the Box-Cox transformation of volatility against the alternative of the logarithmic or the square-root transformation. In addition, the empirical study on major stock indexes showed that the correlation between the errors in the mean and volatility equations plays an important role in SV models and captures the leverage effect. Hence it is important to incorporate the leverage effect into the BCSV model when the volatility of returns on a stock index is under investigation.

Appendix: An Alternative Method to Obtain the Joint Posterior

The Box-Cox stochastic volatility model is described by,

$$
\begin{aligned}
y_t &= \sqrt{g(\alpha_t, \delta)}\varepsilon_t, \\
\alpha_{t+1} &= \mu + \phi(\alpha_t - \mu) + u_{t+1},
\end{aligned}
$$

where

$$g_t = g(\alpha_t, \delta) = \begin{cases} (1+\delta\alpha_t)^{1/\delta} & \delta \neq 0 \\ \exp(\alpha_t) & \delta = 0 \end{cases},$$

the covariance matrix of (ε_t, u_{t+1}) is

$$\Sigma = \begin{pmatrix} 1 & \rho\sigma_u \\ \rho\sigma_u & \sigma_u^2 \end{pmatrix},$$

for $t = 1, 2, \cdots, n-1$, $u_1 \sim N(0, (1-\rho^2)\sigma_u^2/(1-\phi^2))$ and $\varepsilon_n \sim N(0,1)$. Then we re-parameterize ρ and σ through $\varphi = \rho\sigma_u$ and $\tau^2 = (1-\rho^2)\sigma_u^2$ which is the same as the re-parameterization presented in Jacquier, Polson and Rossi (2002).

Let $\theta = (\phi, \delta, \mu, \rho, \sigma_u)'$, $\varepsilon = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_{n-1})'$ and $\mathbf{u} = (u_2, u_3, \cdots, u_n)'$, and let $\mathbf{y}$ be the vector of observed returns and α be the vector of latent volatilities. Given the joint prior, which is the same as that presented in Section 2, we can obtain the posterior,

$$\begin{aligned} p(\theta, \alpha|\mathbf{y}) &\propto p(\phi, \delta, \mu, \rho, \sigma) \times p(y_1, y_2, \cdots, y_{n-1}; \alpha_2, \alpha_3, \cdots, \alpha_n|\theta) \times p(y_n|\theta) \times p(\alpha_1|\theta) \\ &= p(\phi, \delta, \mu, \rho, \sigma) \times \prod_{t=1}^{n-1} g_t^{-1/2} p(y_t\ g_t^{-1/2}, \alpha_{t+1}|\alpha_t, \theta) \times p(y_n|\theta) \times p(\alpha_1|\theta) \\ &= p(\phi, \delta, \mu, \rho, \sigma)|\Sigma|^{-(n-1)/2} \times \left(\prod_{t=1}^{n-1} g_t^{-1/2}\right) \times \exp\left\{-\frac{1}{2}tr(\Sigma^{-1}A)\right\} \times p(y_n|\theta) \times p(\alpha_1|\theta), \end{aligned}$$

where $|\Sigma| = \tau^2$ and

$$A = \sum_{t=1}^{n-1} (\varepsilon_t\ \ u_{t+1})'(\varepsilon_t\ \ u_{t+1}).$$

Jacquier, Polson and Rossi (2002) show that

$$\Sigma^{-1} = \frac{1}{\tau^2}\begin{pmatrix} \varphi^2 & -\varphi \\ -\varphi & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \frac{C}{\tau^2} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Thus,

$$tr(\Sigma^{-1}A) = \frac{1}{\tau^2}tr(CA) + a_{11}.$$

Then the posterior is

$$\begin{aligned} p(\theta, \alpha|\mathbf{y}) &= p(\phi, \delta, \mu, \rho, \sigma) \times \prod_{t=1}^{n-1} g_t^{-1/2} \times \left(\frac{1}{\tau^2}\right)^{(n-1)/2} \exp\left\{-\frac{1}{2\tau^2}tr(CA) - \frac{1}{2}a_{11}\right\} \\ &\quad \times p(y_n|\theta) \times p(\alpha_1|\theta), \end{aligned}$$

where

$$
\begin{aligned}
p(y_n|\theta) &= \frac{1}{\sqrt{2\pi g_n}} \exp\left\{-\frac{y_n^2}{2g_t}\right\}, \\
p(\alpha_1|\theta) &= \frac{1}{\sqrt{2\pi\tau^2/(1-\phi^2)}} \exp\left\{-\frac{(\alpha_1-\mu)^2}{2\tau^2/(1-\phi^2)}\right\}, \\
tr(CA) &= \sum_{t=1}^{n-1} \left(\alpha_{t+1}-\mu-\phi(\alpha_t-\mu)-\varphi g_t^{-1/2} y_t\right)^2.
\end{aligned}
$$

Substituting the priors into the above equation, we obtain the joint posterior

$$
\begin{aligned}
p(\theta,\alpha|\mathbf{y}) &= p(\delta)\times(1+\phi)^{\omega-1/2}(1-\phi)^{\gamma-1/2} \\
&\quad \times \left(\prod_{t=1}^{n} g_t^{-1/2}\right) \exp\left\{-\frac{1}{2}\sum_{t=1}^{n}\frac{y_t^2}{g_t}\right\} \times \left(\frac{1}{\tau^2}\right)^{(n+v+2)/2+1} \times \exp\left\{-\frac{\kappa}{2\tau^2}\right\},
\end{aligned}
$$

where

$$
\kappa = (1-\phi^2)(\alpha_1-\mu)^2 + tr(CA) + p_0(\varphi-\varphi_0)^2 + q_0(\mu-\mu_0)^2 + S_\tau.
$$

Hence the joint posterior obtained here is identical to (12) which is obtained through a transformation of ε_t and u_{t+1}.

References

1. Andersen, T. (1994), "Stochastic Autoregressive Volatility: A Framework for Volatility Modelling," *Mathematical Finance*, **4**, 75-102.

2. Barndorff-Nielsen, N., and Shephard, N. (2001), "Non-Gaussian Ornstein–Uhlenbeck-based Models and Some of Their Uses in Financial Economics," *Journal of the Royal Statistical Society, Series B*, **63**, 167-241.

3. Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroscedasticity," *Journal of Econometrics*, **31**, 307-327.

4. Box, G.E.P., and Cox, D.R. (1964), "An Analysis of Transformations," *Journal of Royal Statistical Society B*, **26**, 211-243.

5. Campbell, J.Y., and Kyle, A.S. (1993), "Smart Money, Noise Trading and Stock Price Behaviour," *Review of Economic Studies*, **60**, 1-34.

6. Chib, S. (2001), "Markov Chain Monte Carlo Methods: Computation and Inference," in *Handbook of Econometrics*, eds. J.J. Heckman and E. Leamer, Amsterdam: North-Holland, **5**, pp.3569-3649.

7. Chib, S., Nardari, F., and Shephard, N. (2002), "Markov Chain Monte Carlo Methods for Stochastic Volatility Models," *Journal of Econometrics*, **108**, 281-316.

8. Clark, P.K. (1973), "A Subordinated Stochastic Process Model with Finite Variance for Speculative Price," *Econometrica*, **41**, 135-155.

9. Cox, J.C., Ingersoll, J.E., and Ross, S.A. (1985), "A Theory of Term Structure of Interest Rates," *Econometrica*, **53**, 385-408.

10. Engle, R.F. (1982), "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, **50**, 987-1007.

11. Engle, R.F., and Ng, V.K. (1993), "Measuring and Testing the Impact of News on Volatility," *The Journal of Finance*, **48**, 1749-1801.

12. Eraker, B., Johannes, M., and Polson, N. (2002), "The Impact of Jumps in Volatility and Returns," *The Journal of Finance*, forthcoming.

13. Fouque, J.-P., Papanicolaou, G., and Sircar, K.R. (2000), *Derivatives in Financial Markets with Stochastic Volatility*, New York: Cambridge University Press.

14. Gallant, A.R., Hsieh, D., and Tauchen, G. (1997), "Estimation of Stochastic Volatility Models with Diagnostics," *Journal of Econometrics*, **81**, 159-192.

15. Gallant, A.R., Rossi, P.E., and Tauchen, G. (1992), "Stock Price and Volume," *Review of Financial Studies*, **5**, 199-242.

16. Gallant, A.R., Rossi, P.E., and Tauchen, G. (1993), "Nonlinear Dynamic Structures," *Econometrica*, **61**, 871-907.

17. Gallant, A.R., and Tauchen, G. (1996), "Which Moments to Match?" *Econometric Theory*, **12**, 657-681.

18. Ghysels, E., Harvey, A.C., and Renault, E. (1996), "Stochastic Volatility," in *Handbook of Statistics*, eds. G.S. Maddala and C.R. Rao, Amsterdam: North-Holland, **14**, pp.119-191.

19. Harvey, A.C., and Shephard, N. (1996), "The Estimation of an Asymmetric Stochastic Volatility Model for Asset Returns," *Journal of Business and Economic Statistics*, **14**, 429-434.

20. Heston, S.L. (1993), "A Closed-Form Solution for Options with Stochastic Volatility, with Application to Bond and Currency Options," *Review of Financial Studies*, **6**, 327-343.

21. Hull, J., and White, A. (1987), "The Pricing of Options on Assets with Stochastic Volatilities," *Journal of Finance*, **42**, 281-300.

22. Jacquier, E., Polson, N.G., and Rossi, P.E. (1994), "Bayesian Analysis of Stochastic Volatility Models," *Journal of Business and Economic Statistics*, **12**, 371-389.

23. Jacquier, E., Polson, N.G., and Rossi, P.E. (2002), "Bayesian Analysis of Stochastic Volatility Models with Fat-Tails and Correlated Errors," *Journal of Econometrics*, forthcoming.

24. Johannes, M., and Polson, N. (2003), "MCMC Methods for Financial Econometrics," unpublished manuscript, Columbia Business School.

25. Johnson, H., and Shanno, D. (1987), "Option Pricing When the Variance Is Changing," *Journal of Financial and Quantitative Analysis*, **22**, 143-152.

26. Kim, S., Shephard, N., and Chib, S. (1998), "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models," *Review of Economic Studies*, **65**, 361-393.

27. Melino, A., and Turnbull, S.M. (1990), "Pricing Foreign Currency Options with Stochastic Volatility," *Journal of Econometrics*, **45**, 239-265.

28. Meyer, R., and Yu, J. (2000), "BUGS for a Bayesian Analysis of Stochastic Volatility Models," *Econometrics Journal*, **3**, 198-215.

29. Nelson, D.B. (1991), "Conditional Heteroskedasticity in Asset Returns: A New Approach," *Econometrica*, **59**, 347-370.

30. Roberts, G.O. (1996), "Markov Chain Concepts Related to Sampling Algorithms," in *Markov chain Monte Carlo in Practice*, eds. W.R. Gilks, S. Richardson, and D.J. Spiegelhalter, London: Chapman & Hall, pp.45-57.

31. Shephard, N. (1993), "Fitting Nonlinear Time Series Models with Applications to Stochastic Variance Models," *Journal of Applied Econometrics*, **8**, S135-S152.

32. Shephard, N., and Pitt, M.K. (1997), "Likelihood Analysis of Non-Gaussian Measurement Time Series," *Biometrika*, **84**, 653-667.

33. Stein, E.M., and Stein, J.C. (1991), "Stock Price Distributions with Stochastic Volatility: An Analytical Approach," *Review of Financial Studies*, **4**, 727-752.

34. Tauchen, G., and Pitts, M. (1983), "The Price Variability-Volume Relationship on Speculative Markets," *Econometrica*, **51**, 485-505.

35. Taylor, S.J. (1982), "Financial Returns Modelled by the Product of Two Stochastic Processes – A Study of the Daily Sugar Prices 1961-75," in *Time Series Analysis: Theory and Practice*, eds. O.D. Anderson, Amsterdam: North-Holland, **1**, pp.203-226,

36. Taylor, S.J. (1986), *Modelling Financial Time Series*, Chichester: Wiley.

37. Taylor, S.J. (1994), "Modelling Stochastic Volatility: A Review and Comparative Study," *Mathematical Finance*, **4**, 183-204.

38. Yu, J. (2002), "MCMC Methods for Estimating Stochastic Volatility Models with Leverage Effects: Comments on Jacquier, Polson and Rossi (2002)," working paper, University of Auckland.

39. Yu, J., Yang, Z., and Zhang, X. (2002), "A Class of Nonlinear Stochastic Volatility Models and Its Implication on Pricing Currency Options," working paper, Monash University.

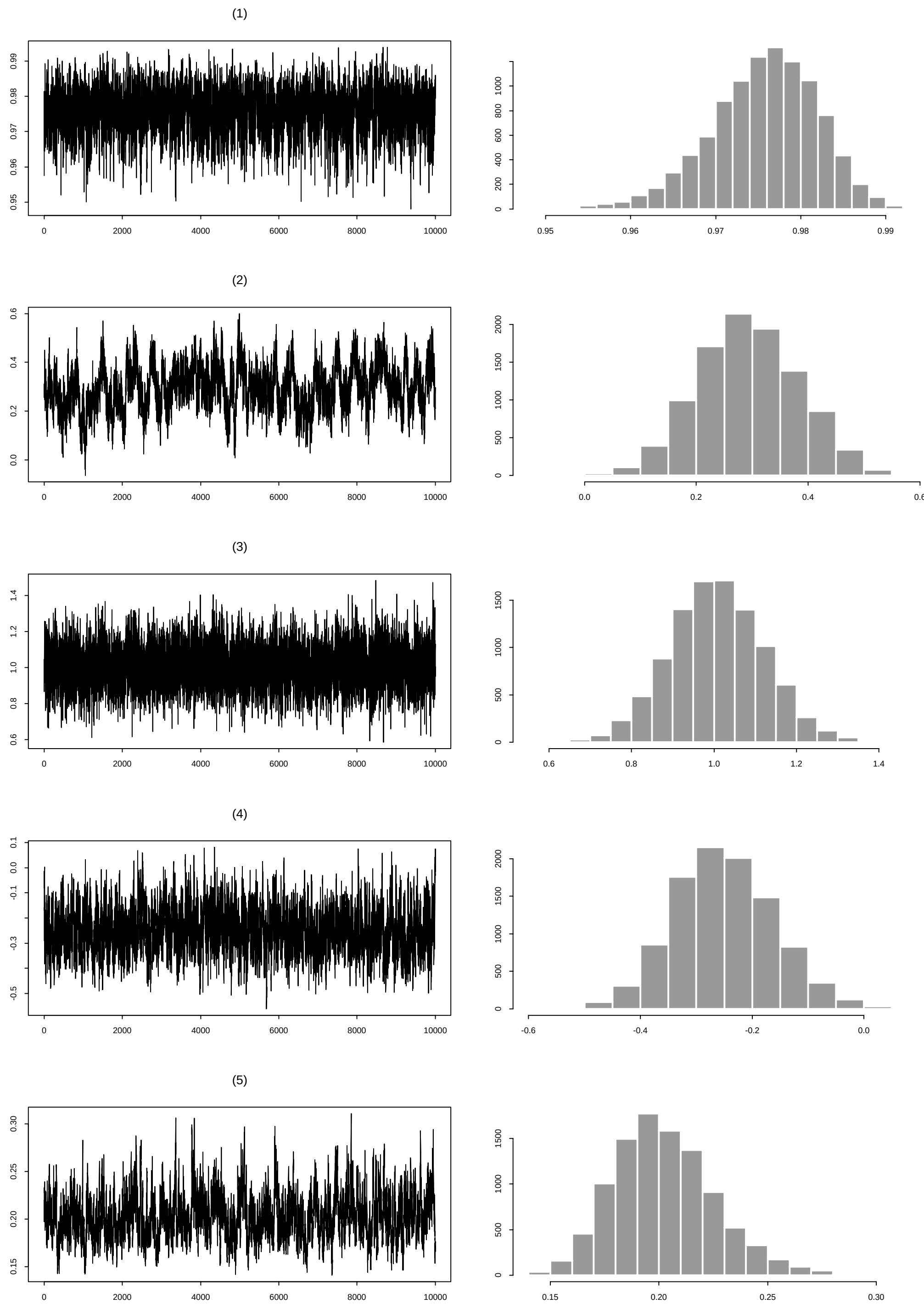

Figure 1. Posterior samples of parameters obtained by applying the BCSV model to a generated dataset. The left-hand panel plots the retained posterior samples of σ, ρ, μ, δ and ϕ, respectively, while the right-hand panel plots the corresponding histograms.

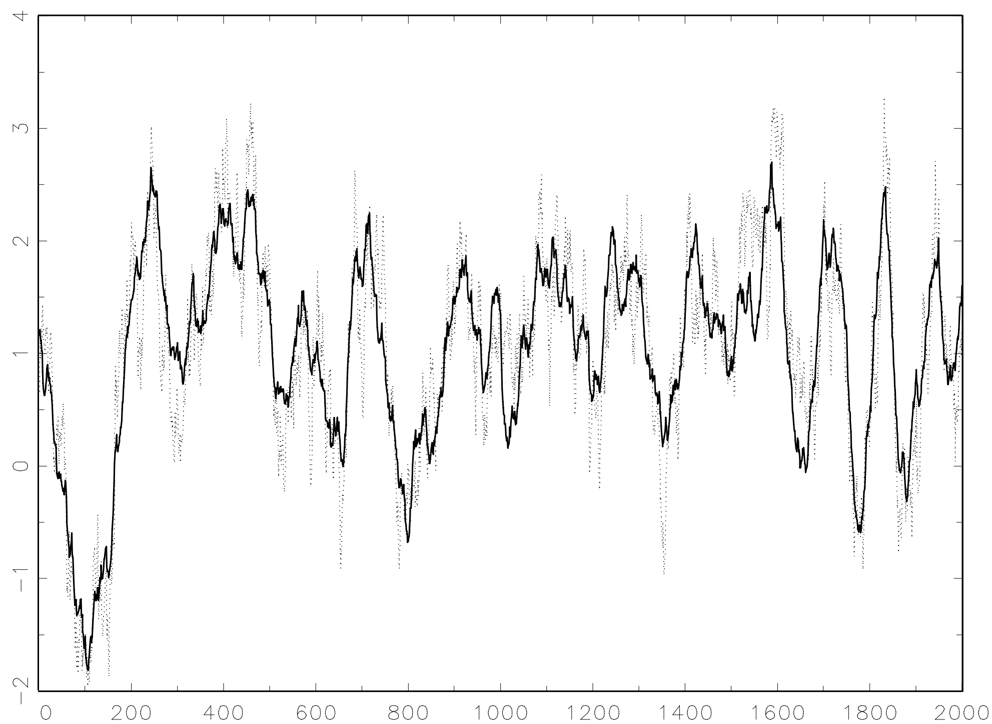

Figure 2. The generated volatilities (dotted line) and sampled volatility (solid line).

Table 1. Summary of the posterior sample obtained from the generated data

true parameter	mean	95% confidence interval	BMSE	s.d.
$\phi = 0.98$	0.97577	(0.9621, 0.9868)	0.00016	0.00629
$\delta = 0.30$	0.29277	(0.1166, 0.4704)	0.00755	0.09059
$\mu = 1.00$	1.00648	(0.7808, 1.2343)	0.00266	0.11438
ρ =-0.25	-0.24995	(-0.4170, -0.0660)	0.00340	0.08937
$\sigma_u = 0.20$	0.20246	(0.1614, 0.2553)	0.00141	0.02356

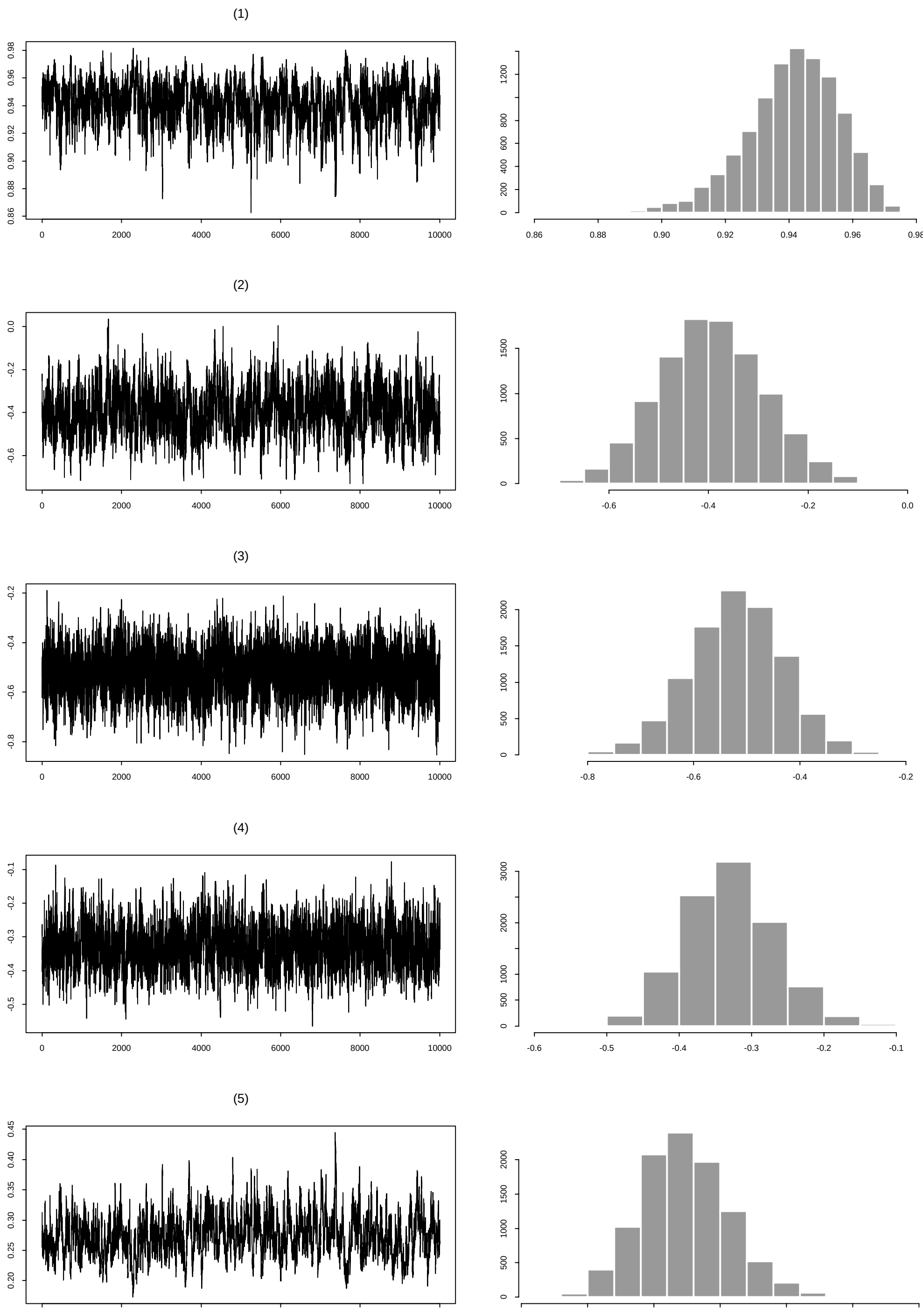

Figure 3. Posterior samples of parameters obtained by applying the BCSV model to daily returns of DJIA index. The left-hand panel plots the retained posterior samples of ϕ, δ, μ, ρ and σ, respectively, while the right-hand panel plots the corresponding histograms.

Table 2. Summary of posterior samples obtained from daily returns of stock indexes

data	parameter	mean	95% confidence interval	BMSE	s.d.
Dow Jones Industrial Average	ϕ	0.94129	(0.9091, 0.9659)	0.00089	0.01462
	δ	-0.39339	(-0.5940, -0.1805)	0.00553	0.10635
	μ	-0.51855	(-0.6974, -0.3508)	0.00270	0.08768
	ρ	-0.33074	(-0.4478, -0.2030)	0.00219	0.06228
	σ_u	0.27368	(0.2127, 0.3442)	0.00224	0.03322
Nasdaq 100	ϕ	0.98337	(0.9767, 0.9891)	0.00012	0.00315
	δ	-0.19324	(-0.3124, -0.0678)	0.00547	0.06120
	μ	-0.67996	(-0.8605, -0.5010)	0.00177	0.09201
	ρ	-0.36513	(-0.4574, -0.2697)	0.00164	0.04791
	σ_u	0.18075	(0.1555, 0.2096)	0.00077	0.01380
NYSE Composite	ϕ	0.93845	(0.9014, 0.9621)	0.00075	0.01549
	δ	-0.25038	(-0.6629, -0.0484)	0.00544	0.10025
	μ	-0.52076	(-0.7010, -0.3484)	0.00267	0.08986
	ρ	-0.35607	(-0.4583, -0.2338)	0.00140	0.05681
	σ_u	0.29939	(0.2396, 0.3857)	0.00196	0.03679
S&P 500	ϕ	0.94549	(0.9139, 0.9703)	0.00083	0.01431
	δ	-0.31273	(-0.5147, -0.1108)	0.00566	0.10218
	μ	-0.60931	(-0.8305, -0.4179)	0.00360	0.10440
	ρ	-0.35311	(-0.4624, -0.2304)	0.00156	0.05903
	σ_u	0.28813	(0.2251, 0.3597)	0.00210	0.03395
Nikkei 225	ϕ	0.97640	(0.9665, 0.9846)	0.00020	0.00460
	δ	-0.24566	(-0.3888, -0.1021)	0.00677	0.07540
	μ	-0.83664	(-1.0391, -0.6357)	0.00307	0.10209
	ρ	-0.63456	(-0.7248, -0.5340)	0.00248	0.04905
	σ_u	0.22521	(0.1918, 0.2630)	0.00101	0.01822
Hang Seng	ϕ	0.92644	(0.8987, 0.9501)	0.00038	0.01321
	δ	-0.28059	(-0.3928, -0.1619)	0.00248	0.05832
	μ	-0.93793	(-1.1619, -0.7191)	0.00253	0.11327
	ρ	-0.29893	(-0.3972, -0.1995)	0.00119	0.05006
	σ_u	0.42457	(0.3583, 0.4995)	0.00121	0.03642

Note: The batch size for computing BMSE is 10,000 and there are 50 batches. s.d. refers to the standard deviation computed through (23) based on draws by retaining one draw for every 50 draws.