A Class of Nonlinear Stochastic Volatility Models and Its Implications on Pricing Currency Options[1]

Jun Yu
Department of Economics
University of Auckland
Private Bag 92019
Auckland, New Zealand
j.yu@auckland.ac.nz

Zhenlin Yang
School of Economics and Social Sciences
Singapore Management University
Singapore 259756
zlyang@smu.edu.sg

Xibin Zhang
Department of Econometrics and Business Statistics
Monash University
Clayton, Vic., 3800, Australia
bill.zhang@buseco.monash.edu.au

This version: November 2002
First version: April 2002

JEL classification: C22, C52, G12
Keywords: Box-Cox Transformation; Stochastic Volatility; MCMC; Exchange Rate Volatility; Option Pricing.

[1]This is a revised version of a paper with the title "A Class of Nonlinear Stochastic Volatility Models" by Jun Yu and Zhenlin Yang. That version of the paper used different datasets as well as different estimation techniques and did not have the section on option pricing. Jun Yu has benefited greatly from conversations with Ronald Gallant and Genshiro Kitagawa, from comments and suggestions from Henk Berkman, Charles Corrado, Andrew Harvey, Steve Heston, John Knight, Michael McAleer, Nour Meddahi, Esther Ruiz, Steve Satchell, Ken Singleton, Neil Shephard and Yiu Kuen Tse, and seminar participants at the University of Auckland (both Department of Economics and Department of Accounting and Finance), Yale University, Econometric Society Australasian Meeting in Brisbane, and New Zealand Econometrics Study Group Meetings in Dunedin. He also gratefully acknowledges financial support from the Royal Society of New Zealand Marsden Fund under number 01-UOA-015. This research was started while Jun Yu was visiting the National University of Singapore in 2001 and he would like to thank the Department of Economics at National University of Singapore for their hospitality. Thanks also go to Calvin Chan and Shirley Huang at the University of Auckland for computing assistant with the first version of the paper. Corresponding author is Jun Yu.

Abstract

This paper proposes a class of stochastic volatility (SV) models which offers an alternative to the one introduced in Andersen (1994). The class encompasses all standard SV models that have appeared in the literature, including the well known lognormal model, and allows us to empirically test all standard specifications in a convenient way. We develop a likelihood-based technique for analyzing the class. Daily dollar/pound exchange rate data reject all the standard models and suggest evidence of nonlinear SV. An efficient algorithm is proposed to study the implications of this nonlinear SV on pricing currency options and it is found that the lognormal model overprices options.

1 Introduction

Modelling the volatility of financial time series via stochastic volatility (SV) models has received a great deal of attention in the theoretic finance literature as well as in the empirical literature. Prices of options based on SV models are shown to be more accurate than those based on the Black-Scholes model (see, for example, Melino and Turnbull (1990)). Moreover, the SV model offers a powerful alternative to GARCH-type models to explain the well documented time varying volatility. Empirical successes of the lognormal SV model relative to GARCH-type models are documented in Danielsson (1994), Geweke (1994b), and Kim, Shephard and Chib (1998) in terms of in-sample fitting, and in Yu (2002) in terms of out-of-sample forecasting.

The most widely used SV model is the lognormal specification which is built upon the models of Clark (1973) and Tauchen and Pitt (1983) and first introduced by Taylor (1982, 1986 and 1994). It has been used to price stock options in Wiggins (1987) and Scott (1987) and currency options in Chesney and Scott (1989). Since it assumes that the logarithmic volatility follows an Ornstein-Uhlenbeck (OU) process, an implication of this specification is that the marginal distribution of logarithmic volatility is normal. This assumption has very important implications for financial economics and risk management.

Many other SV models coexist in the theoretical finance literature as well as in the empirical literature. For example, Stein and Stein (1991) and Johnson and Shanno (1987) assume the square root of volatility follows, respectively, an OU process and a geometric Brownian motion, while Hull and White (1987) and Heston (1993) assume a geometric Brownian motion and a square-root process for volatility. In the discrete time case, various SV models can be regarded as generalizations to the corresponding GARCH models. For example, a polynomial SV model is a generalization of GARCH(1,1) (Bollerslev (1986)) while a square root polynomial SV model is a generalization of standard deviation (SD)-GARCH(1,1). Andersen (1994) introduces a general class of SV models, of which a class of polynomial SV models has been emphasized. This class encompasses most of the discrete time SV models in the literature. Other more recent classes of SV models include those proposed by Barndorff-Nielsen and Shephard (2001) and by

Meddahi (2001).

Despite all these alternative specifications, there is a lack of procedure for selecting an appropriate functional form of stochastic volatility.[2] The specification of the correct stochastic volatility function, on the other hand, is very important in several respects. First, different functional forms lead to different formulae for option pricing. Misspecification of the stochastic volatility function can result in incorrect option prices. Second, the marginal distribution of volatility depends upon the functional form of stochastic volatility.

In this paper, we propose a new class of SV models, namely, nonlinear SV models. Like the class of Andersen (1994), it includes as special cases many SV models that have appeared in the literature. It overlaps with but does not encompass the class of Andersen. Different from his class which precludes a simple comparison of different SV models, an advantage of our proposed class is the ease with which different specifications on stochastic volatility can be tested. In fact, the specification test is based on a single parameter. Furthermore, as a byproduct of this general way of modelling stochastic volatility, one obtains the functional form of transformation which induces marginal normality of volatility. We empirically test all standard specifications against our general specification using daily dollar/pound data. Our empirical test of all standard SV models is, to the best of our knowledge, the first in the literature. The empirical test rejects all standard SV models and favors a nonlinear SV specification. Implications of this nonlinearity on some important financial variables are examined. For example, without sacrificing the overall goodness-of-fit, our nonlinear SV model improves the fit to data when the market has little movement. We also find that our model implies a smoother volatility series. Moreover, the marginal distribution of volatility is different from a lognormal distribution. Most importantly, an application of our nonlinear SV model to option pricing shows that the lognormal SV model overprices currency options, particularly out-of-the-money options.

The paper is organized as follows. Section 2 presents this class of nonlinear SV

[2]It is well known that a GARCH process converges to a relevant stochastic volatility process (Nelson (1990)). A specification test based on a GARCH family can be suggestive of an appropriate stochastic volatility specification; see for example, Hentschel (1995). Such a test, however, is by no mean a direct test of stochastic volatility specifications.

models. In Section 3, a Markov Chain Monte Carlo (MCMC) method is developed to provide likelihood-based analysis of the proposed class of models. The class is fitted to daily observations on dollar/pound exchange rate series in Section 4. In Section 5 we illustrate the importance of the proposed models in terms of their implications on pricing currency options. In Section 6 we apply the new models to analyze four other exchange rates. Finally in Section 7 we present conclusions and possible extensions.

2 A Class of Nonlinear SV Models

In the theoretic finance literature on option pricing, the SV model is often formulated in terms of stochastic differential equations. For instance, Wiggins (1987), Chesney and Scott (1989), and Scott (1991) specify the following model for the asset price $P(t)$ and the corresponding volatility $\sigma^2(t)$,

$$dP(t)/P(t) = \alpha dt + \sigma(t) dB_1(t), \tag{2.1}$$

$$d\ln\sigma^2(t) = \lambda(\xi - \ln\sigma^2(t))dt + \gamma dB_2(t), \tag{2.2}$$

where $B_1(t)$ and $B_2(t)$ are two Brownian motions and $corr(dB_1(t), dB_2(t)) = \rho$ with ρ capturing the so-called leverage effect.

In the empirical literature, the above continuous time model is often discretized. The discrete time SV model may be obtained, for example, via the Euler-Maruyama approximation. The approximation, after a location shift and reparameterization, leads to the *lognormal* SV model given by

$$X_t = \sigma_t e_t, \tag{2.3}$$

$$\ln\sigma_t^2 = \mu + \phi(\ln\sigma_{t-1}^2 - \mu) + \sigma v_t, \tag{2.4}$$

where X_t is a continuously compounded return and e_t, v_t are two sequences of independent and identically distributed (iid) $N(0,1)$ random variables with $corr(e_t, v_{t+1}) = \rho$. The above model is equivalently represented, in the majority of empirical literature, by

$$X_t = \exp(\frac{1}{2}h_t)e_t, \tag{2.5}$$

$$h_t = \mu + \phi(h_{t-1} - \mu) + \sigma v_t, \tag{2.6}$$

where $h_t = \ln \sigma_t^2$.

The lognormal SV model specifies that the logarithmic volatility follows an AR(1) process. However, this relationship may not always be warranted by the data. A natural generalization to this relationship is to allow a general (nonlinear) smooth function of volatility to follow an AR(1) process. That is,

$$X_t = \sigma_t e_t, \tag{2.7}$$

$$h(\sigma_t^2, \delta) = \mu + \phi[h(\sigma_{t-1}^2, \delta) - \mu] + \sigma v_t, \tag{2.8}$$

where e_t and v_t are two $N(0,1)$ sequences with $corr(e_t, v_{t+1}) = \rho$, and $h(\cdot, \delta)$ is a smooth function indexed by a parameter δ. A nice choice of this function is the Box-Cox power function (Box and Cox (1964)):

$$h(t, \delta) = \begin{cases} (t^\delta - 1)/\delta, & \text{if } \delta \neq 0, \\ \ln t, & \text{if } \delta = 0. \end{cases} \tag{2.9}$$

As the function $h(\cdot, \delta)$ is specified as a general nonlinear function, the model is thus termed in this paper the *nonlinear* SV (N-SV hereafter) model. Several attractive features of this new class of SV models include: i) as we will show below it includes the lognormal SV model and the other popular SV models as special cases, ii) it adds great flexibility to the functional form, and iii) it allows a simple test for the lognormal SV specification, i.e., a test of $H_0 : \delta = 0$, and some other "classical" SV specifications. If we write $h_t = h(\sigma_t^2, \delta)$, then we can re-write the N-SV models as

$$X_t = [g(h_t, \delta)]^{1/2} e_t, \tag{2.10}$$

$$h_t = \mu + \phi(h_{t-1} - \mu) + \sigma v_t, \tag{2.11}$$

where $g(h_t, \delta)$ is the inverse Box-Cox transformation of the form

$$g(h_t, \delta) = \begin{cases} (1 + \delta h_t)^{1/\delta}, & \text{if } \delta \neq 0, \\ \exp(h_t), & \text{if } \delta = 0. \end{cases} \tag{2.12}$$

Denote the vector of model parameters by $\theta = (\mu, \delta, \phi, \sigma, \rho)$.

The idea of our proposed N-SV models is similar to that made in Higgins and Bera (1992) from the linear ARCH model (Engle (1982)) to the nonlinear ARCH (NARCH) model. Obviously, our model provides a stochastic volatility generalization of a nonlinear GARCH(1,1) model.

It can be seen as $\delta \to 0$, $(1+\delta h_t)^{1/(2\delta)} \to \exp(0.5h_t)$ and $((\sigma_t^2)^\delta - 1)/\delta \to \ln \sigma_t^2$. Hence the proposed N-SV model includes the lognormal SV model as a special case. If $\delta = 1$, the variance equation (2.8) becomes

$$\sigma_t^2 = \mu' + \phi(\sigma_{t-1}^2 - \mu') + \sigma v_t, \tag{2.13}$$

where $\mu' = \mu + 1$. This is a polynomial SV model in Andersen (1994). According to this specification, volatility follows a normal distribution as its marginal distribution. If $\delta = 0.5$, the variance equation (2.8) becomes

$$\sigma_t = \mu'' + \phi(\sigma_{t-1} - \mu'') + 0.5\sigma v_t, \tag{2.14}$$

where $\mu'' = 0.5\mu + 1$. This is a square root polynomial SV model in Andersen (1994) and can be regarded as a discrete time version of the continuous time SV model in Scott (1987) and Stein and Stein (1991). As a result, the marginal distribution of the square root of volatility is Gaussian.

In Table 1 we summarize some well-known SV models and show their parameter relations with our model. For the continuous time SV models, their Euler discrete time versions are considered. It can be seen that all these models can be obtained from our model by placing the appropriate restrictions on the three parameters δ, μ and ϕ. In fact, all the models except our model require δ to be 0, 0.5, or 1.[3] For a general δ, our model is different from any of them and δ provides some idea about the degree of departure from a "classical" parametric SV model. See Figure 1 for the comparison of the square root of inverse Box-Cox transformation, $(1+\delta h_t)^{1/(2\delta)}$ (or σ_t), as a function of h_t for various values of δ over the interval $[-2, 2]$. This is a possible range that actual h_t may lie within in the framework of lognormal SV model.

The Box-Cox transformation has been applied in various areas in finance. One of the most relevant applications to our work may be that proposed by Higgins and Bera (1992)

[3]Some specifications in Table 1 may be different from the actual specifications used in the original references. However, they are equivalent to each other via Ito's lemma. For example, Heston (1993) adopts a square root specification for σ_t^2 which is identical to assuming σ_t follows a particular OU process.

who introduce the NARCH model. Another relevant application is Hentschel (1995) who introduces a family of GARCH models by applying the Box-Cox transformation to the conditional standard deviation. A nice feature of our proposed class is that it provides a simple way to test the null hypothesis of polynomial SV specifications against a variety of non-polynomial alternatives. Moreover, as a consequence of specification testing, our proposed class provides an effective channel to check the marginal distribution of unobserved volatility.

We now establish some basic statistical properties of the N-SV models. It is easy to see that h_t is stationary and ergodic if $\phi < 1$ and that if so

$$\mu_h \equiv E(h_t) = \mu,\ \sigma_h^2 \equiv Var(h_t) = \frac{\sigma^2}{1-\phi}, \text{ and } \rho(\ell) \equiv Corr(h_t, h_{t-\ell}) = \phi^{\ell}.$$

It follows that X_t is stationary and ergodic as it is the product of two stationary and ergodic processes. For the moments of X_t, a distributional constraint has to be imposed on v_t or h_t. As σ_t^2 is nonnegative, the exact normality of v_t is incompatible unless $\delta = 0$ or $1/\delta$ is an even integer.[4] Our experience suggests that, as far as statistical inferences and pricing options are concerned, the assumption of the exact normality of v_t works well for all the empirically possible values of parameters that we have encountered.[5] Unfortunately, even in the case where $1/\delta$ is an even integer, it does not seem to be possible to obtain an analytic form for the moments of the model. Moreover, unlike the lognormal SV model, it appears that there is no obvious way to linearize the mean equation (2.10). These two undesirable properties make the classical econometric treatments of SV models, such as generalized method of moments (GMM) and quasi maximum likelihood (QML), difficult to implement for the N-SV model.

To conclude this section, we attempt to offer a heuristic interpretation of δ from a finance perspective.[6] For ease of interpretation, we restrict ourselves to the range of

[4]This is the well known truncation problem with the Box-Cox power transformation. The truncation effect is negligible if $\delta\sigma_h/(1+\delta\mu)$ is small, which is achieved when i) δ is small, or ii) μ is large, or iii) σ_h is small. See Yang (1999) for a discussion on this.

[5]The same problem occurs in the model proposed by Stein and Stein (1991). They claim that "for a wide range of empirically reasonable parameter values, the probability of passing the barrier at $\sigma = 0$ is so small as to be of no significant consequence."

[6]Our treatment here is analogous with the introduction of continuously compounded returns. We are grateful to Steve Satchell for pointing this out to us.

positive δ. Define $m = 1/\delta$ and re-write the inverse Box-Cox transformation as

$$\sigma_t^2 = (1 + \frac{h_t}{m})^m = \prod_{i=1}^{m}(1 + h_{it}), \tag{2.15}$$

where $\{h_{it}\}$ can be understood as a sequence of intra-day volatility movement. From a market microstructure perspective, intra-day volatility movement are caused primarily by the arrival of new information. Therefore, according to equation (2.15) one can argue that on average there are m times per day of new information arrivals and h_t represents the average impact of the information on volatility. In the lognormal SV model, as $m \to \infty$ and $\sigma_t^2 \to \exp(h_t)$, new information arrives at the market very frequently. In the N-SV model with a positive, finite value of δ, say $\delta = 0.25$, on average new information arrives at the market 4 times per day.

3 Likelihood-Based Analysis of Nonlinear SV Models

3.1 Why Use MCMC?

The literature on estimating SV models is vast. This is in part due to the fact that the likelihood function has no closed form expression for SV models and hence the maximum likelihood approach is extremely difficult to implement. As a consequence, the SV model becomes a central example to compare the relative merits of alternative estimation procedures.

Less computationally intensive methods often involve no simulation. These include GMM (Andersen and Sorensen (1996) and Hansen and Scheinkman (1995)), QML (Harvey, Ruiz and Shephard (1994) and Ruiz (1994)), and the estimation method via the empirical characteristic function (Knight, Satchell and Yu (2002) and Singleton (2001)). Unfortunately, it is difficult to apply these methods in our settings due to the nonlinear structure in the mean equation. More efficient estimation methods often involve simulations and are generally computationally more expensive. These include simulated maximum likelihood methods proposed by Danielsson and Richard (1993) and Danielsson (1994) for estimating the discrete time SV model and by Durham and Gallant

(2002) for estimating the continuous time SV model; the maximum likelihood Monte Carlo method of Sandmann and Koopman (1998); the simulation method using important sampling and antithetic variables proposed by Durbin and Koopman (2000); the indirect inference proposed of Gouriéroux, Monfort and Renault; simulated method of moments of Duffie and Singleton (1993) and its refinement, efficient method of moments of Gallant and Tauchen (1996); Markov Chain Monte Carlo methods proposed by Jacquier, Polson and Rossi (1994) and improved by Kim et al. (1998) in the discrete time context and by Eraker (2001) in the continuous time context.

The relative merits of the alternative methods depend not only on the finite sample efficiency but also on the flexibility to adapt to modifications of model specification. Moreover, in the framework of SV models, a good method should also allow one to extract the unobserved volatility model with a low cost and to do simple but useful model diagnostics. Judged by these criteria, MCMC is our choice for inferences since it provides a flexible and highly efficient approach to analyzing SV models.

Andersen, Chung and Sorensen (1999) document a finite sample comparison of various methods in Monte Carlo studies and find that MCMC is the most efficient estimation tool. Their finding is not surprising since MCMC provides a full likelihood-based inference. Moreover, Meyer and Yu (2000) and Chib, Nardari and Shephard (2002) discuss its flexibility of modelling modifications of the lognormal SV model. Furthermore, as a byproduct of parameter estimation, MCMC methods provide estimates of latent volatility and predictive distributions for volatility (see for example Jacquier et al. (1994) and Eraker, Johannes and Polson (2001)). In addition to providing an efficient way for Bayesian inference, MCMC can also be used to calculate the likelihood, compute the filtered volatility estimates, and do diagnostic checking (see for example Kim et al. (1998)). As a consequence of likelihood evaluation, the likelihood ratio test can be used to compare model performance of alternative specifications. Alternatively, one can use Bayesian methods for model comparison. Examples include Bayes factors (Chib (1995)) and deviance information criterion (Berg, Meyer and Yu (2002)) and both can be obtained based on the MCMC output.

In this paper the proposed N-SV models are to be applied to exchange rate series. Although the leverage effect is particularly important for stock returns, it has been found to be much less severe for exchange rates (Meyer and Yu (2000)). Consequently, we impose a restriction into the N-SV model, that is, $\rho = 0$. Hence the vector of model parameters reduces to $\theta = (\mu, \delta, \phi, \sigma)$.[7]

3.2 Estimating Nonlinear SV Models

Instead of searching for the analytic expression of a posterior density, MCMC methods aim to provide a general mechanism to sample the parameter vector from its posterior density. In the context of SV models, it is well known that the intractable likelihood function $f(\mathbf{X}|\theta)$ makes the direct analysis of the posterior density $f(\theta|\mathbf{X})$ extremely difficult, where $\mathbf{X} = (X_1, X_2, \cdots, X_T)$. To circumvent this problem, a common practice is to augment the parameter vector to $(\theta, \mathbf{h})$ where $\mathbf{h} = (h_1, h_2, \cdots, h_T)$. MCMC procedures can then be developed to sample the posterior density $f(\theta, \mathbf{h}|\mathbf{X})$ without dealing with $f(\mathbf{X}|\theta)$.

Instead of simulating directly from the posterior distribution which is often intractable, MCMC methods set up a Markov chain for each variate, and its stationary distribution is the same as the posterior density. When the Markov chain converges, the simulated values may be regarded as a sample obtained from the posterior and hence can be used as the basis for making statistical inferences.

Many MCMC algorithms have been proposed to sample the parameters and the latent volatility process in the context of lognormal SV model. Examples include the Metropolis-Hastings algorithm developed by Jacquier et al. (1994), the single-move Gibbs sampling algorithm discussed in Shephard (1993), Geweke (1994a), Shephard and Kim (1994), and multi-move or block-wise Gibbs sampling algorithms proposed in Shephard and Pitt (1997) and Kim et al. (1998). The most simplistic sampler for analyzing a lognormal SV model is the single-move algorithm which updates one variate at a time. Kim et al. (1998) showed that for the lognormal SV model the convergence of the single-move algorithm is slow due to the high posterior correlations among components of $\mathbf{h}$. Kim et al. (1998) developed several multi-move algorithms by approximating the log-chi square distribution with a discrete mixture of normals (i.e.

[7] Although we do not consider the SV model with the leverage effect in this paper, we expect our algorithm can be modified to estimate ρ if one adopts the representation of Meyer and Yu (2000).

the so-called *offset mixture of normals* approximation) to facilitate a joint draw of the vector $\mathbf{h}$.

The MCMC algorithm developed in this paper falls somewhere between the single-move and multi-move algorithms. It is different from the single-move algorithm in that ϕ and δ are sampled simultaneously according to the Metropolis-Hastings rule. In terms of sampling h_t, our algorithm involves a single-move procedure and may encounter a slow convergence. However, we calculate the partial posterior of h_t when updating components of $\mathbf{h}$ sequentially. As a consequence this procedure enables us to improve the simulation efficiency over the single-move algorithm. Moreover, the N-SV model has no obvious offset mixture of normals representation and this precludes a straightforward generalization of the multi-move algorithms of Kim et al. (1998). Finally, our method does not use an approximation.

To develop our sampling algorithm, we assume the priors of model parameters are respectively, $\sigma^2 \sim IG(p/2, S_\sigma/2)$, $(\phi+1)/2 \sim \text{Beta}(\omega,\gamma)$ and $\delta \sim N(\mu_\delta, \sigma_\delta^2)$, where IG denotes the inverse-gamma distribution. The joint posterior density for model parameters and latent volatilities is

$$\begin{aligned} f(\theta, \mathbf{h}|\mathbf{X}) &= \text{prior}(\theta) \times p(h_1|\theta) \times \prod_{t=2}^{T} p(h_t|h_{t-1},\theta) \times \prod_{t=1}^{T} p(X_t|h_t,\theta) \\ &\propto (1+\phi)^{\omega-0.5}(1-\phi)^{\gamma-0.5} \exp\left\{-\frac{(\delta-\mu_\delta)^2}{2\sigma_\delta^2}\right\} \qquad (3.16) \\ &\times \left[\prod_{t=1}^{T} g(h_t,\delta)^{-1/2}\right] \exp\left\{-\sum_{t=1}^{T} \frac{X_t^2}{2g(h_t,\delta)}\right\} \left[\frac{1}{\sigma^2}\right]^{\frac{T+p}{2}+1} \\ &\times \exp\left\{-\frac{(1-\phi^2)(h_1-\mu)^2 + \sum_{t=2}^{T}\left[(h_t-\mu)-\phi(h_{t-1}-\mu)\right]^2 + S_\sigma}{2\sigma^2}\right\}, \end{aligned}$$

where p, S_σ, ω, γ, μ_δ and σ_δ^2 are all hyperparameters to be defined by users. After integrating out σ^2 from the joint posterior, we obtain the logarithm of the marginal posterior of $(\phi, \delta, \mathbf{h})$,

$$\begin{aligned}\ln f(\phi, \delta, \mathbf{h}|\mathbf{X}) \;\; \propto \;\; & (\omega - 0.5)\ln(1+\phi) + (\gamma - 0.5)\ln(1-\phi) \\ & -\frac{(\delta - \mu_\delta)^2}{2\sigma_\delta^2} - \frac{1}{2}\sum_{t=1}^{T}\ln\; g(h_t, \delta) - \sum_{t=1}^{T}\frac{X_t^2}{2g(h_t, \delta)} \\ & -\frac{T+p}{2}\ln\left\{\frac{(1-\phi^2)(h_1-\mu)^2 + \sum_{t=2}^{T}\left[(h_t-\mu) - \phi(h_{t-1}-\mu)\right]^2 + S_\sigma}{2}\right\}.\end{aligned} \tag{3.17}$$

The Gibbs sampling algorithm can then be used to sample ϕ, δ and $\mathbf{h}$. Given the posterior samples of ϕ, δ and $\mathbf{h}$ obtained from the marginal posterior $\ln f(\phi, \delta, \mathbf{h}|\mathbf{X})$, the parameter σ^2 can be sampled directly from,

$$\sigma^2 \sim IG\left(\frac{T+p}{2}, \frac{1}{2}\left[(1-\phi^2)(h_1-\mu)^2 + \sum_{t=2}^{T}\left[(h_t-\mu) - \phi(h_{t-1}-\mu)\right]^2 + S_\sigma\right]\right). \tag{3.18}$$

Kim et al. (1998) showed that the marginal posterior of μ is $N(\hat{\mu}^*, \hat{\sigma}_\mu^2)$ with

$$\begin{cases} \hat{\mu}^* & = \hat{\sigma}_\mu^2\{\frac{1-\phi^2}{\sigma^2}h_1 + \frac{1-\phi}{\sigma^2}\sum_{t=2}^{T}(h_t - \phi h_{t-1})\} \\ \hat{\sigma}_\mu^2 & = \sigma^2\left\{(T-1)(1-\phi^2) + (1-\phi^2)\right\}^{-1} \end{cases}. \tag{3.19}$$

Given the posterior samples of ϕ, δ, σ and $\mathbf{h}$, the parameter μ can be sampled directly from this marginal posterior.[8] Hence our sampling algorithm may be summarized as follows:

1. Initialize θ and $\mathbf{h}$;
2. Sample ϕ and δ from (3.17) given all the other parameters and $\mathbf{h}$;
3. Sample the components of $\mathbf{h}$ sequentially based on (3.16) given θ;
4. Sample σ^2 from (3.18) given all the other parameters and $\mathbf{h}$;
5. Sample μ from (3.19) given σ^2, ϕ and $\mathbf{h}$;

[8]Since μ can be sampled independently, the prior of μ has no effect on sampling the other parameters and the vector of latent volatilities. That is why we do not put the prior of μ into the joint posterior. When the prior of μ is required for further inferences, for instance for calculating the marginal likelihood, it can be assumed to be Gaussian with constant mean and variance.

6. Goto 2 and iterate for $N_0 + N$ times;

where N_0 is the number of iterations in the burn-in period and N is the simulation sample size.

Two important points should be noted. First, ϕ and δ are sampled simultaneously according to the Metropolis-Hastings rule, rather than a single-move procedure.[9] Second, when updating h_t ($t = 1, 2, \cdots, T$) sequentially in Step 3, we only calculate the partial posterior of h_t which is the product of relevant terms containing h_t in (3.16). For instance when $\delta \neq 0$, the partial log-posterior of h_t is (ignoring the end conditions to save space)

$$\begin{aligned} \ln\, p(h_t|\theta) &\propto -\frac{1}{2\delta}\log(1+\delta h_t) - \frac{1}{2}X_t^2(1+\delta h_t)^{-1/\delta} \\ &\quad -\frac{1}{2\sigma^2}\left[(h_t-\mu)-\phi(h_{t-1}-\mu)\right]^2 - \frac{1}{2\sigma^2}\left[(h_{t+1}-\mu)-\phi(h_t-\mu)\right]^2, \end{aligned}$$

and when $\delta = 0$, the partial log-posterior of h_t becomes

$$\begin{aligned} \ln\, p(h_t|\theta) &\propto -\frac{1}{2}h_t - \frac{1}{2}X_t^2\exp(-h_t) \\ &\quad -\frac{1}{2\sigma^2}\left[(h_t-\mu)-\phi(h_{t-1}-\mu)\right]^2 - \frac{1}{2\sigma^2}\left[(h_{t+1}-\mu)-\phi(h_t-\mu)\right]^2. \end{aligned}$$

In such a way to update h_t, the computational cost is greatly reduced.

As in Meyer and Yu (2000) we use the convergence checking criteria available in the CODA software to check whether convergence has been achieved. All the results we report in this paper are based on samples which have passed the Heidelberger and Welch convergence test for all parameters.

To measure the simulation inefficiency, we use the integrated autocorrelation time, IACT (Sokal (1996)), which is also referred to as the *inefficiency factor* by Kim et al. (1998). Following Meyer and Yu (2000), we calculate IACT of a parameter (say z) using

$$\text{IACT} = \frac{\text{var}(\bar{z}_{MC}) \times N}{\text{var}(z)},$$

where the square root of $\text{var}(\bar{z}_{MC})$ is the Monte Carlo standard error and $\text{var}(z)$ is the variance of the posterior distribution. To estimate $\text{var}(z)$, we use the empirical variance

[9]See Chib and Greenberg (1995) for detailed discussion on the Metropolis-Hastings algorithms. When updating ϕ and δ, the random numbers are generated from the proposal Gaussian density on an elliptical contour. This strategy may increase the sampling efficiency.

from an MCMC output. To estimate the square root of $\text{var}(\bar{z}_{MC})$ we use the estimate suggested in Geweke (1992) based on estimating the spectral density.

3.3 Volatility Estimate, Likelihood Evaluation, and Likelihood Ratio Test

Since MCMC methods provide samples from the joint posterior distribution of all the parameters (including both model parameters and latent volatilities), a natural way for estimating volatility is to integrate out the model parameters from the posterior. This is a Bayesian approach and has been suggested in Jacquier et al. (1994). Alternatively, one can make the use of the so-called *particle filter* techniques, a class of simulation-based methods developed in recent statistics literature for filtering nonlinear non-Gaussian state space models. Important contributions in this area of research include Gordon, Salmond and Smith (1993), Kitagawa (1996), and Pitt and Shephard (1999). As a byproduct of filtering, one can do diagnostic checking to look for some suggestion of what is wrong with the model, and to evaluate the likelihood function of the model at the posterior mean.

In the context of SV models, Kim et al. (1998) explain how to use the method developed by Pitt and Shephard (1999) while Berg et al. (2002) discuss the method proposed by Kitagawa (1996). In this paper we employ Kitagawa's filtering algorithm using $50,000$ particle points. However, we should point out that Kitagawa's algorithm is not necessarily the most efficient. Perhaps a more efficient algorithm for filtering a SV model is in, for example, Pitt and Shephard (1999).

Once likelihood is evaluated at the posterior mean, one can make statistical comparisons of the proposed N-SV model and any standard SV model. Since the N-SV model nests all standard SV models, a simple test statistic is the likelihood ratio test defined by

$$LR = 2\{\ln f(x|M_1, \hat{\theta}) - \ln f(x|M_0, \hat{\theta})\},$$

where M_1 and M_0 denote the N-SV model and a standard SV model respectively. For non-nested model comparison, one can use the non-nested likelihood ratio test developed by Atkinson (1986) for classical inferences, or for Bayesian inferences use the Bayesian factor (Chib (1995)) if the prior is proper or deviance information criterion (Berg et al. (2002)) regardless of properties of the prior. We focus on the likelihood ratio test in this paper.

3.4 Simulation Studies

To check the reliability of the proposed MCMC algorithm for estimation of N-SV models and for model comparison, we apply our algorithm to a generated dataset. We generate one data series of 2000 observations from the N-SV model using the following parameter values: $\mu = -0.2$, $\sigma = 0.2$, $\phi = 0.95$ and $\delta = 0.2$. All these parameter values are selected to be representatives of typical daily exchange rates.[10] The generated return and volatility series are plotted in the first two panels in Figure 2.

In both the simulation and empirical studies (in Section 4), we estimate the N-SV model using the proposed MCMC algorithm. For comparison purposes, we also estimate the lognormal SV model and for this we employ the all purpose Bayesian software package BUGS based on the single-move Gibbs sampler as described in Meyer and Yu (2000) for ease of implementation. In all cases we choose a burn-in period of 50,000 iterations and a follow-up period of 500,000, and store every 50th iteration. The MCMC sampler is initialized by setting $\phi = 0.95, \sigma^2 = 0.02$, and $\mu = 0$ for the lognormal SV model and arbitrarily initialized for the N-SV model. The same prior distributions are used for the common parameters in both models.[11] The hyperparameters are, respectively, $p = 10.0$, $\omega = 20.0$, $\gamma = 1.5$, $S_\sigma = 0.1$, $\mu_\delta = 0.2$ and $\sigma_\delta^2 = 0.25$.

In Table 2 we summarize the results from estimation and model comparison, including the posterior means, standard deviations, Monte Carlo standard errors (MC SE), IACT's for all the parameters, the likelihood values for both models, and the likelihood ratio statistic and associated p value for the null hypothesis of the lognormal SV model against the N-SV model. For the N-SV model we also report the 90% Bayesian confidence intervals for all the parameters.

[10] See the empirical results below and Shephard and Pitt (1997) on parameter settings for simulation purposes.

[11] The only exception is for μ. In the lognormal SV model we choose an informative but reasonably flat prior distribution for μ (i.e. a normal distribution with mean 0 and variance 25) while in the N-SV model we use a diffuse prior for the reason argued above.

First, it can be seen that the proposed MCMC procedure can estimate very precisely all the parameters in the N-SV model, including the key parameter, δ. Second, the 90% Bayesian confidence interval of δ includes the true value and excludes 0, 0.5 and 1. Consequently, we are able to reject all standard SV models as we wish. Moreover, the likelihood ratio statistic favors the true specification and suggests some evidence against the lognormal model. Third, comparison of IACT's across two models shows that the inefficiency factors in the N-SV model are substantially smaller and suggests that better mixing is achieved in the N-SV model.

To understand the implications of the mis-specification on volatility estimates, we obtain two filtered volatility estimates and plot the difference between the true volatility and two estimated volatility series in panels 3-4 of Figure 2. From these two panels, the two estimated volatility series are almost indistinguishable. To highlight the differences between the models, we plot the differences between the two estimated volatility series in the last panel of Figure 2. It can be seen that the estimated volatilities from both models are very close to each other during times of normal volatility. During times of high volatility, the differences are larger. Closer inspection shows that the two estimated volatility sequences have a similar sample mean (0.9514 versus 0.9625) but the sample variance of estimated volatilities is considerably smaller for the N-SV model (0.1940 versus 0.2303), indicating that while two models imply a similar level of long term variance the N-SV model tends to generate a smoother volatility series.

4 Empirical Results for Dollar/Pound Exchange Rates

4.1 Data

SV models are often used to model the volatility of exchange rates (see for example, Melino and Turnbull (1990), Harvey et al. (1994), Mahieu and Schotman (1998), and Jacquier, Polson and Rossi (2002)). In this section we empirically test all standard SV models against the proposed models using daily dollar/pound exchange rates for the period from January 1, 1986 to December 31, 1998. The dataset is available from the H-10 Federal Reserve Statistical Release. For convergence purposes we use the meancorrected and variance-scaled returns defined by

$$X_t = \frac{Y_t}{s(Y_t)}, \text{ with } Y_t = (\ln S_t - \ln S_{t-1}) - \frac{1}{n}\sum(\ln S_t - \ln S_{t-1}),$$

where $s(Y_t)$ is the sample standard deviation of Y_t and S_t is the exchange rate at time t. The sample size is 3268. Since the lognormal SV model is the most widely used one in the literature, we also estimate it for comparison.

4.2 Empirical Results

Figure 3 displays the adjusted return series. Figure 4 plots the MCMC iterations and kernel density estimates of the marginal posterior distribution of the model parameters in the N-SV model. In Table 3 we summarize the empirical results, including the posterior means, standard deviations, Monte Carlo standard errors (MC SE), IACT's for all the parameters, the likelihood values for both models, and the likelihood ratio statistic and associated p value for the null hypothesis of the lognormal SV model against the N-SV model. For the N-SV model we also report the 90% Bayesian confidence intervals for all the parameters.

A few results emerge from Table 3 and Figure 4. First, the posterior mean of δ in the proposed N-SV model is 0.172 and the 90% Bayesian confidence interval does not include 0, or 0.5, or 1. This is the evidence of nonlinear stochastic volatility. As a consequence, one has to reject all the standard SV models used in the literature, including the lognormal, Stein-Stein, and Heston specifications. Although all the standard SV models are rejected, the posterior quantities of δ seem to suggest that the lognormal model is closer to the true specification than other SV models with either $\delta = 0.5$ or $\delta = 1$. Second, the posterior mean of ϕ (0.9676) is close to 1 in the lognormal model and suggestive of high persistency of volatility. In the proposed N-SV model, it remains at a similar level. In fact all the estimated parameters have similar magnitudes and similar standard deviations across both models. Third, the likelihood ratio statistic and the associated p value suggest that the lognormal model is rejected at the 10 percent level. Fourth, as in the simulation study, IACT's are large for most parameters and indicate a slow convergence. However, all the chains mix well and the mixing is not affected in the N-SV model. On the contrary, the inefficiency factors in the N-SV model are considerably smaller than those in the lognormal model. Fifth, compared with other parameters, δ appears more difficult to estimate and has the largest value of standard deviation. Finally, according to our interpretation of δ, for the dollar/pound exchange rate on average new information arrives at the market about 6 times per day.

To provide diagnostic checks for the observed series and two SV models, we follow Kim et al. (1998, Section 4.2) and compute the forecast uniforms from one-step-ahead forecasts for both models. Figure 5 gives the QQ-plot of the normalized innovations obtained from the lognormal model and N-SV model respectively. The plot suggests that there are more outliers in the normalized innovations that the lognormal SV cannot explain than the N-SV model. Similar to Kim et al., we find that these ouliers correspond to small values of $|X_t|$ which are the inliers of returns. Consequently, we can conclude that the N-SV model explains the inlier behavior better than the lognormal SV model.

As argued in Section 2, a byproduct of the new volatility modelling is that the marginal distributions of volatility is obtained. The marginal distributions of volatility implied from the estimated lognormal and N-SV models are plotted in Figure 6, where the solid line is for the lognormal SV model and hence is the density function of a lognormal distribution. It can be seen that these two distributions are not very close to each other. For example, it appears that very little daily movement on the market is more possible in the N-SV model than in the lognormal SV model. The finding is quite interesting and may have important implications for risk management.

As a final comparison of the performances of the two SV models, we obtain two filtered estimates of volatility and plot them in the second and third panels in Figure 7. For comparison purposes, we also plot the absolute value of returns in the first panel. The two filtered volatility series are almost indistinguishable. To highlight the differences between the models we plot the difference between the two estimated volatility series in the last panel. It can be seen that the estimated volatilities from both models are very close to each other during times of normal volatility. During times of high volatility, the differences are larger. Similar to what we have found in the simulation study, we find that the two estimated volatility sequences have a similar sample mean (0.995 versus 1.004) but the sample variance of estimated volatilities is considerably smaller for the N-SV model (0.3297 versus 0.3782), indicating that while two models imply a similar level of long term variance the N-SV model tends to generate a smoother volatility series. As we will see below, this property has important implications on option pricing.

5 Implications on Option Pricing

Probably the most important application of the SV model is the pricing of options. Under a set of assumptions, Hull and White (1987) show that the value of a European call option on stocks based on a general specification of stochastic volatility is the Black-Scholes price integrated over the distribution of the mean volatility. Using the characteristic function approach, Heston (1993) derives a closed form solution for a European call option based on a square-root specification of volatility. For most other SV models, including our newly proposed N-SV model, option prices have no closed form solution and hence have to be approximated. A flexible way for approximating option prices is via Monte Carlo simulations. For example, Hull and White (1987) design an efficient procedure of carrying out the Monte Carlo simulation to calculate a European call option on stocks.

To examine the implication of our N-SV models on option pricing, we modify Hull and White's procedure to price currency options by taking into account the difference between stock and currency options (which is the currency options pay a "dividend" rate equal to the foreign interest rate; see for example Hull (1996, Ch12)).

Let C be the value of a European call option on a currency with maturity τ (measured in number of days), strike price X, current volatility σ_0^2, current exchange rate S_0, and the difference between the domestic and the foreign interest rates $r_d - r_f$. Under the same set of assumptions in Hull and White (1987), it can be shown that

$$C = e^{-\tau r_d} \int_0^\infty \mathrm{BS}(w_\tau) pdf(w_\tau | h_0) dw_\tau, \tag{5.20}$$

where w_τ^2 is given by

$$w_\tau^2 = \int_0^\tau g(h_s, \delta) ds, \tag{5.21}$$

and $\mathrm{BS}(w_\tau)$ is the Black-Scholes price for a currency option

$$\mathrm{BS}(w_\tau) = F_0 \mathrm{N}(d_1) - X\mathrm{N}(d_2), \tag{5.22}$$

in which $F_0 = S_0 e^{(r_d - r_f)\tau}$ is the forward exchange rate applying to time τ, d_1 and d_2 are given, respectively, by

$$d_1 = \frac{\ln(F_0/X) + w_\tau^2}{w_\tau}, \tag{5.23}$$

$$d_2 = d_1 - w_\tau. \tag{5.24}$$

In discrete time we have to approximate w_τ^2. In this paper we follow the suggestion of Amin and Ng (1993):

$$w_\tau^2 \approx \sum_{t=1}^{n} g(h_i, \delta), \tag{5.25}$$

where n is the number of discrete time periods until maturity of the option. In this paper, we choose the unit discrete time period to be one trading day and hence n $(= \tau)$ is the number of trading days before the maturity.

The Monte Carlo algorithm for calculating the value of a European call option on a currency may be summarized as follows:

1. Obtain the initial value of h_0 based on the initial value of σ_0^2;
2. Draw independent standard normal variates ν_i for $1 \leq i \leq n$;
3. Generate h_i according to

$$h_i = \mu + \phi(h_{i-1} - \mu) + \sigma\nu_i, \text{ for } i = 1, ..., n;$$

4. Calculate w_τ^2 using equation (5.25);
5. Calculate $\mathrm{BS}(w_\tau)$ using equation (5.22) and call it p_1;
6. Repeat Steps 3-5 using $\{-\nu_i\}$ and define the value of $\mathrm{BS}(w_\tau)$ by p_2;
7. Calculate the average value of p_1 and p_2 and call it y;
8. Repeat Steps 2-7 for K times and hence we should have a sequence of y's;
9. Calculate the mean of y's and this is the estimate of the option price.

Our algorithm is related to the one suggested by Mahieu and Schotman (1998), but there are several important differences. The first difference is we use an antithetic method in Step 6 to reduce the variance of simulation errors. Secondly, our algorithm can price not only at-the-money options but also in-the-money and out-of-the-money options while Mahieu and Schotman only price at-the-money options. The third difference is we can price options based on the N-SV models. Finally, we use a much larger value of K (10,000 as opposed to 500) to ensure that the approximation errors in calculating equation (5.20) are very small.

The algorithm is then applied to price a half-year call option based on the lognormal and N-SV models with the estimated parameter values in Table 3 imposed.[12] In both models, we choose $n = 126$, $S_0 = 1.5$, $r_d = 0$, $r_f = 0$, $K = 10,000$, $\sigma_0 = 0.006349$,[13] and S_0/X takes each of the following values, $0.75, 0.8, 0.85, 0.9, 0.95, 1, 1.05, 1.1, 1.15, 1.2, 1.25$. Table 4 compares the option prices and percentage differences between the prices based on the two estimated SV models.

The main conclusion we draw from the table is that the lognormal SV model tends to overprice the options. In fact the N-SV option price is always no bigger than the lognormal option prices. This finding is not surprising because we have found that while both models have a similar value of long term variance the N-SV model tends to generate a smoother volatility series. Prices of all the out-of-money options based on the N-SV model are systematically lower than those based on the lognormal model and the deep-out-of-the-money options show the largest percentage of discrepancies. The differences in the percentage term are much smaller for in-the-money options and eventually disappear when the in-the-money option goes very deep. Since near out-of-money options where the strike price is within about 10% of the spot price are traded very frequently over the counter and on exchanges, our results have important practical implications.

[12]Since the parameter estimates reported in Table 3 are based on the scaled data, for the purpose of pricing options, we have to scale the data back by multiplying the mean equation by the sample standard error of raw data which equals 0.006321 for the dollar/pound exchange rate.

[13]This initial value of standard error is very close to the sample standard error of the dollar/pound exchange rate and corresponds to a square root of volatility of 160% per year.

6 Empirical Results for Other Exchange Rates

In this section we apply the N-SV models to daily exchange rates of four other major currencies against the US dollar for the same sampling period as for the dollar/pound rate. The currencies are Canadian dollar (CD), French franc (FF), German mark (GM), and Japanese yen (JY). These data are also available from the H-10 Federal Reserve Statistical Release. As for the dollar/pound exchange rate, we use the demeaned and variance-scaled return. The sample size is 3268 in all cases.

Figure 8 displays all the other currencies. In Table 5 we summarize the empirical results, including the posterior means, standard deviations, 90% Bayesian confidence intervals for all the parameters, the likelihood values for both the N-SV and lognormal models, and the likelihood ratio statistic and associated p value for the null hypothesis of the lognormal SV model against the N-SV model. The number in parentheses is the standard deviation while the number in brackets represents the 90% Bayesian confidence interval.

A few results emerge from Table 5. First, and most importantly, in all cases the posterior mean of δ is very close to zero and the 90% Bayesian confidence interval contains zero, indicating suitability of the lognormal SV model. The same conclusion is drawn from the LR test. However, in no case the 90% Bayesian confidence interval contains 0.5 or 1 and hence rejects all the other standard SV models, including the Stein-Stein and Heston specifications. Second, the estimated volatility process for all the currencies except for JY is highly persistent in both models. Also, all the estimated parameters have very similar magnitudes and similar standard deviations across both models. Although not reported, it appears that all the chains mix well, indicating the empirical results are reliable. Finally, since estimated δ is so close to 0, the marginal distribution of volatility should be well approximated by the lognormal distribution.

7 Conclusions and Extensions

In this paper a class of nonlinear SV models has been proposed. The new class is based on the Box-Cox power transformation and encompasses all standard parametric SV models which have appeared in the literature, including the well known lognormal SV model. As a result, different SV specifications in the literature can be easily tested. The MCMC approach is developed to provide a likelihood-based inference for the analysis of proposed models. Simulation studies confirm that the proposed MCMC algorithm works well for the new models. Empirical applications are performed first using daily dollar/pound exchange rate series. Empirical results show that all the standard SV models are rejected and hence suggest evidence of nonlinear stochastic volatility. Furthermore, model diagnostics indicate that, without sacrificing the overall goodness-of-fit the nonlinear SV model improves the fit to the data when the market has little movement. Moreover, this nonlinearity has important implications for pricing currency options. In particular the lognormal models tend to overprice out-of-money options. The deeper the out-of-money options, the larger the percentage bias. For all the other four major exchange rate series considered, the only standard "classical" SV model which cannot be rejected is the lognormal model. As a result, daily exchange volatility is well described by the lognormal distribution as its marginal distribution, consistent with the results found in recent literature (Andersen et al. (2001)).

There are some other possible extensions to our work. One possibility is to use the suggested methodology to analyze stock data. However, since stock data often display a strong volatility feedback feature, one has to incorporate this leverage effect into the nonlinear SV model. Other interesting extensions would be to incorporate jumps and long memory volatility into the model; see for example Duffie, Pan and Singleton (2000), Andersen, Benzoni and Lund (2001), Eraker et al. (2001), Breidt, Crato and De Lima (1998) and Robinson (2001). Finally, it would be interesting to evaluate the out-of-sample forecasting performances of the nonlinear SV models relative to standard SV models.

References

[1] Amin, K.I and V.K. Ng (1993). Jump diffusion option valuation in discrete time. *Journal of Finance* 48, 881-909.

[2] Andersen, T. (1994). Stochastic autoregressive volatility: A framework for volatility modelling. *Mathematical Finance* 4, 75–102.

[3] Andersen, T., L. Benzoni, and J. Lund (2001). An empirical investigation of continuous-time equity returns models. *Journal of Finance*, forthcoming.

[4] Andersen, T., T. Bollerslev, F.X. Diebold, and P. Labys (2001). The distribution of realized exchange rate volatility. *Journal of the American Statistical Asscociation* 96, 42–55.

[5] Andersen, T., H. Chung, and B. Sorensen (1999). Efficient method of moments estimation of a stochastic volatility model: A Monte Carlo study. *Journal of Econometrics* 91, 61–87.

[6] Andersen, T. and B. Sorensen (1996). GMM estimation of a stochastic volatility model: A Monte Carlo study. *Journal of Business and Economic Statistics* 14, 329–352.

[7] Atkinson, A.C. (1986). Monte Carlo tests of separate families of hypothesis. Unpublished Paper, Imperial College, London.

[8] Barndorff-Nielsen, N. and N. Shephard (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.

[9] Berg, A., R. Meyer, and J. Yu (2002). Deviance Information Criterion for Comparing Stochastic Volatility Models. Working Paper, Department of Economics, University of Auckland.

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

[11] Box, G.E.P. and D.R. Cox (1964). An analysis of transformations. *Journal of Royal Statistical Society B* 26, 211–243.

[12] Breidt, F.J., N. Crato, and De Lima, P. (1998). The detection and estimation of long memory in stochastic volatility. *Journal of Econometrics* 83, 325–348.

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

[14] Chib, S. (1995). Marginal likelihood from the Gibbs output. *The Journal of the American Statistical Association* 90, 1313-1321.

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

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

[17] Clark, P.K. (1973). A subordinated stochastic process model with finite variance for speculative price. *Econometrica* 41, 135–155.

[18] Danielsson, J. (1994). Stochastic volatility in asset prices: Estimation with simulated maximum likelihood. *Journal of Econometrics* 64, 375–400.

[19] Danielsson, J. and J.F. Richard (1993). Accelerated Gaussian importance sampler with application to dynamic latent variable models. *Journal of Applied Econometrics* 8, s153–s173.

[20] Duffie, D., J. Pan and K.J. Singleton (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica* 68, 1343–1376.

[21] Duffie, D. and K.J. Singleton (1993). Simulated moments estimation of Markov models of asset prices. *Econometrica* 61, 929–952.

[22] Durbin, J. and S.J. Koopman (2000). Time series analysis of non-Gaussian observations based on state space models from both classical and Bayesian perspectives (with discussion). *Journal of the Royal Statistical Society Series B* 62, 3–56.

[23] Durham, Garland and A.R. Gallant (2002). Numerical techniques for maximum likelihood estimation of continuous-time diffusion processes. *Journal of Business and Economic Statistics* 20, 297–316.

[24] Engle, R.F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* 50, 987–1007.

[25] Eraker, B. (2001). MCMC analysis of diffusion models with application to finance *Journal of Business and Economic Statistics* 19, 177-191.

[26] Eraker, B., Johannes, M. and N. Polson (2001) The impact of jumps in volatility and returns. *Journal of Finance*, forthcoming.

[27] Gallant, A.R. and G. Tauchen (1996). Which moments to match? *Econometric Theory* 12, 657-681.

[28] Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In Bernardo, J.M., Beger, J.O., David, A.P., and Smith, A.F.M. (eds), *Bayesian Statistics* 4, pp. 169-93. Oxford: Oxford University Press.

[29] Geweke, J. (1994a). Comments on 'Bayesian analysis of stochastic volatility models' by Jacquier, E., N.G. Polson and P.E. Rossi. *Journal of Business and Economic Statistics* 11, 397–399.

[30] Geweke, J. (1994b). Bayesian comparison of econometric models. Working Paper, Federal Reserve Bank of Minneapolis, Minnesota.

[31] Gordon, N.J., D.J. Salmond, and A.E.M. Smith (1993). A novel approach to nonlinear and non-Gaussian Bayesian state estimation. *IEEE-Proceedings F* 140, 107–133.

[32] Gouriéroux, C., Monfort, A. and E. Renault (1993). Indirect inference. *Journal of Applied Econometrics* 8, 85–118.

[33] Hansen, L.P. and J.A. Scheinkman (1995). Back to the future: Generating moment implications for continuous-time Markov processes. *Econometrica* 63, 767–804.

[34] Harvey, A.C., E. Ruiz, and N. Shephard (1994). Multivariate stochastic variance models. *Review of Economic Studies* 61, 247–264.

[35] Hentschel, L. (1995). All in the family: Nesting symmetric and asymmetric GARCH models. *Journal Financial Economics* 39, 71-104.

[36] 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.

[37] Higgins, M.L. and A. Bera (1992). A class of nonlinear ARCH models. *International Economic Review* 33, 137–158.

[38] Hull, J. (1996). *Options, Futures, and other Derivatives.* Prentice Hall, Englewood Cliffs, NJ.

[39] Hull, J. and A. White (1987). The pricing of options on assets with stochastic volatilities. *Journal of Finance* 42, 281–300.

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

[41] Jacquier, E., N.G. Polson, and P.E. Rossi (2002). Bayesian analysis of stochastic volatility models with fat-tails and correlated errors. *Journal of Econometrics*, forthcoming.

[42] Johnson, H. and D. Shanno (1987). Option pricing when the variance is changing. *Journal of Financial and Quantitative Analysis* 22, 143–152.

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

[44] Kitagawa, G. (1996). Monte Carlo filter and smoother for Gaussian nonlinear state space models. *Journal of Computational and Graphical Statistics* 5, 1–25.

[45] Knight, J.L., S.S. Satchell and J. Yu (2002). Estimation of the stochastic volatility model by the empirical characteristic function method. *Australian and New Zealand Journal of Statistics* 44, 319–335.

[46] Mahieu, R. J. and P.C. Schotman (1998). An empirical application of stochastic volatility models. *Journal of Applied Econometrics* 13, 330–360.

[47] Meddahi, Nour (2001). An eigenfunction approach for volatility modelling. Manuscript, Department of Economics, University of Montreal.

[48] Melino, A. and S.M. Turnbull (1990). Pricing foreign currency options with stochastic volatility. *Journal of Econometrics* 45, 239–265.

[49] Meyer, R. and J. Yu (2000). BUGS for a Bayesian analysis of stochastic volatility models. *Econometrics Journal*, 3, 198–215.

[50] Nelson, D. (1990). ARCH models as diffusion approximations. *Journal of Econometrics* 45, 7–38.

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

[52] Robinson, P. (2001). The memory of stochastic volatility models . *Journal of Econometrics* 101, 195–218.

[53] Ruiz, E. (1994). Quasi-maximum likelihood estimation of stochastic volatility models. *Journal of Econometrics* 63, 289–306.

[54] Sandmann, G. and S.J. Koopman (1998). Maximum likelihood estimation of stochastic volatility models. *Journal of Econometrics* 63, 289–306.

[55] Scott, L.O. (1987). Option pricing when the variance changes randomly: Theory, estimation and an application. *Journal of Financial and Quantitative Analysis* 22, 419–439.

[56] Shephard, N. (1993). Fitting nonlinear time series models with applications to stochastic variance models. *Journal of Applied Econometrics* 8, S135–52.

[57] Shephard, N. and Kim, S. (1994). Comments on 'Bayesian analysis of stochastic volatility models' by Jacquier, E., N.G. Polson and P.E. Rossi. *Journal of Business and Economic Statistics* 11, 406–410.

[58] Shephard, N. and Pitt, M.K. (1997). Likelihood analysis of non-Gaussian measurement time series. *Biometrika* 84, 653–667.

[59] Singleton, K.J. (2001). Estimation of affine asset pricing models using the empirical characteristic function. *Journal of Econometrics* 102, 111–141.

[60] Sokal, A.D. (1996). Monte Carlo Methods in Statistical Mechanics: Foundations and New Algorithms. Lectures at the Cargese Summer School on "Functional Integration: Basics and Applications".

[61] Stein, E.M. and J.C. Stein (1991). Stock price distributions with stochastic volatility: An analytical approach. *Review of Financial Studies* 4, 727–752.

[62] Tauchen, G. and M. Pitts (1983). The price variability-volume relationship on speculative markets. *Econometrica* 51, 485–505.

[63] Taylor, S.J (1982). Financial returns modelled by the product of two stochastic processes — a study of the daily sugar prices 1961-75. In Anderson, O.D., Editor, *Time Series Analysis: Theory and Practice, 1*, 203–226. North-Holland, Amsterdam.

[64] Taylor, S.J (1986). *Modelling Financial Time Series.* Chichester, Wiley.

[65] Taylor, S.J (1994). Modelling stochastic volatility: A review and comparative study. *Mathematical Finance* 4, 183–204.

[66] Wiggins, J.B. (1987) Option values under stochastic volatility: Theory and empirical estimate. *Journal of Financial Economics* 19, 351–372.

[67] Yang, Z. (1999) Estimating a transformation and its effect on Box-Cox T-ratio. *Test* 8, 167-190.

[68] Yu, J. (2002) Forecasting volatility in the New Zealand stock market. *Applied Financial Economics* 12, 193–202.

Table 1: Alternative Stochastic Volatility Models and Parameter Relationship

	Models	δ	μ	ϕ
Wiggins (1987) Scott (1987) Chesney and Scott (1989) Taylor (1994) Jacquier, Polson and Rossi (1994) Harvey, Ruiz and Shephard (1994) Kim, Shephard and Chib (1998)	$\ln \sigma_t^2 = \mu + \phi(\ln \sigma_{t-1}^2 - \mu) + \sigma v_t$	0		
Scott (1987) Stein and Stein (1991) Andersen (1994)	$\sigma_t = \mu + \phi(\sigma_{t-1} - \mu) + \sigma v_t$	0.5		
Heston (1993)	$\sigma_t = \phi\sigma_{t-1} + \sigma v_t$	0.5	0	
Hull and White (1987) Johnson and Shanno (1987)	$\ln \sigma_t^2 = \mu + \ln \sigma_{t-1}^2 + \sigma v_t$	0		1
Andersen (1994)	$\sigma_t^2 = \mu + \phi(\sigma_{t-1}^2 - \mu) + \sigma v_t$	1		
Clark (1973)	$\ln \sigma_t^2 = \mu + \sigma v_t$	0		0
Nonlinear SV	$\frac{(\sigma_t^2)^\delta - 1}{\delta} = \mu + \phi[\frac{(\sigma_{t-1}^2)^\delta - 1}{\delta} - \mu] + \sigma v_t$			

Table 2: Results for Simulated Rate

	True Val	N-SV					Lognormal SV			
		Mean	SD	90% CI	MC SE	IACT	Mean	SD	MC SE	IACT
ϕ	0.95	.9564	.0121	(.9348, .9741)	.00019	121.9	.9598	.0120	.00050	883.7
σ	0.2	.1893	.0261	(.15, .2359)	.00048	169.1	.1924	.0269	.00138	1319.7
μ	-0.2	-.2105	.1144	(-.3968, -.0236)	.00091	31.5	-.2137	.1256	.00269	229.0
δ	0.2	.2105	.1444	(.0011, .4355)	.00250	149.9				
Loglik		-2657.346					-2658.990			
LR Stat		3.287								
p-Val		0.0698								

Table 3: Empirical Results for dollar/pound Exchange Rate

	N-SV					Lognormal SV			
	Mean	SD	90% CI	MC SE	IACT	Mean	SD	MC SE	IACT
ϕ	.9595	.0101	(.9417, .9745)	.00017	138.3	.9676	..0091	.00026	408.2
σ	.2066	.0269	(.1672, .2543)	.00050	174.9	.1873	..0268	.00090	568.1
μ	-.2244	.1044	(-.3913, -.0495)	.00087	35.0	-.2579	..1095	.00103	44.1
δ	.1716	.1203	(.0039, .3684)	.00214	189.0				
Loglik	-4369.792					-4371.606			
LR Stat	3.628								
p-Val	0.0568								

Table 4: Comparison of Call Option Prices on Currency Based on Lognormal SV and N-SV Models; Option Parameters: $\tau = 126$ Days, $S_0 = 1.5$, $r_d = 0$, $r_f = 0$, $\sigma_0 = 0.006349$ Per Day

	Lognormal SV	N- SV	Percentage
S_0/X	Option Price	Option Price	Difference
0.75	2.401e-5	1.172e-5	-104.86
0.8	1.511e-4	1.032e-4	-46.41
0.85	8.645e-4	7.231e-4	-19.55
0.9	0.00415	0.00386	-7.513
0.95	0.01548	0.01507	-2.721
1	0.04257	0.04213	-1.044
1.05	0.08701	0.08661	-0.462
1.1	0.1413	0.1410	-0.213
1.15	0.1971	0.1969	-0.102
1.2	0.2504	0.2503	-0.040
1.25	0.3001	0.3001	0.000

Note: In all cases, the parameter estimates in Table 3 are used.

Table 5: Empirical Results for Other Exchange Rates

		CD	FF	GM	JY
Log-Normal SV	ϕ	.9575 (.0094) [.9372, .9741]	.9618 (.0099) [.939, .9781]	.9617 (.0099) [.9395, .9733]	.8654 (.0278) [.8158, .9069]
	σ	.233 (.0248) [.1864, .284]	.1841 (.0244) [.1417, .2361]	.1772 (.0231) [.1383, .2285]	.4284 (.0515) [.3473, .5169]
	μ	-.3024 (.1027) [-.5016, -.0946]	-.2293 (.0918) [-.4096, -.0476]	-.2219 (.0883) [-.3956, -.0464]	-.378 (0.0654) [-.4842, -.2695]
	Loglik	-4316.61	-4434.54	-4440.67	-4330.89
N-SV	ϕ	.9579 (.0092) [.9414, .9716]	.9538 (.0106) [.9346, .9691]	.9585 (.0105) [.9395, .9733]	.8601 (.0285) [.8078, .9005]
	σ	.2363 (.0248) [.1989, .28]	.2053 (.0234) [.1715, .2472]	.1869 (.0235) [.1519, .2284]	.4380 (.0508) [.3647, .5293]
	μ	-.3308 (.1120) [-.516, -.1489]	-.2223 (.0934) [-.3769, -.0708]	-.2364 (.0945) [-.3924, -.0829]	-.3880 (0.0730) [-.5105, -.2704]
	δ	-.0486 (.0940) [-.2108, .1003]	.0441 (.1534) [-.2104, .2975]	-.0786 (.1668) [-.3514, .1988]	-.0158 (.0902) [-.1662, .1296]
	Loglik	-4316.44	-4434.52	-4440.37	-4330.66
LR Stat		0.3418	0.0524	0.6174	0.4654
p-Value		0.5588	0.8189	0.4320	0.4951

Note: The number in parentheses is the standard deviation while the number in brackets represents the 90% Bayesian confidence interval.

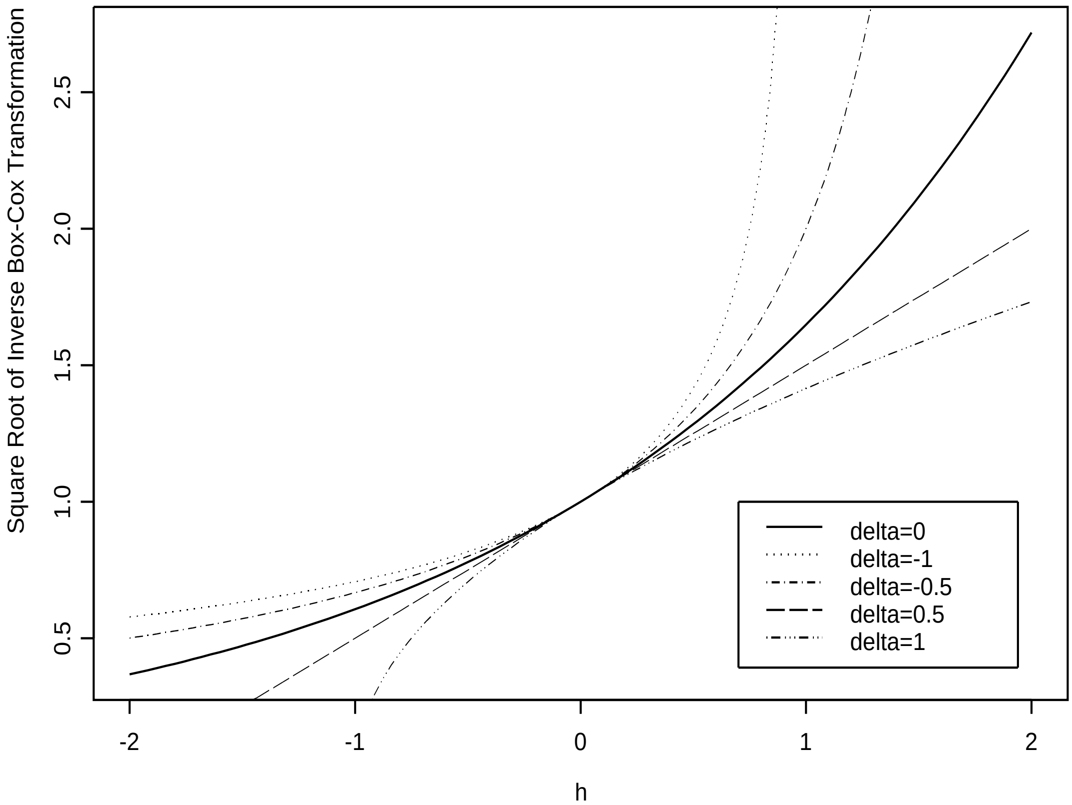

Figure 1: Inversion Box-Cox transformation for various values of δ.

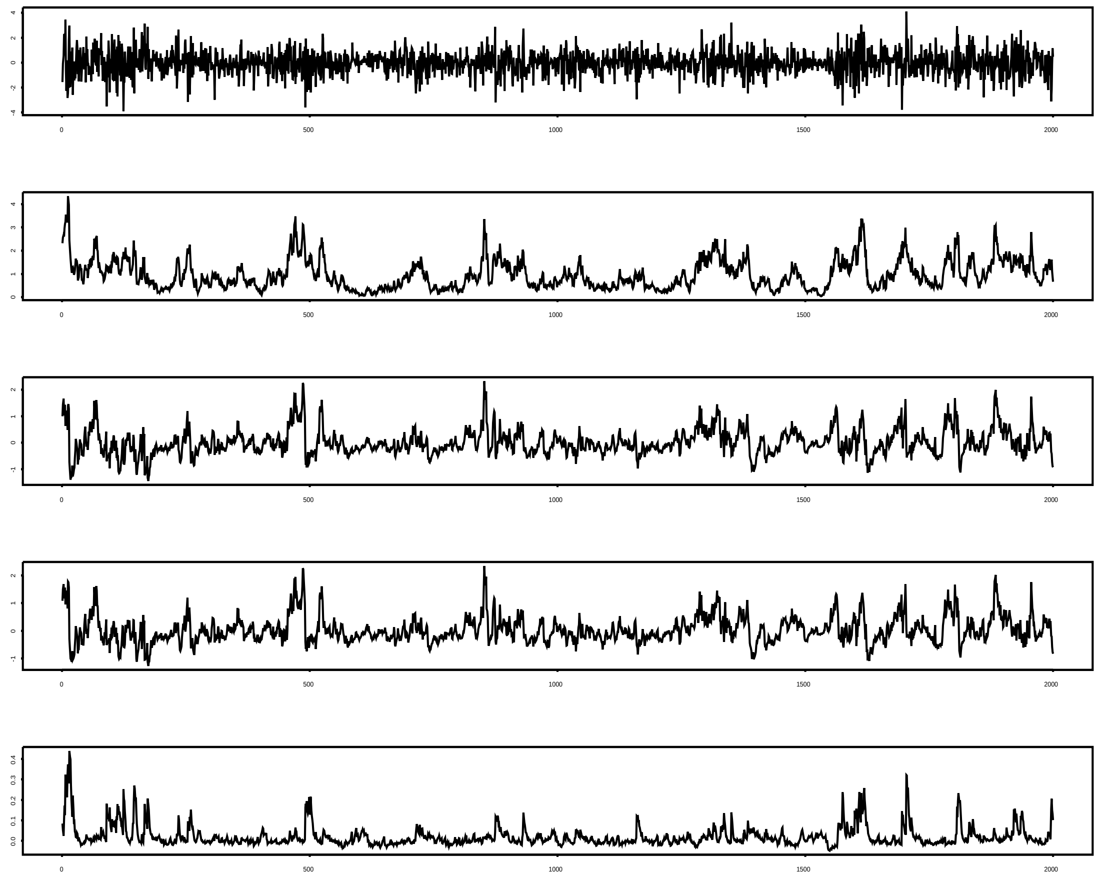

Figure 2: Simulated data and filtered volatility: The first panel is the time series plot of simulated return; the second panel is the time series plot of true volatility; the third panel is the difference between the true volatility and filtered volatility from the lognormal SV model; the fourth panel is the difference between the true volatility and filtered volatility from the N-SV model; the fifth panel is the difference between the third and fourth panels.

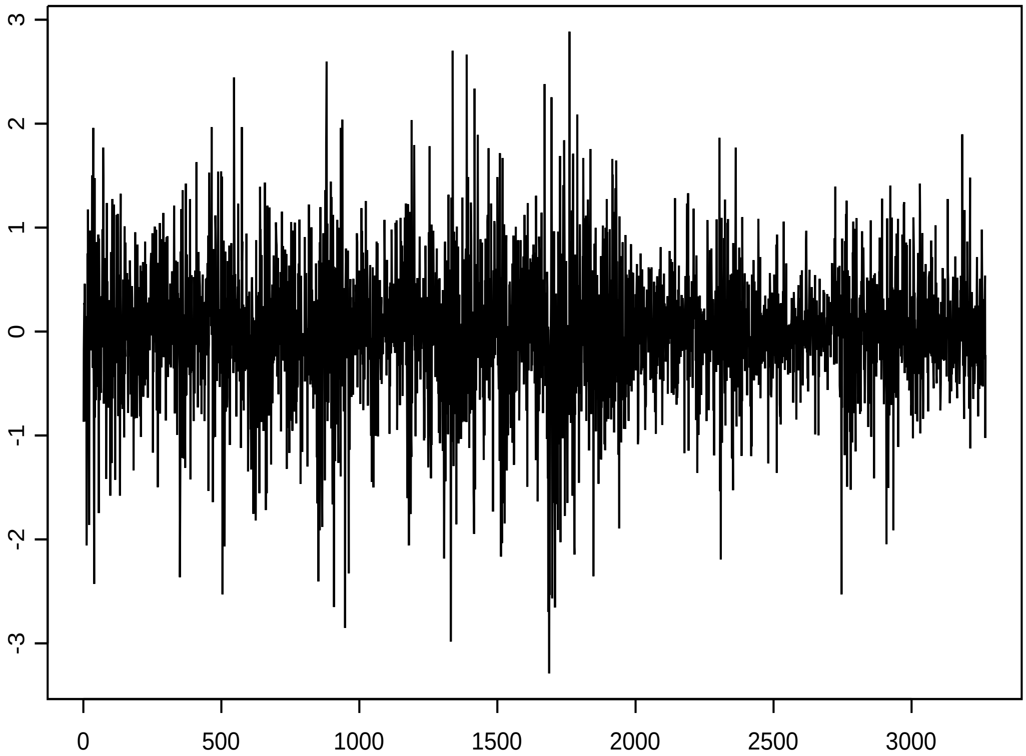

Figure 3: Time series plots for dollar/pound exchange rate return

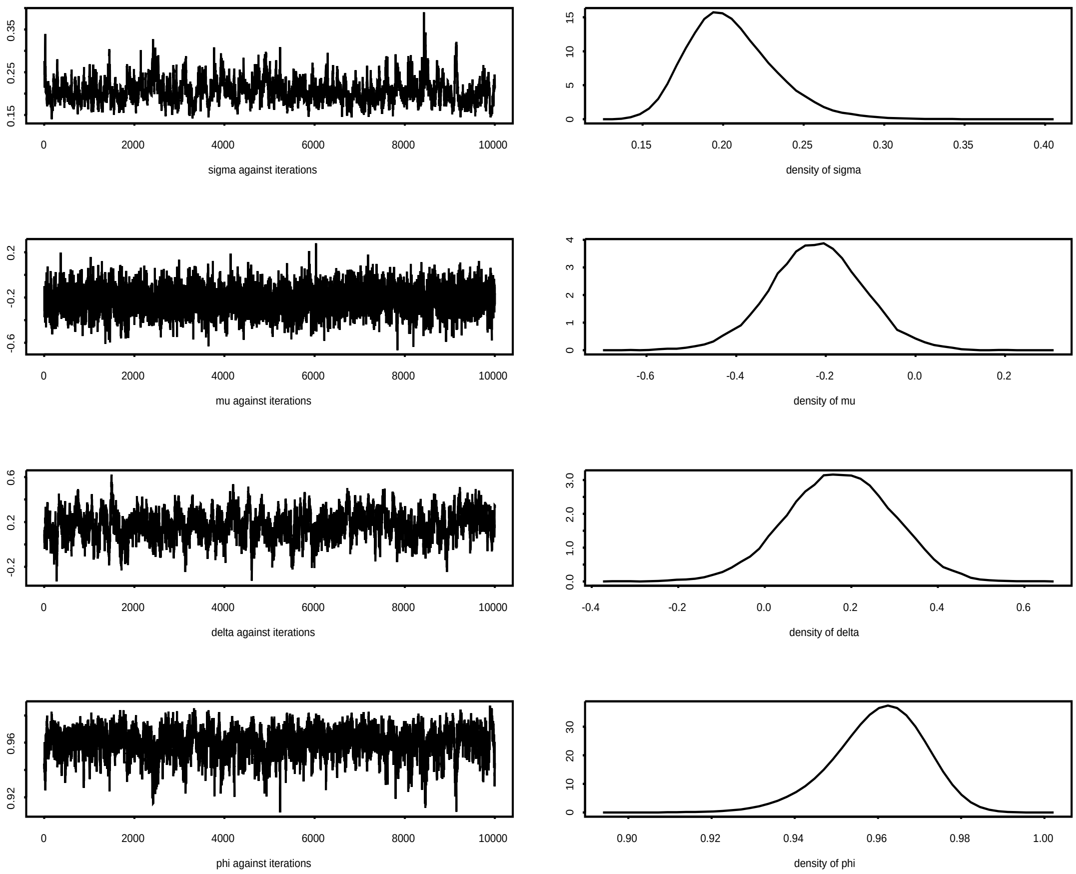

Figure 4: MCMC sample for dollar/pound exchange rate returns in the N-SV model.

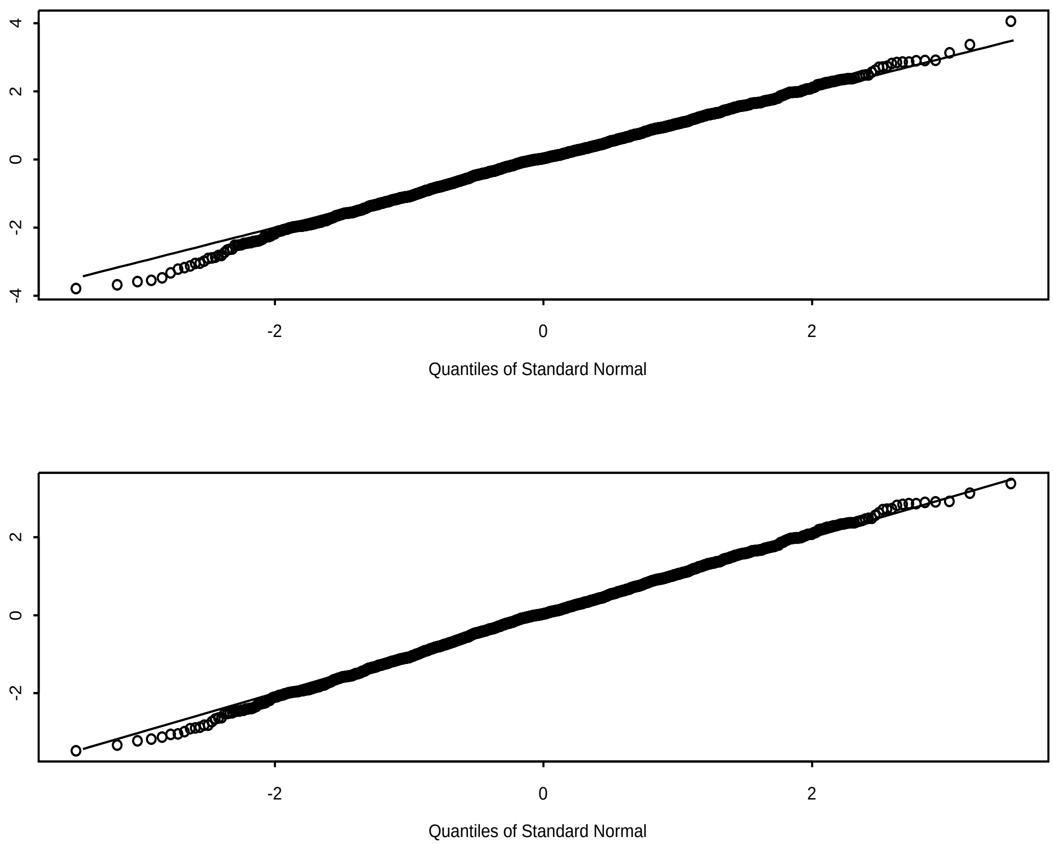

Figure 5: Diagnostic checks of two SV models for dollar/pound exchange rate returns. The first panel is the QQ-plot of the normalized residuals from the lognormal SV model; the second panel is the QQ-plot of the normalized residuals from the N-SV model.

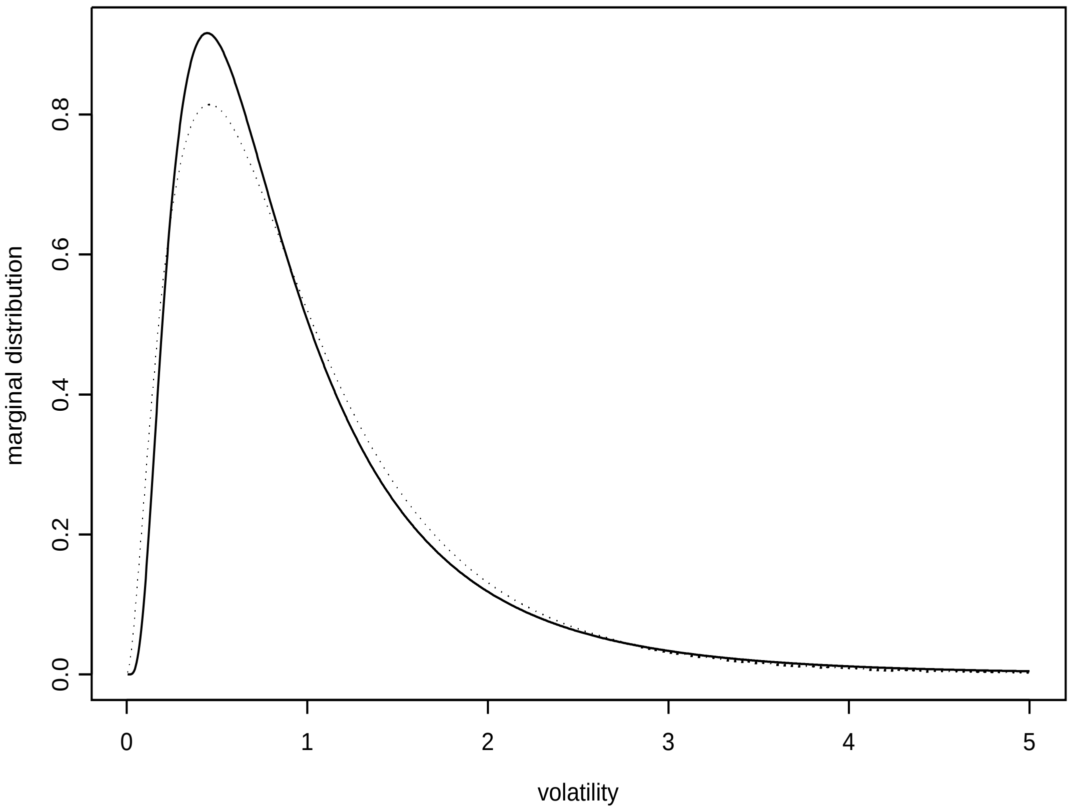

Figure 6: Marginal densities of dollar/pound exchange rate volatility implies from the lognormal SV model and the N-SV model. The solid line is for the lognormal SV model; the point line is for the N-SV model.

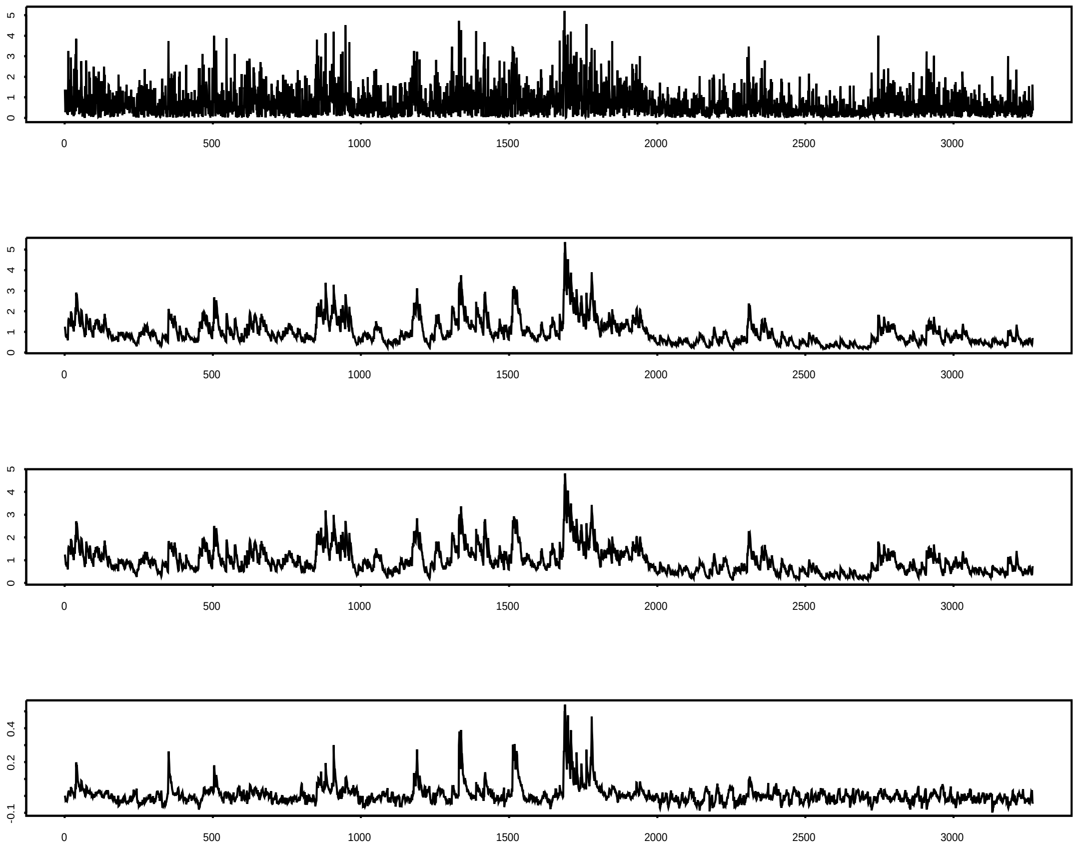

Figure 7: Filtered volatility for dollar/pound exchange rate returns. The first panel is the absolute value of the return series; the second panel is the filtered volatility from the lognormal SV model; the third panel is the filtered volatility from the N-SV model; the fourth panel is the difference between the second and the third panels.

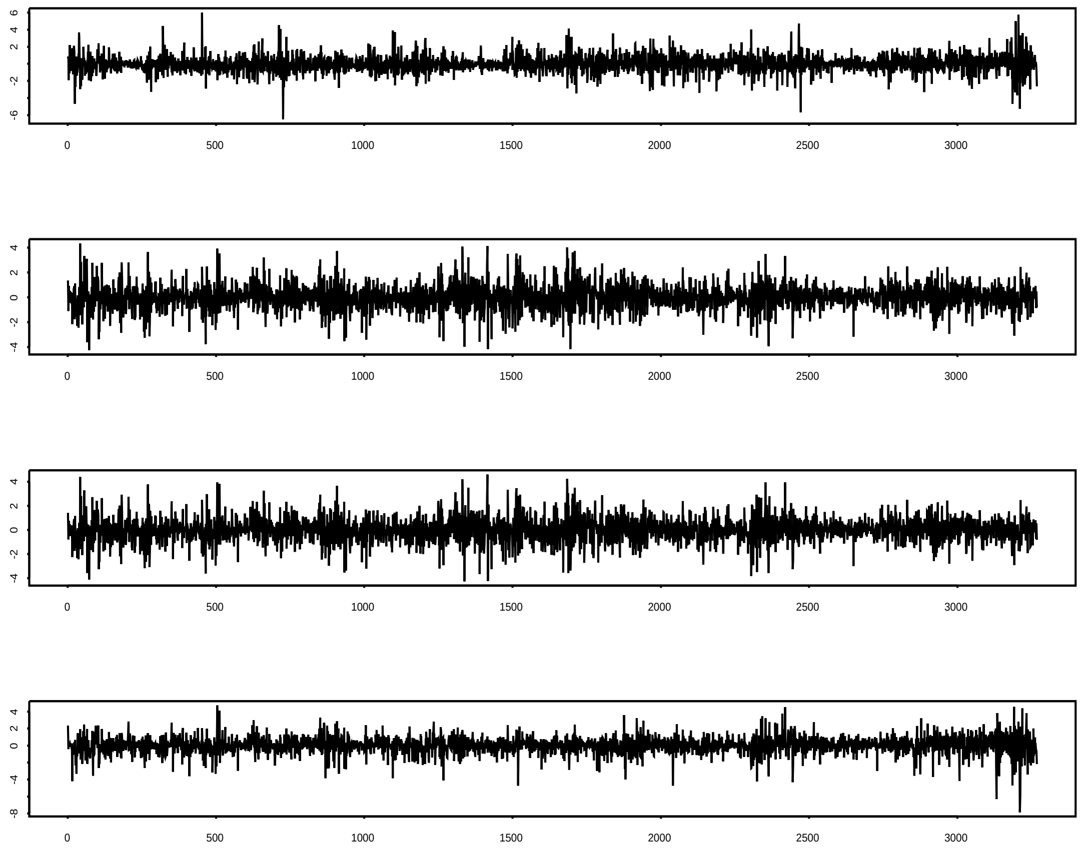

Figure 8: Time series plots for Canadian dollar, French franc, German mark, and Japanese yen/US dollar exchange rate returns.