Available online at www.sciencedirect.com
ScienceDirect
Computational Statistics & Data Analysis 51 (2006) 2218 – 2231
COMPUTATIONAL STATISTICS & DATA ANALYSIS
www.elsevier.com/locate/csda

A class of nonlinear stochastic volatility models and its implications for pricing currency options

Jun Yu[a,*], Zhenlin Yang[a], Xibin Zhang[b]

[a] School of Economics and Social Sciences, Singapore Management University, Singapore
[b] Department of Econometrics and Business Statistics, Monash University, Australia

Available online 11 September 2006

Abstract

A class of stochastic volatility (SV) models is proposed by applying the Box–Cox transformation to the volatility equation. This class of nonlinear SV (N-SV) models encompasses all standard SV models, including the well-known lognormal (LN) SV model. It allows to empirically compare and test all standard specifications in a very convenient way and provides a measure of the degree of departure from the classical models. A likelihood-based technique is developed for analyzing the model. Daily dollar/pound exchange rate data provide some evidence against LN model and strong evidence against all the other classical specifications. An efficient algorithm is proposed to study the economic importance of the proposed model on pricing currency options.

Keywords: Box–Cox transformation; GARCH; MCMC; Volatility; Option pricing

1. Introduction

Modeling the volatility of financial time series via stochastic volatility (SV) models has received a great deal of attention in the theoretical finance literature as well as in the empirical finance 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 in explaining the well-documented time varying volatility. Empirical successes of the lognormal SV model relative to GARCH-type models were documented in Kim et al. (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 perhaps the lognormal (LN) specification, which was first introduced by Taylor (1982). It has been used to price stock options in Wiggins (1987) and Scott (1987) and currency options in Chesney and Scott (1989). As it assumes that the logarithmic volatility follows an Ornstein–Uhlenbeck (OU) process, an implication for this specification is that the marginal distribution of logarithmic volatility is normal. This assumption has very important implications for financial economics and risk management.

* Corresponding author. Tel.: +65 68280858; fax: +65 68280833.
E-mail address: yujun@smu.edu.sg (J. Yu)
URL: http://www.mysmu.edu/faculty/yujun/default.htm (J. Yu).

0167-9473/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.csda.2006.08.024

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) assumed the square root of volatility follows, respectively, an OU process and a geometric Brownian motion, while Hull and White (1987) and Heston (1993) assumed a geometric Brownian motion and a square-root process, respectively, for volatility. In particular, Heston's model has received a great deal of attention in the option pricing literature, because it provides a closed-form expression for option pricing formula. 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 the standard deviation (SD) GARCH(1,1) model. Andersen (1994) introduced a class of polynomial SV models, which encompasses most discrete-time SV models in the literature. Other recent classes of SV models include those proposed by Barndorff-Nielsen and Shephard (2001), Jones (2003) and Meddahi (2001), respectively.

In spite of all these alternative specifications, there is a lack of simple procedure for selecting an appropriate functional form for a SV process. It is well-known that a GARCH process converges to a relevant SV process (Nelson, 1990). A specification test based on a GARCH family can be suggestive for an appropriate SV specification (see, for example, Duan, 1997; Hentschel, 1995). However, such a test is by no means a direct test of SV specifications. The specification of the correct SV 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 SV function can result in incorrect option prices. Second, the marginal distribution of volatility depends upon the functional form of SV.

In this paper, we propose a new class of SV models, namely, the nonlinear SV (N-SV) models. Similar to the class of Andersen (1994), this new class of SV models includes as special cases many SV models that have appeared in the literature. It overlaps with but does not encompass the class of Andersen, the classes of Jones (2003), and the class of Meddahi (2001). Unlike these alternative classes that preclude a simple comparison of competing SV models, our proposed class allows easy testing on the functional form specifications for the SV, based on a single parameter. This parameter also provides a measure of degree of departure from the classical SV models. Furthermore, with this general approach to modeling SV, one obtains the functional form of transformation, which induces marginal normality of volatility.

We have empirically tested all the standard specifications against ours using daily dollar/pound exchange rate data. We found that our empirical test presents significant evidence against all the standard SV models and favors a N-SV specification. Our empirical test of all standard SV models is, to the best of our knowledge, the first in the literature. Economic importance of this nonlinearity is also examined. For example, without sacrificing the overall goodness-of-fit, our N-SV model improves the fit to data when the market has little movement. We also found that our model implies a smoother volatility series. Moreover, the marginal distribution of volatility is different from the LN distribution. Most importantly, an application of our N-SV model to option pricing shows that the LN SV model overprices currency options, particularly out-of-the-money options, when the true model is the empirically estimated nonlinear model.

The paper is organized as follows. Section 2 introduces this class of N-SV models. In Section 3, a Markov Chain Monte Carlo (MCMC) method is developed to conduct likelihood-based analysis of the proposed class of models. Section 4 presents empirical results from fitting daily returns of the dollar/pound exchange rate. Section 5 illustrates the economic importance of the proposed models in terms of their implications for pricing currency options. Finally Section 6 concludes the paper with some discussion on possible extensions of this work.

2. A class of N-SV models

In the theoretical finance literature on option pricing, the SV model is often formulated in terms of stochastic differential equations. For instance, Wiggins (1987), and Chesney and Scott (1989) specified the following model for the asset price $P(t)$

$$\mathrm{d}P(t)/P(t) = \alpha\,\mathrm{d}t + \sigma(t)\,\mathrm{d}B_1(t), \tag{1}$$

and its associated volatility $\sigma^2(t)$,

$$\mathrm{d}\ \ln\,\sigma^2(t) = \lambda(\xi - \ln\sigma^2(t))\,\mathrm{d}t + \gamma\,\mathrm{d}B_2(t), \tag{2}$$

where $B_1(t)$ and $B_2(t)$ are two Brownian motions and $corr(\mathrm{d}B_1(t), \mathrm{d}B_2(t)) = \rho$ 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 can be obtained, for example, via the Euler–Maruyama approximation, which, after a location shift and reparameterization, leads to the LN SV model given by

$$X_t = \sigma_t e_t \tag{3}$$

and

$$\ln \sigma_t^2 = \mu + \phi(\ln \sigma_{t-1}^2 - \mu) + \sigma v_t, \tag{4}$$

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

$$X_t = \exp(\tfrac{1}{2} h_t) e_t \tag{5}$$

and

$$h_t = \mu + \phi(h_{t-1} - \mu) + \sigma v_t, \tag{6}$$

where $h_t = \ln \sigma_t^2$. See, for example, Yu (2005) for a detailed account on the leverage effect.

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

$$X_t = \sigma_t e_t, \tag{7}$$
$$h(\sigma_t^2, \delta) = \mu + \phi[h(\sigma_{t-1}^2, \delta) - \mu] + \sigma v_t, \tag{8}$$

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) given by

$$h(t, \delta) = \begin{cases} (t^{\delta} - 1)/\delta & \text{if } \delta \neq 0, \\ \ln t & \text{if } \delta = 0. \end{cases} \tag{9}$$

As the function $h(\cdot, \delta)$ is specified as a general nonlinear function, the model is thus termed in this paper the N-SV model. Several attractive features of this new class of SV models include: (i) as we will show below, it includes the LN SV model and the other popular SV models as special cases; (ii) it adds a greater flexibility to the functional form; (iii) it provides a measure on the degree of departure from a specific classical SV model; and (iv) it allows a simple test for the LN SV specification, i.e., a test of $\mathrm{H}_0 : \delta = 0$, and some other "classical" SV specifications. If we write $h_t = h(\sigma_t^2, \delta)$, then the N-SV model can be rewritten as

$$X_t = [g(h_t, \delta)]^{1/2} e_t, \tag{10}$$

$$h_t = \mu + \phi(h_{t-1} - \mu) + \sigma v_t, \tag{11}$$

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

$$g(h_t, \delta) = \begin{cases} (1 + \delta h_t)^{1/\delta} & \text{if } \delta \neq 0, \\ \exp(h_t) & \text{if } \delta = 0. \end{cases} \tag{12}$$

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

Table 1
Alternative SV models and parameter relationship

	Models	δ	μ	ϕ
Taylor (1982), Wiggins (1987), Chesney and Scott (1989), Jacquier et al. (1994), Harvey et al. (1998), Kim et al. (1998) and Scott (1987)	$\ln \sigma_t^2 = \mu + \phi(\ln \sigma_{t-1}^2 - \mu) + \sigma v_t$	0		
Scott (1987), Andersen (1994) and Stein and Stein (1991)	$\sigma_t = \mu + \phi(\sigma_{t-1} - \mu) + \sigma v_t$	0.5		
Heston (1993)	$\sigma_t = \phi\sigma_{t-1} + \sigma v_t$	0.5	0	
Hull and White (1987) and 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	$\frac{(\sigma_t^2)^\delta - 1}{\delta} = \mu + \phi\left[\frac{(\sigma_{t-1}^2)^\delta - 1}{\delta} - \mu\right] + \sigma v_t$			

The idea of our N-SV models is similar to that of the nonlinear ARCH (NARCH) model proposed by Higgins and Bera (1992). Obviously, our model provides an SV generalization of a nonlinear GARCH(1,1) model. Similar to the NARCH model, the proposed N-SV model can be used to test the nested models based on one parameter. However, compared with the NARCH model, our N-SV model is closely related to the option pricing literature, because the nested models have been used for pricing options.

It can be seen that 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 LN SV model as a special case. If $\delta = 1$, the variance equation (8) becomes

$$\sigma_t^2 = \mu' + \phi(\sigma_{t-1}^2 - \mu') + \sigma v_t, \tag{13}$$

where $\mu' = \mu + 1$. This is a polynomial SV model in Andersen (1994). According to this specification, the volatility has a marginal normal distribution. If $\delta = 0.5$, the variance equation (8) becomes

$$\sigma_t = \mu'' + \phi(\sigma_{t-1} - \mu'') + 0.5\sigma v_t, \tag{14}$$

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 the volatility is Gaussian.

Table 1 summarizes 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. Some specifications in Table 1 may be different from the actual specifications given in the original papers. However, they are equivalent to each other via Ito's lemma. For example, Heston (1993) adopted a square root specification for σ_t^2 which is identical to assuming σ_t follows a particular OU process. It can be seen that all these models can be obtained from our model by placing appropriate restrictions on the three parameters δ, μ and ϕ. In fact, all the models except ours require δ to be 0, 0.5, or 1. 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.

The Box–Cox transformation has been applied in various areas in finance. The most relevant application to our work is perhaps that made by Higgins and Bera (1992) for reasons mentioned above. Also relevant to our work is made by Hentschel (1995) where a family of GARCH models is introduced by applying the Box–Cox transformation to the conditional SD. 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 the 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 model. It is easy to see that h_t is stationary and ergodic if $\phi < 1$ and that if so

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

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. 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. Our experience suggests that, however, 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. This finding is similar to what had been found by Stein and Stein (1991). 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. As an alternative, Yu and Yang (2006) approximated the distribution of $u_t = \sigma_t^2 = (1 + \delta h_t)^{1/\delta}$ by a generalized LN distribution. This alternative specification gives rise to analytical expressions for model moments and can be thought of nesting the standard SV models in approximation.

To conclude this section, we attempt to offer a heuristic interpretation of δ from a finance perspective which is analogous to the introduction of continuously compounded returns. For ease of interpretation, we restrict ourselves to the range of positive δ. Define $m = 1/\delta$ and re-write the inverse Box–Cox transformation as

$$\sigma_t^2 = \left(1 + \frac{h_t}{m}\right)^m = \prod_{i=1}^{m} (1 + h_{it}), \tag{15}$$

where $\{h_{it}\}$ can be understood as a sequence of intra-day volatility movements. From a market microstructure perspective, intra-day volatility movements are caused primarily by the arrival of new information. From (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 LN 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 finite positive value of δ, say $\delta = 0.25$, on average new information arrives at the market 4 times per day.

3. Estimation and inference using MCMC

In this section, we develop a likelihood-based technique for model estimation and inference using MCMC. The literature on estimating SV models is vast. Broto and Ruiz (2002) provided a recent survey on numerous estimating techniques for SV models. The choice of MCMC is mainly due to the good finite sample performance of MCMC in the context of SV models. See, for example, Jacquier et al. (1994), Kim et al. (1998), and Meyer and Yu (2000).

As volatility in SV models is latent, the calculation of the likelihood function requires integrating out the latent variables, which in turn make the direct likelihood-based analysis numerically difficult. Let $\mathbf{X} = (X_1, X_2, \ldots, X_T)$ be the vector of observations and $f(\mathbf{X}|\boldsymbol{\theta})$ the likelihood function. To circumvent the problem caused by the latent process, a common practice is to augment the parameter vector to $(\theta, \mathbf{h})$, where $\mathbf{h} = (h_1, h_2, \ldots, h_T)$. Given a set of priors, we can obtain the joint posterior, denoted as $f(\theta, \mathbf{h}|\mathbf{X})$, based on the likelihood of the augmented parameter vector. The sequence of sampled augmented parameter vector forms a Markov chain whose stationary transition density is the same as the joint posterior. When the simulated chain converges, the chain of simulated values is regarded as a sample obtained from the joint posterior and hence can be used for statistical inferences.

3.1. Estimating the N-SV model

Assume that the priors of the 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 inverted gamma distribution. The joint posterior density for the parameters and the latent volatilities is

$$
\begin{aligned}
f(\theta,\mathbf{h}|\mathbf{X}) &= \text{prior}(\theta)\times p(h_1|\theta)\times\prod_{t=2}^{T}p(h_t|h_{t-1},\theta)\times\prod_{t=1}^{T}p(X_t|h_t,\theta)\\
&\propto (1+\phi)^{\omega-0.5}(1-\phi)^{\gamma-0.5}\exp\left\{-\frac{(\delta-\mu_\delta)^2}{2\sigma_\delta^2}\right\}\\
&\quad\times\left[\prod_{t=1}^{T}g(h_t,\delta)^{-1/2}\right]\exp\left\{-\sum_{t=1}^{T}\frac{X_t^2}{2g(h_t,\delta)}\right\}\left[\frac{1}{\sigma^2}\right]^{((T+p)/2)+1}\\
&\quad\times\exp\left\{-\frac{(1-\phi^2)(h_1-\mu)^2+\sum_{t=2}^{T}[(h_t-\mu)-\phi(h_{t-1}-\mu)]^2+S_\sigma}{2\sigma^2}\right\},
\end{aligned}
\tag{16}
$$

where p, S_σ, ω, γ, μ_δ and σ_δ^2 are hyperparameters to be defined by users. When σ^2 is integrated out of (16), we can obtain the logarithmic marginal posterior of $(\phi,\delta,\mu,\mathbf{h})$,

$$
\begin{aligned}
\ln f(\phi,\delta,\mu,\mathbf{h}|\mathbf{X}) &\propto (\omega-0.5)\ln(1+\phi)+(\gamma-0.5)\ln(1-\phi)\\
&\quad-\frac{(\delta-\mu_\delta)^2}{2\sigma_\delta^2}-\frac{1}{2}\sum_{t=1}^{T}\ln g(h_t,\delta)-\sum_{t=1}^{T}\frac{X_t^2}{2g(h_t,\delta)}\\
&\quad-\frac{T+p}{2}\ln\left\{\frac{(1-\phi^2)(h_1-\mu)^2+\sum_{t=2}^{T}[(h_t-\mu)-\phi(h_{t-1}-\mu)]^2+S_\sigma}{2}\right\}.
\end{aligned}
\tag{17}
$$

The sampling algorithm for the proposed model is based on (16) and (17). First, we use the random-walk Metropolis–Hastings algorithm to sample (ϕ,δ) simultaneously, given all the other parameters and latent volatilities. Second, we have found that the posterior of μ, which is conditional on all the other parameters and latent volatilities, is Gaussian with mean and variance defined by

$$
\begin{cases}
\hat{\mu}^*=\hat{\sigma}_\mu^2\{\frac{1-\phi^2}{\sigma^2}h_1+\frac{1-\phi}{\sigma^2}\sum_{t=2}^{T}(h_t-\phi h_{t-1})\},\\
\hat{\sigma}_\mu^2=\sigma^2\{(T-1)(1-\phi^2)+(1-\phi^2)\}^{-1}.
\end{cases}
\tag{18}
$$

Thus, μ can be sampled directly from $N(\hat{\mu}^*,\hat{\sigma}_\mu^2)$. Third, we sample σ^2 directly from its conditional posterior,

$$
\sigma^2\sim IG\left(\frac{T+p}{2},\frac{1}{2}\left[(1-\phi^2)(h_1-\mu)^2+\sum_{t=2}^{T}[(h_t-\mu)-\phi(h_{t-1}-\mu)]^2+S_\sigma\right]\right).
\tag{19}
$$

Finally, we sample each component of $\mathbf{h}$ sequentially, where the random-walk Metropolis–Hastings algorithm is employed to update each component of $\mathbf{h}$. Hence our sampling algorithm is summarized as follows:

1. Initialize θ and $\mathbf{h}$.
2. Sample ϕ and δ based on (17), given all the other parameters and $\mathbf{h}$.
3. Sample each component of $\mathbf{h}$ sequentially according to (16), given $\boldsymbol{\theta}$.
4. Sample σ^2 from (19), given all the other parameters and $\mathbf{h}$.
5. Sample μ from its conditional posterior, given σ^2, ϕ and $\mathbf{h}$.
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 number of iterations after the burn-in period.

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 (see, for example, Chib and Greenberg, 1995). Second, when we calculate the acceptance probability to update a component of $\mathbf{h}$, say h_t, the conditional posterior of h_t, given the parameters and the other components of $\mathbf{h}$, is (ignoring the end conditions to save space)

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

for $\delta \neq 0$, where $\mathbf{h}_{\backslash t}$ denotes $\mathbf{h}$ with h_t deleted. When $\delta = 0$, the conditional posterior of h_t becomes

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

Hence, h_t can be sampled by using the random-walk Metropolis–Hastings algorithm, where the acceptance probability is computed based on the above two equations.

Following Meyer and Yu (2000), we use the convergence checking criteria available in the CODA software to check whether convergence has been achieved, as well as the integrated autocorrelation time (IACT) to measure the simulation inefficiency. All the reported results in this paper are based on samples, which have passed the Heidelberger and Welch convergence test for all parameters.

3.2. Volatility estimate, likelihood evaluation and likelihood ratio test

As MCMC methods provide samples from the joint posterior distribution of all the parameters and latent volatilities, a natural way for estimating volatility is to integrate out the model parameters from the posterior. This Bayesian approach was suggested in Jacquier et al. (1994). Alternatively, one can make use of the so-called *particle filter* techniques. Particle filter is a class of simulation-based methods developed in recent statistics literature for filtering nonlinear non-Gaussian state space models (see, for example, Gordon et al., 1993; Kitagawa, 1996; Pitt and Shephard, 1999). With the particle filtering algorithm, one can conduct diagnostic checking to seek for evidence on what is wrong with the model, as well as to evaluate the likelihood of the model at the posterior mean. In this paper, we follow Berg et al. (2004) and employ Kitagawa's filtering algorithm with 50,000 particle points.

Once likelihood is evaluated at the posterior mean, one can make a statistical comparison between the proposed N-SV model and a standard SV model. When the N-SV model nests a standard SV model under consideration, 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 the comparison between non-nested models, one may use the non-nested likelihood ratio test developed by Atkinson (1986) for classical inferences. Bayesian methods are also available for the case of non-nested model comparisons, e.g., Bayes factors of Chib (1995) for Bayesian inference when the prior is proper and the deviance information criterion (Spiegelhalter et al., 2002; Berg et al., 2004) when the prior is not necessarily proper. In this paper, we use the likelihood ratio test.

3.3. Simulation studies

To check the reliability of the proposed MCMC algorithm for model estimation and comparison, we apply our algorithm to a simulated data set. We generate 2000 observations from the N-SV model using the following parameter values: $\mu = -0.2$, $\sigma = 0.2$, $\phi = 0.95$ and $\delta = 0.2$. This parameter setting is selected to be empirically realistic for daily exchange rates.

In both the simulation study and the empirical study given latter, we estimate the N-SV model using the proposed MCMC algorithm. For comparison purposes, we also estimate the LN SV model by using the all purpose Bayesian software package BUGS based on the single-move Gibbs sampler for the ease of implementation (see, for example, Meyer and Yu, 2000). In both samplers, we choose a burn-in period of 50,000 iterations and a follow-up period of 500,000 iterations, and retain 1 draw for every 50 draws. The MCMC sampler is initialized by setting $\phi = 0.95$, $\sigma^2 = 0.02$

Table 2
Simulation results

	True value	N-SV					LN SV			
		Mean	SD	90% CI	MC SE	IACT	Mean	SD	MC SE	IACT
ϕ	0.95	0.9564	0.0121	(0.9348, 0.9741)	0.00019	121.9	0.9598	0.0120	0.00050	883.7
σ	0.2	0.1893	0.0261	(0.15, 0.2359)	0.00048	169.1	0.1924	0.0269	0.00138	1319.7
μ	−0.2	−0.2105	0.1144	(−0.3968, −0.0236)	0.00091	31.5	−0.2137	0.1256	0.00269	229.0
δ	0.2	0.2105	0.1444	(0.0011, 0.4355)	0.00250	149.9				
Loglik.				−2657.346				−2658.990		
LR stat					3.287					
p-value					0.0698					

and $\mu = 0$ for the LN SV model and arbitrarily initialized for the N-SV model. The same prior distributions are used for the common parameters in both models.The only exception is for μ. In the LN SV model, we choose an informative but reasonably flat prior distribution for μ, such as a normal density with mean 0 and variance 25, while in the N-SV model we use a diffuse prior for the reason argued above. The hyperparameters are, respectively, $p = 10.0$, $\omega = 20.0$, $\gamma = 1.5$, $S_\sigma = 0.1$, $\mu_\delta = 0.2$ and $\sigma_\delta^2 = 1$.

Table 2 summarizes the results from estimation and model comparison, including the posterior means, SDs, 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 LN SV model against the N-SV model. For the N-SV model we also report the 90% Bayesian credible (highest probability) intervals for all the parameters.

First, it can be seen that the proposed MCMC procedure can reliably estimate all the parameters in the N-SV model, including the key parameter, δ. Second, the 90% Bayesian credible interval for δ includes the true value but not 0. Although not reported, we find that even 99% Bayesian credible interval of δ does not include 0.5 or 1. The likelihood ratio statistic favors the true specification and shows evidence against the standard SV model. Third, the 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 has been achieved in the N-SV model.

4. Empirical results for exchange rates

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). In this section, we estimate the proposed model using daily returns of the dollar/pound exchange rate for the period from January 1, 1986 to December 31, 1998. The data set is available from the H-10 Federal Reserve Statistical Release. We use the mean-corrected and variance-scaled returns defined by

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

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

Fig. 1 displays the adjusted return series. Table 3 summarizes the empirical results, including the posterior means, SDs, 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 LN SV model against the N-SV model. For the N-SV model we also report the 90% Bayesian credible intervals for all the parameters.

A few results emerge from Table 3. First, the posterior mean of δ in the proposed N-SV model is 0.172 and the 90% Bayesian credible interval does not include 0. Although not reported, we find that even 99% Bayesian credible interval of δ does not include 0.5 or 1. Thus, data provide significant evidence against the LN SV model and highly significant evidence against all the other SV models, including the Stein-Stein and Heston specifications. Although all the standard SV models are rejected, the posterior quantities of δ seem to suggest that the LN 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 LN model, which is suggestive of a high persistency of volatility. In the proposed N-SV model, ϕ remains at a similar level. In fact all the estimated parameters have similar magnitudes and similar SDs across both models.

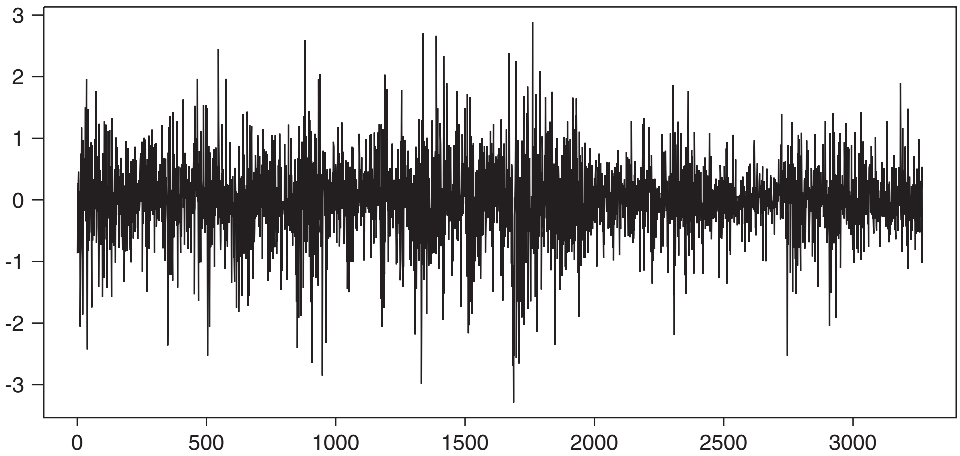

Fig. 1. Time series plots for dollar/pound exchange rate return.

Table 3
Empirical results

	N-SV					LN SV			
	Mean	SD	90% CI	MC SE	IACT	Mean	SD	MC SE	IACT
ϕ	0.9595	0.0101	(0.9417, 0.9745)	0.00017	138.3	0.9676	0.0091	.00026	408.2
σ	0.2066	0.0269	(0.1672, 0.2543)	0.00050	174.9	0.1873	0.0268	0.00090	568.1
μ	−0.2244	0.1044	(−0.3913, −0.0495)	0.00087	35.0	−0.2579	0.1095	0.00103	44.1
δ	0.1716	0.1203	(0.0039, 0.3684)	0.00214	189.0				
Loglik.			−4369.792				−4371.606		
LR Stat				3.628					
p-value				0.0568					

Third, the likelihood ratio statistic and the associated p value suggest that the LN model is rejected at the 10% 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 LN model. Fifth, compared with other parameters, δ appears to be more difficult to estimate and has the largest value of SD. Finally, according to our interpretation of δ, for the dollar/pound exchange rate data, on average new information arrives at the market about six times per day.

To provide diagnostic checks for the observed series and two SV models, we follow Kim et al. (1998) and compute the forecast uniforms from one-step-ahead forecasts for both models. Fig. 2 gives the QQ-plot of the normalized innovations obtained from the LN model and N-SV model, respectively. The plot suggests that there are more outliers in the normalized innovations that the LN SV cannot explain than the N-SV model. Similar to Kim et al. (1998), we find that these outliers 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 does the LN SV model in this case.

As argued in Section 2, a by-product of the new way to model volatility is that the marginal distribution of volatility is obtained. The marginal distributions of volatility implied from the estimated LN and N-SV models are plotted in Fig. 3, where the solid line is for the LN SV model and hence is the density function of a LN 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 LN SV model. This 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 and compare two filtered volatilities. To conserve space, we do not plot them but merely summarize the results. In general, the two filtered volatilities are very close to each other when volatility is high. When volatility is high, the differences become large. Moreover, we find

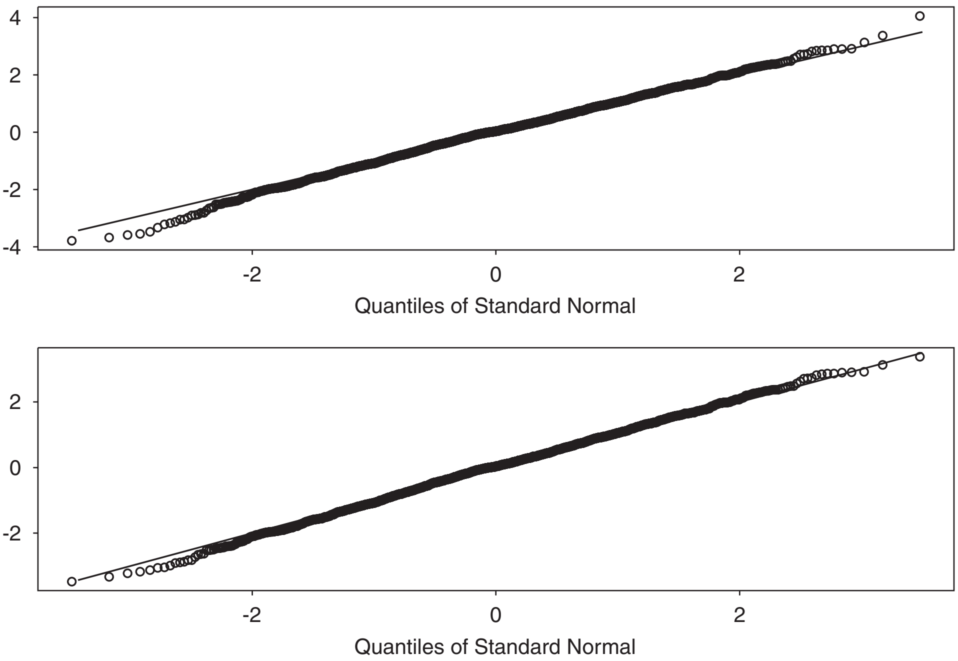

Fig. 2. Diagnostic checks of two SV models for daily returns of the dollar/pound exchange rate. The first panel is the QQ-plot of the normalized residuals from the LN SV model; the second panel is the QQ-plot of the normalized residuals from the N-SV model.

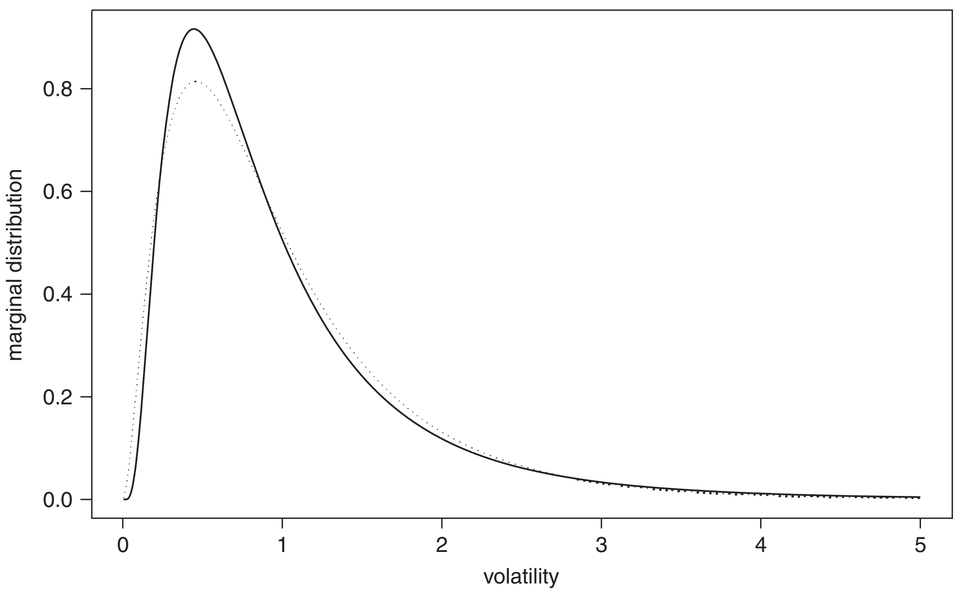

Fig. 3. Marginal densities of the dollar/pound exchange rate volatility implied from the LN SV model and the N-SV model. The solid line is for the LN SV model; the dotted line is for the N-SV model.

that the two filtered volatilities have similar sample means (0.995 versus 1.004) but rather different sample variances (0.3297 for the N-SV model versus 0.3782 for the LN SV model), 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 for option pricing.

5. Implications for option pricing

The most important application of the SV model is probably the pricing of options. Under a set of assumptions, Hull and White (1987) showed that the value of an 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 a characteristic function approach, Heston (1993) derived 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. Hull and White (1987) outlined an efficient procedure for conducting Monte Carlo simulations to calculate an European call option on stocks.

To examine the economic importance of our N-SV model on option pricing, we price options using both the LN SV and N-SV models, provided that the true model is the estimated N-SV model. To price options, we follow Mahieu and Schotman (1998).

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 $r_\mathrm{d} - r_\mathrm{f}$. Under the same set of assumptions as in Hull and White (1987), it can be shown that

$$C = \mathrm{e}^{-\tau r_d} \int_0^\infty \mathrm{BS}(w_\tau)\, p\, \mathrm{d} f(w_\tau | h_0)\, \mathrm{d}w_\tau, \tag{20}$$

where w_τ^2 is given by

$$w_\tau^2 = \int_0^\tau g(h_s, \delta)\, \mathrm{d}s, \tag{21}$$

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

$$\mathrm{BS}(w_\tau) = F_0 \mathrm{N}(d_1) - X \mathrm{N}(d_2), \tag{22}$$

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

$$d_1 = \frac{\ln(F_0/X) + w_\tau^2}{w_\tau} \tag{23}$$

and

$$d_2 = d_1 - w_\tau. \tag{24}$$

In discrete time setup one to approximate w_τ^2. In this paper, we follow the suggestion of Amin and Ng (1993) and approximate w_τ^2 by

$$w_\tau^2 \approx \sum_{t=1}^{n} g(h_i, \delta), \tag{25}$$

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 v_i for $1 \leqslant i \leqslant n$.
3. Generate h_i according to

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

4. Calculate w_τ^2 using Eq. (25).

Table 4
Comparison of call option prices on currency

S_0/X	LN SV Option price	N-SV Option price	Percentage 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

5. Calculate $\mathrm{BS}(w_\tau)$ using Eq. (22) and call it p_1.
6. Repeat Steps 3–5 using $\{-v_i\}$ 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 to give a sequence of y values.
9. Calculate the mean of y's and this is the estimate of the option price.

Our algorithm is closely related to the one given by Mahieu and Schotman (1998) but differs from theirs in two aspects. First, we use an antithetic method in Step 6 to reduce the variance of simulation errors. Second, we use a much larger value of K (10,000 as opposed to 500) to ensure small approximation errors in calculating Eq. (20).

The algorithm is then applied to pricing a half-year call option based on the LN SV and N-SV models with the estimated parameter values in Table 2 imposed. Since the parameter estimates reported in Table 2 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. In both models, we choose $n = 126$, $S_0 = 1.5$, $r_\mathrm{d} = 0$, $r_\mathrm{f} = 0$, $K = 10,000$, $\sigma_0 = 0.006349$ 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. The 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. 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 LN SV model tends to overprice the options. In fact the N-SV option price is always no bigger than the LN 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 LN 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.

6. Conclusions and extensions

In this paper, a class of nonlinear SV (N-SV) models has been proposed. The new class facilitates comparing and testing all standard parametric SV models. Since these alternative parametric SV models coexist in the literature, our approach is useful in the sense that it can provide evidence to support or against some of the classical specifications. 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. An empirical application is given using the daily dollar/pound exchange rate series. Empirical results show that all the standard SV models are rejected and hence provide evidence of nonlinear stochastic volatility. Furthermore, model diagnostics indicate that without sacrificing the overall goodness of fit, the N-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 LN models tend to overprice out-of-money options. The deeper the out-of-money options, the larger the percentage bias.

Although in this paper we focused on one volatility factor which is free of jumps, there are some possible extensions to our work. One possible extension is to use the suggested methodology to analyze stock data. However, as stock data often display a strong feature of asymmetric volatility, together with a higher kurtosis than implied by the standard SV models, one has to incorporate a leverage effect and a fat-tailed error distribution into the N-SV model. Also, as equity data often have more than one volatility factors (Gallant and Tauchen, 2001), one needs to apply the Box–Cox transformation to all the factors. Other interesting extensions would be to incorporate jumps and long memory volatility into the model (see, for example, Duffie et al., 2000; Breidt et al., 1998). Furthermore, although we present our model and estimation method in terms of discrete-time SV models, one can estimate the continuous-time N-SV models using alternative estimation methods. Finally, it would be interesting to evaluate the out-of-sample forecasting performances of the N-SV model relative to standard SV models.

Acknowledgments

We would like to thank Henk Berkman, Charles Corrado, Andrew Harvey, Steve Heston, Genshiro Kitagawa, John Knight, Michael McAleer, Esther Ruiz, Steve Satchell, Neil Shephard, Yiu Kuen Tse and Shu Yan, and seminar participants at the University of Auckland (both Department of Economics and Department of Accounting and Finance), Yale University, Econometric Society Meeting in Brisbane, and New Zealand Econometric Society Meeting, New Zealand Finance Colloquium, and Canadian Econometrics Study Group Meeting for comments and suggestions. Jun Yu gratefully acknowledges financial support from the Royal Society of New Zealand Marsden Fund under number 01-UOA-015. This research was started while the Jun Yu was visiting National University of Singapore in 2001 and he would like to thank Department of Economics at National University of Singapore for hospitality. Thanks also go to the Victorian Partnership for Advanced Computing (VPAC) Calvin Chan and Shirley Huang for computing support.

References

Amin, K.I., Ng, V.K., 1993. Jump diffusion option valuation in discrete time. J. Finance 48, 881–909.

Andersen, T., 1994. Stochastic autoregressive volatility: a framework for volatility modeling. Math. Finance 4, 75–102.

Atkinson, A.C., 1986. Monte Carlo tests of separate families of hypothesis. Unpublished Paper, Imperial College, London.

Barndorff-Nielsen, N., Shephard, N., 2001. Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics. J. Roy. Statist. Soc. B 63, 167–241.

Berg, A., Meyer, R., Yu, J., 2004. Deviance information criterion for comparing stochastic volatility models. J. Business Econom. Statist. 22, 107–120.

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. J. Econometrics 31, 307–327.

Box, G.E.P., Cox, D.R., 1964. An analysis of transformations. J. Roy. Statist. Soc. B 26, 211–243.

Breidt, F.J., Crato, N., De Lima, P., 1998. The detection and estimation of long memory in stochastic volatility. J. Econometrics 83, 325–348.

Broto, C., Ruiz, E., 2002. Estimation methods for stochastic volatility models: a survey. Working Paper 02-54, Universidad Carlos III de Madrid.

Chesney, M., Scott, L.O., 1989. Pricing European currency options: a comparison of the modified Black-Scholes model and a random variance model. J. Finan. Quant. Anal. 24, 267–284.

Chib, S., 1995. Marginal likelihood from the Gibbs output. J. Amer. Statist. Assoc. 90, 1313–1321.

Chib, S., Greenberg, E., 1995. Understanding the Metropolis–Hastings algorithm. The Amer. Statist. 49, 327–335.

Clark, P.K., 1973. A subordinated stochastic process model with finite variance for speculative price. Econometrica 41, 135–155.

Duan, J., 1997. Augmented GARCH(p, q) process and its diffusion limit. J. Econometrics 79, 97–127.

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

Gallant, A.R., Tauchen, G., 2001. Efficient method of moments. Working paper, Department of Economics, University of North Carolina.

Gordon, N.J., Salmond, D.J., Smith, A.F.M., 1993. A novel approach to nonlinear and non-Gaussian Bayesian state estimation. IEE Proc. F 140, 107–133.

Harvey, A.C., Ruiz, E., Shephard, N., 1994. Multivariate stochastic variance models. Rev. Econom. Stud. 61, 247–264.

Hentschel, L., 1995. All in the family: nesting symmetric and asymmetric GARCH models. J. Finan. Econom. 39, 71–104.

Heston, S.L., 1993. A closed-form solution for options with stochastic volatility, with application to bond and currency options. Rev. Finan. Stud. 6, 327–343.

Higgins, M.L., Bera, A., 1992. A class of nonlinear ARCH models. Internat. Econom. Rev. 33, 137–158.

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

Jacquier, E., Polson, N.G., Rossi, P.E., 1994. Bayesian analysis of stochastic volatility models. J. Business Econom. Statist. 12, 371–389.

Johnson, H., Shanno, D., 1987. Option pricing when the variance is changing. J. Finan. Quant. Anal. 22, 143–152.

Jones, C., 2003. The dynamics of stochastic volatility models: evidence from underlying and option markets. J. Econometrics 116, 181–224.

Kim, S., Shephard, N., Chib, S., 1998. Stochastic volatility: likelihood inference and comparison with ARCH models. Rev. Econom. Stud. 65, 361–393.

Kitagawa, G., 1996. Monte Carlo filter and smoother for Gaussian nonlinear state space models. J. Comput. Graphical Statist. 5, 1–25.

Mahieu, R.J., Schotman, P.C., 1998. An empirical application of stochastic volatility models. J. Appl. Econom. 13, 330–360.

Meddahi, N., 2001. An eigenfunction approach for volatility modeling. Mimeo, Department of Economics, University of Montreal.

Melino, A., Turnbull, S.M., 1990. Pricing foreign currency options with stochastic volatility. J. Econometrics 45, 239–265.

Meyer, R., Yu, J., 2000. BUGS for a Bayesian analysis of stochastic volatility models. Econometrics J. 3, 198–215.

Nelson, D., 1990. ARCH models as diffusion approximations. J. Econometrics 45, 7–38.

Pitt, M., Shephard, N., 1999. Filtering via simulation: auxiliary particle filter. J. Amer. Statist. Assoc. 94, 590–599.

Scott, L.O., 1987. Option pricing when the variance changes randomly: theory, estimation and an application. J. Finan. Quant. Anal. 22, 419–439.

Spiegelhalter, D.J., Best, N.G., Carlin, B.P., van der Linde, A., 2002. Bayesian measures of model complexity and fit (with discussion). J. Roy. Statist. Soc. B 64, 583–639.

Stein, E.M., Stein, J.C., 1991. Stock price distributions with stochastic volatility: an analytical approach. Rev. Finan. Stud. 4, 727–752.

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. (Ed.), Time Series Analysis: Theory and Practice, vol. 1. North-Holland, Amsterdam, pp. 203–226.

Wiggins, J.B., 1987. Option values under stochastic volatility: theory and empirical estimate. J. Finan. Econom. 19, 351–372.

Yang, Z., 1999. Estimating a transformation and its effect on Box–Cox T-ratio. Test 8, 167–190.

Yu, J., 2002. Forecasting volatility in the New Zealand stock market. Appl. Finan. Econom. 12, 193–202.

Yu, J., 2005. On leverage in a stochastic volatility model. J. Econom. 127, 165–178.

Yu, J., Yang, Z., 2006. A class of nonlinear stochastic volatility models. Fifth International Conference on Computational Intelligence in Economics and Finance, forthcoming.